

GATE Bibliography

The GATE Team

August 16, 2023

References

- [Abdel-Hafez *et al.* 14]
A. Abdel-Hafez, Q. V. Phung, and Y. Xu. Utilizing voting systems for ranking user tweets. In *Proceedings of the 2014 Recommender Systems Challenge, RecSysChallenge '14*, pages 23:23–23:28, New York, NY, USA, 2014. ACM.
- [Abecker *et al.* 03]
A. Abecker, S. Tabor, R. Traphöner, J. Franz, W. Maas, and W. Eickhoff. The INKASS Information Ontology for Knowledge Asset Trading. In *submitted to Wirtschaftsinformatik 2003*, Dresden, Germany, 2003.
- [Abekawa *et al.* 10]
T. Abekawa, M. Utiyama, E. Sumita, and K. Kageura. Community-based construction of draft and final translation corpus through a translation hosting site minna no hon'yaku (mnh), 2010.
- [Abel *et al.* 11a]
F. Abel, I. Celik, G.-J. Houben, and P. Siehdnel. Leveraging the semantics of tweets for adaptive faceted search on Twitter. In *Proceedings of the 10th international conference on The semantic web - Volume Part I, ISWC'11*, pages 1–17, Berlin, Heidelberg, 2011. Springer-Verlag.
- [Abel *et al.* 11b]
F. Abel, Q. Gao, G. J. Houben, and K. Tao. Semantic enrichment of Twitter posts for user profile construction on the social web. In *ESWC (2)*, pages 375–389, 2011.
- [Abel *et al.* 11c]
F. Abel, Q. Gao, G.-J. Houben, and K. Tao. Analyzing temporal dynamics in twitter profiles for personalized recommendations in the social web. In *Proceedings of the 3rd International Web Science Conference, WebSci '11*, pages 2:1–2:8, New York, NY, USA, 2011. ACM.
- [Abelson *et al.* 85]
H. Abelson, G. Sussman, and J. Sussman. *The Structure and Interpretation of Computer Programs*. MIT Press, Cambridge, Mass., 1985.
- [Aberdeen *et al.* 95]
J. Aberdeen, J. Burger, D. Day, L. Hirschman, P. Robinson, and M. Vilain. MITRE: Description of the *Alembic* System Used for MUC-6. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*. Morgan Kaufmann, California, 1995.
- [Abney 96]
S. Abney. Partial parsing via finite state cascades. *Natural Language Engineering*, 2(4):337–344, 1996.

- [Abney 02]
S. Abney. Bootstrapping. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics: Proceedings of the Conference*, 2002.
- [ACE04a]
Annotation Guidelines for Entity Detection and Tracking (EDT), Feb 2004. Available at <http://www ldc.upenn.edu/Projects/ACE/>.
- [ACE04b]
Annotation Guidelines for Event Detection and Characterization (EDC), Feb 2004. Available at <http://www ldc.upenn.edu/Projects/ACE/>.
- [ACE04c]
Annotation Guidelines for Relation Detection and Characterization (RDC), Feb 2004. Available at <http://www ldc.upenn.edu/Projects/ACE/>.
- [Ackely *et al.* 85]
D. Ackely, G. Hinton, and T. Sejnowski. A Learning Algorithm for Boltzmann Machines. *Cognitive Science* 9, pages 147–169, 1985.
- [Ackoff 89]
R. L. Ackoff. From data to wisdom. *Journal of Applied System Analysis*, 16:3–9, 1989.
- [ACL 97]
ACL. *Proceedings of the Fifth Conference on Applied Natural Language Processing (ANLP-97)*. Association for Computational Linguistics, 1997.
- [Adamic *et al.* 16]
L. A. Adamic, T. M. Lento, E. Adar, and P. C. Ng. Information evolution in social networks. In *WSDM*, pages 473–482. ACM, 2016.
- [Adams 82]
D. Adams. *Life, the Universe and Everything*. Pan, 1982.
- [Adams *et al.* 11]
B. Adams, D. Phung, and S. Venkatesh. Eventscapes: Visualizing events over time with emotive facets. In *Proceedings of the 19th ACM International Conference on Multimedia*, pages 1477–1480, 2011.
- [address= Dordrecht Kamp & Reyle 93]
H. address = Dordrecht Kamp and U. Reyle. *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Kluwer, 1993.
- [Adida *et al.* 08]
B. Adida, M. Birbeck, S. McCarron, and S. Pemberton. Rdfa in xhtml: Syntax and processing. Technical Report www.w3.org/TR/2008/REC-rdfa-syntax-20081014/, W3C, 2008.
- [Adriaens & Schreurs]
G. Adriaens and D. Schreurs. From COGRAM to ALCOGRAM: Toward a controlled English grammar checker. In *Conference on Computational Linguistics (COLING'92)*, pages 595–601, Nantes, France.
- [Agarwal & Searls 08]
P. Agarwal and D. B. Searls. Literature mining in support of drug discovery. *Briefings in Bioinformatics*, 9(6):479–492, 2008.

- [Agarwal *et al.* 14]
D. Agarwal, B. Chen, R. Gupta, J. Hartman, Q. He, A. Iyer, S. Kolar, Y. Ma, P. Shivaswamy, A. Singh, and L. Zhang. Activity ranking in linkedin feed. In S. A. Macskassy, C. Perlich, J. Leskovec, W. Wang, and R. Ghani, editors, *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '14, New York, NY, USA - August 24 - 27, 2014*, pages 1603–1612. ACM, 2014.
- [Agatonovic *et al.* 08a]
M. Agatonovic, N. Aswani, K. Bontcheva, H. Cunningham, T. Heitz, Y. Li, I. Roberts, and V. Tablan. Large-scale, parallel automatic patent annotation. In *Proceedings of the 1st ACM workshop on Patent information retrieval (PaIR '08, 30 October 2008, PaIR '08*, pages 1–8, New York, NY, USA, October 2008. ACM.
- [Agatonovic *et al.* 08b]
M. Agatonovic, N. Aswani, K. Bontcheva, H. Cunningham, T. Heitz, Y. Li, I. Roberts, and V. Tablan. Large-scale, parallel automatic patent annotation. In *Proceedings of the 1st ACM workshop on Patent information retrieval (PaIR '08)*, pages 1–8, 2008.
- [Agency 11]
H. E. S. Agency. Statistics - Students and qualifiers at UK HE institutions. <http://www.hesa.ac.uk/content/view/1897/239/>, 2011.
- [Agerwalak & Arvind 82]
T. Agerwalak and Arvind. Data flow systems: Guest editor's introduction. *IEEE Computer*, 15:10–13, 1982.
- [Agirre & Rigau 96]
E. Agirre and G. Rigau. Word sense disambiguation using conceptual density. In *Proc. of 16th International Conference on Computational Linguistics*, volume 1, pages 16–23, Copenhagen, Denmark, 1996.
- [Agirre *et al.* 09]
E. Agirre, A. X. Chang, D. S. Jurafsky, C. D. Manning, V. I. Spitzkovsky, and E. Yeh. Stanford-ubc at tac-kbp. In *Proceedings of the Second Text Analysis Conference (TAC 2009)*, Gaithersburg, Maryland, USA, November 2009.
- [Agrawal & An 12]
A. Agrawal and A. An. Unsupervised emotion detection from text using semantic and syntactic relations. In *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology - Volume 01, WI-IAT '12*, pages 346–353, Washington, DC, USA, 2012. IEEE Computer Society.
- [Agrawal & Siddiqui 09]
S. Agrawal and T. j. Siddiqui. Using syntactic and contextual information for sentiment polarity analysis. In *ICIS '09: Proceedings of the 2nd International Conference on Interaction Sciences*, pages 620–623, New York, NY, USA, 2009. ACM.
- [Aguado *et al.* 98]
G. Aguado, A. Bañón, J. A. Bateman, S. Bernardos, M. Fernández, A. Gómez-Pérez, E. Nieto, A. Olalla, R. Plaza, and A. Sánchez. ONTOGENERATION: Reusing domain and linguistic ontologies for Spanish text generation. In *Workshop on Applications of Ontologies and Problem Solving Methods, ECAI'98*, 1998.
- [Ahmad & Fulford 92]
K. Ahmad and H. Fulford. Knowledge processing 4: Semantic relations and their use in elaborating terminology. Technical Report CS Report No. CS-92-07, University of Surrey, Guildford, UK, 1992.

- [Ahmad & Gillam 05]
K. Ahmad and L. Gillam. Automatic ontology extraction from unstructured texts. In *On the Move to Meaningful Internet Systems 2005: CoopIS, DOA, and ODBASE*, volume 3761/2005, pages 1330–1346. Springer Berlin / Heidelberg, 2005.
- [Ahmad *et al.* 06]
K. Ahmad, L. Gillam, and D. Cheng. Sentiments on a grid: Analysis of streaming news and views. In *5th Language Resources and Evaluation Conference*, 2006.
- [Aho & Ullman 72]
A. V. Aho and J. D. Ullman. *The Theory of Parsing, Translating and Compiling. Vol 1 : Parsing*. Prentice-Hall, Englewood Cliffs, N.J., 1972.
- [Aho *et al.* 86]
A. V. Aho, R. Sethi, and J. D. Ullman. *Compilers Principles, Techniques, and Tools*. Addison-Wesley, Reading, Massachusetts, 1986.
- [Ainsworth 88]
W. Ainsworth. *Speech Recognition by Machine*. Peter Peregrinus / IEE, London, 1988.
- [Aitchison 94]
J. Aitchison. *Words in the Mind: an introduction to the mental lexicon*. Blackwell, Oxford, 1994.
- [Aker *et al.* 17a]
A. Aker, L. Derczynski, and K. Bontcheva. Simple open stance classification for rumour analysis. *CoRR*, abs/1708.05286, 2017.
- [Aker *et al.* 17b]
A. Aker, J. Petrak, and F. Sabbah. An extensible multilingual open source lemmatizer. In *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017*, pages 40–45, Varna, Bulgaria, September 2017. INCOMA Ltd.
- [A.Kilgarrieff 03]
A.Kilgarrieff. No-bureaucracy evaluation. In *Proceedings of the EACL 2003 Workshop on Evaluation Initiatives in Natural Language Processing*, Budapest, Hungary, 2003.
- [Al-Mhabis & Cunningham 17]
N. Al-Mhabis and H. Cunningham. Socio-political perspectives on surveillance and censorship: Implications for on-line privacy in the age of cloud computing. In *IEEE Computing Conference*, July 2017.
- [Alan W. Biermanna & Sigmon 83]
B. W. B. Alan W. Biermanna and A. H. Sigmon. An Experimental Study of Natural Language Programming. *International Journal of Man-Machine Studies*, 18(1):71–87, 1983.
- [Alani *et al.* 02]
H. Alani, S. Dasmahapatra, N. Gibbins, H. Glaser, S. Harris, Y. Kalfoglou, K. O’Hara, and N. Shadbolt. Managing Reference: Ensuring Referential Integrity of Ontologies for the Semantic Web. In *13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02)*, pages 317–334, Siguenza, Spain, 2002.
- [Alani *et al.* 03]
H. Alani, S. Kim, D. Millard, M. Weal, W. Hall, P. Lewis, and N. Shadbolt. Web-based Knowledge Extraction and Consolidation for Automatic Ontology Instantiation. In *Proceedings of the Knowledge Markup and Semantic Annotation Workshop (SEMANNOT’03)*, Sanibel, Florida, 2003.

- [Albakour *et al.* 10]
M.-D. Albakour, U. Kruschwitz, and S. Lucas. Sentence-level attachment prediction. In H. Cunningham, A. Hanbury, and S. Rüger, editors, *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference)*. LNCS volume number: 6107, Lecture Notes in Computer Science, Vienna, Austria, May 2010. Springer.
- [Aleman-Meza *et al.* 06]
B. Aleman-Meza, M. Nagarajan, C. Ramakrishnan, L. Ding, P. Kolari, A. P. Sheth, I. B. Arpinar, A. Joshi, and T. Finin. Semantic analytics on social networks: experiences in addressing the problem of conflict of interest detection. In *WWW '06: Proceedings of the 15th international conference on World Wide Web*, pages 407–416, New York, NY, USA, 2006. ACM Press.
- [Allan *et al.* 98a]
J. Allan, J. Carbonell, G. Doddington, J. Yamron, and Y. Yang. Topic detection and tracking pilot study: Final report. In *Proceedings of the DARPA Broadcast News Transcription and Understanding Workshop*, pages 194–218, Lansdowne, VA, USA, Feb 1998. 007.
- [Allan *et al.* 98b]
J. Allan, R. Papka, and V. Lavrenko. On-line new event detection and tracking. In *Proceedings of the 21st SIGIR Conference on Research and Development in Information Retrieval*, pages 37–45, 1998.
- [Allen 95]
J. Allen. *Natural Language Understanding*. Benjamin/Cummings, Redwood City, CA, 2 edition, 1995.
- [Allerton 87]
D. Allerton. The linguistic and sociolinguistic status of proper names. *Journal of Linguistics*, 11(3), 1987.
- [Alleva *et al.* 93]
F. Alleva, X. Huang, and M.-Y. Hwang. An improved search algorithm using incremental knowledge for continuous speech recognition. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 2, pages 307–310, 1993.
- [Alm *et al.* 05]
C. O. Alm, D. Roth, and R. Sproat. Emotions from text: machine learning for text-based emotion prediction. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 579–586, 2005.
- [Alonso & Lease 11]
O. Alonso and M. Lease. Crowdsourcing for information retrieval: principles, methods, and applications. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, pages 1299–1300. ACM, 2011.
- [Alonso & Mizzaro 12]
O. Alonso and S. Mizzaro. Using crowdsourcing for trec relevance assessment. *Information Processing and Management*, (0):–, 2012.
- [Alonso *et al.* 13]
O. Alonso, C. C. Marshall, and M. Najork. Are some tweets more interesting than others? #hardquestion. In *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval*, HCIR '13, pages 2:1–2:10, New York, NY, USA, 2013. ACM.
- [(ALPAC) 66]
T. A. L. P. A. C. (ALPAC). Language and Machines. Computers in Translation and Linguistics. Technical Report Publication 1416, Division of Behavioural Sciences, National Academy of Sciences, National Research Council, Washington, D.C., 1966.

- [Alshawi 87]
H. Alshawi. *Memory and Context for Language Interpretation*. Cambridge University Press, Cambridge, 1987.
- [Alshawi 92]
H. Alshawi, editor. *The Core Language Engine*. MIT Press, Cambridge MA, 1992.
- [Altun *et al.* 03]
Y. Altun, I. Tsochantaridis, and T. Hofmann. Hidden Markov Support Vector Machines. In *Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003)*, Washington DC, USA, 2003.
- [Amardeilh & Damjanovic 09]
F. Amardeilh and D. Damjanovic. Du texte à la connaissance : annotation sémantique et peuplement d'ontologie appliqués à des artefacts logiciels. In *Proceedings of IC 2009, 20èmes Journées Francophones d'Ingénierie des Connaissances (20th French conference on Knowledge Engineering)*, Hammamet, Tunisia, 25-29 May 2009.
- [Amardeilh & Francart 04]
F. Amardeilh and T. Francart. A semantic web portal with hlt capabilities. *Veille Stratégique Scientifique et Technologique (VSST'04)*, 2:481–492, 2004.
- [Amardeilh 06]
F. Amardeilh. OntoPop or how to annotate documents and populate ontologies from texts. In *Proceedings of the Workshop on Mastering the Gap: From Information Extraction to Semantic Representation (ESWC'06)*, Budva, Montenegro, 2006.
- [Amardeilh 08]
F. Amardeilh. Semantic annotation and ontology population. In J. Cardoso and M. Lytras, editors, *Semantic Web Engineering in the Knowledge Society*. Idea Group Publishing, 2008.
- [Amardeilh *et al.* 07]
F. Amardeilh, B. Vatan, N. Gibbins, T. R. Payne, A. Saleh, and H. H. Wang. Sws bootstrapping methodology. Technical Report D1.2, TAO Project Deliverable, 2007. <http://www.tao-project.eu/resources/publicdeliverables/d1-2.pdf>.
- [Amardeilh *et al.* 08]
F. Amardeilh, M. Gibson, K. Bontcheva, and D. Damjanovic. Cross-media content augmentation. Technical Report D3.2, TAO Project Deliverable, 2008. <http://www.tao-project.eu/resources/publicdeliverables/d3-2-final.pdf>.
- [Amardeilh *et al.* 09]
F. Amardeilh, D. Damjanovic, and K. Bontcheva. CA Manager: a Framework for Creating Customised Workflows for Ontology Population and Semantic Annotation. In *Proceedings of the Semantic Authoring, Annotation and Knowledge Markup Workshop (SAAKM 2009) co-located with the 5th International Conference on Knowledge Capture (K-CAP 2009)*, Redondo Beach, California, USA, September 2009.
- [Ambati & Vogel 10]
V. Ambati and S. Vogel. Can crowds build parallel corpora for machine translation systems? In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 62–65, 2010.
- [Ambati *et al.* 10]
V. Ambati, S. Vogel, and J. G. Carbonell. Active learning and crowd-sourcing for machine translation. In *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10), Valletta, Malta, may 2010*, 2010.

- [Amer-Yahia & Lalmas 06]
S. Amer-Yahia and M. Lalmas. Xml search: languages, inex and scoring. *SIGMOD Rec.*, 35(4):16–23, December 2006.
- [Amer-Yahia *et al.* 04]
S. Amer-Yahia, C. Botev, and J. Shanmugasundaram. Texquery: A full-text search extension to xquery. In *Proc. of the 13th Int. Conf. on World Wide Web, WWW '04*, pages 583–594, 2004.
- [Amigó *et al.* 12]
E. Amigó, A. Corujo, J. Gonzalo, E. Meij, and M. d. Rijke. Overview of RepLab 2012: Evaluating Online Reputation Management Systems. In *CLEF 2012 Labs and Workshop Notebook Papers*, 2012.
- [Ammicht *et al.* 01]
E. Ammicht, A. Potamianos, and E. Fosler-Lussier. Ambiguity Representation and Resolution in Spoken Dialogue Systems. In *European Conference on Speech Communication and Technology (EUROSPEECH)*, volume 3, pages 2217–2220, Aalborg, Denmark, September 2001.
- [Amsler & White 79]
R. A. Amsler and J. S. White. Development of a Computational Methodology for Deriving Natural Language Semantic Structures via Analysis of Machine-readable Dictionaries. Technical Report MCS77-01315, NSF, 1979.
- [Amsler 80]
R. A. Amsler. The Structure of the Merriam-Webster Pocket Dictionary. Technical Report TR-164, University of Texas at Austin, 1980.
- [Amsler 81]
R. A. Amsler. A Taxonomy of English Nouns and Verbs. In *Proceedings of the 19th Annual Meeting of the Association for Computational Linguistics*, pages 133–138, Stanford, CA, 1981.
- [Amtrup 95]
J. Amtrup. ICE – INTARC Communication Environment User Guide and Reference Manual Version 1.4. Technical report, University of Hamburg, 1995.
- [Amtrup 99]
J. Amtrup. Architecture of the Shiraz Machine Translation System. <http://cr1.nmsu.edu/~shiraz/archi.html>, 1999.
- [Ananiadou 88]
S. Ananiadou. *Towards a methodology for automatic term recognition*. Unpublished PhD thesis, University of Manchester, UK, 1988.
- [Ananiadou 94]
S. Ananiadou. A methodology for automatic term recognition. In *Proc. of 15th International Conference on Computational Linguistics (COLING)*, Kyoto, Japan, 1994.
- [Ananiadou *et al.* 95]
S. Ananiadou, I. Radford, and J.-I. Tsujii. Sublanguage knowledge acquisition for hypertext optimisation. In *Proc. of the Natural Language Processing Pacific Rim Symposium*, Seoul. Korea, 1995.
- [Andersen *et al.* 92a]
P. M. Andersen, P. J. Hayes, A. K. Huettner, I. B. Nirenburg, L. M. Schmandt, and S. P. Weinstein. Automatic extraction of facts from press releases to generate news stories. In *Proceedings of the Third Conference on Applied Natural Language Processing*, pages 170–177. Association for Computational Linguistics, 1992.

- [Andersen *et al.* 92b]
P. Andersen, P. Hayes, A. Huettner, I. Nirenburg, L. Schmandt, and S. Weinstein. Automatic Extraction of Facts from Press Releases to Generate News Stories. In *Proceedings of the Third Conference on Applied Natural Language Processing*, pages 170–177, 1992.
- [Anderson 77]
S. Anderson. Formal Syntax. In Wasow and Akmajian, editors, *Comment on the paper by Wasow in Culicover.*, pages 361–376. Academic Press, London, 1977.
- [André 99]
E. André, editor. *Applied Artificial Intelligence Journal, Special Double Issue on Animated Interface Agents*. 1999. Vol. 13, No. 4-5.
- [André *et al.* 98]
E. André, T. Rist, and J. Müller. WebPersona: A Life-Like Presentation Agent for the World-Wide Web. *Knowledge-Based Systems*, 11(1):25–36, 1998.
- [Andreasen 83]
N. C. Andreasen. *Scale for the Assessment of Negative Symptoms*. University of Iowa Press, Iowa City, 1983. Cited by 0000.
- [Androutsopoulos *et al.* 94]
I. Androutsopoulos, G. Ritchie, and P. Thanish. Natural Language Interfaces to Databases - An Introduction. Technical report, Department of Artificial Intelligence, University of Edinburgh, 1994. Research Paper no. 709.
- [Angeletou *et al.* 11]
S. Angeletou, M. Rowe, and H. Alani. Modelling and analysis of user behaviour in online communities. In *Proceedings of the 10th International Conference on the Semantic Web, ISWC'11*, pages 35–50. Springer-Verlag, 2011.
- [Angelova *et al.* 01]
G. Angelova, K. Bontcheva, R. Mitkov, N. Nikolov, and N. Nicolov, editors. Tzigov Chark, Bulgaria, Sep 2001.
- [Angelova *et al.* 03]
G. Angelova, K. Bontcheva, R. Mitkov, N. Nicolov, and N. Nikolov, editors. Borovets, Bulgaria, Sep 2003.
- [Angluin & Smith 83]
D. Angluin and C. Smith. Inductive Inference: Theory and Methods. *Computer Surveys*, 15(3), 1983.
- [Ankolekar *et al.* 06]
A. Ankolekar, K. Sycara, J. Herbsleb, and R. Kraut. Supporting Online Problem Solving Communities with the Semantic Web. In *Proc. of WWW*, 2006.
- [Ankolekar *et al.* 08]
A. Ankolekar, M. Krotzsch, T. Tran, and D. Vrandečić. The two cultures: Mashing up web 2.0 and the semantic web. *Journal of Web Semantics*, 6(1):70–75, February 2008. Web Semantics: Science, Services and Agents on the World Wide Web.
- [ANLP'97 97]
Proceedings of the Fifth Conference on Applied Natural Language Processing, Washington, DC, 1997.
- [Anonymous 04]
Anonymous. Open-source Tools for Creation, Maintenance, and Storage of Lexical Resources for Language Generation from Ontologies. In *Proceedings of 4th Language Resources and Evaluation Conference (LREC'04)*, 2004.

- [Antoniou & vanHermelen 08]
G. Antoniou and F. van Hermelen. *A Semantic Web Primer*. MIT Press, 2nd edition, 2008.
- [Ao & Takagi 05]
H. Ao and T. Takagi. ALICE: an algorithm to extract abbreviations from MEDLINE. *J Am Med Inform Assoc*, 12(5):576–586, 2005.
- [Appelt & Martin 99]
D. Appelt and D. Martin. Named entity extraction from speech: Approach and results using the TextPro system. In *Proceedings of DARPA Broadcast News Workshop*, pages 51–54, 1999.
- [Appelt & Onyshkevych 96]
D. E. Appelt and B. Onyshkevych. The common pattern specification language. In *Proceedings of TIPSTER Text Program, Phase III Workshop*, pages 23–30, Morristown, NJ, USA, 1996. Association for Computational Linguistics.
- [Appelt 85]
D. E. Appelt. *Planning English Sentences*. Cambridge University Press, Cambridge, 1985.
- [Appelt 86]
D. E. Appelt. Planning English referring expressions. In B. L. W. B. Grosz, K. S. Jones, editor, *Readings in Natural Language Processing*, pages 501 – 517. Morgan Kaufmann, California, 1986.
- [Appelt 96]
D. Appelt. The Common Pattern Specification Language. Technical report, SRI International, Artificial Intelligence Center, 1996.
- [Appelt 99]
D. E. Appelt. An Introduction to Information Extraction. *Artificial Intelligence Communications*, 12(3):161–172, 1999.
- [Appelt *et al.* 93a]
D. Appelt, J. Hobbs, J. Bear, D. Israel, and M. Tyson. FASTUS: A finite-state processor for information extraction from real-world text. In *Proceedings. IJCAI-93*, 1993.
- [Appelt *et al.* 93b]
D. Appelt, J. Hobbs, J. Bear, D. Israel, M. Kameyama, and M. Tyson. Description of the JV-FASTUS system as used for MUC-5. In *Proceedings of the Fourth Message Understanding Conference MUC-5*, pages 221–235. Morgan Kaufmann, California, 1993.
- [Appelt *et al.* 95]
E. R. Appelt, J. Hobbs, J. Bear, D. Israel, M. Kameyama, A. Kehler, D. Martin, K. Myers, and M. Tyson. SRI International FASTUS system: MUC-6 Test Results and Analysis. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*. Morgan Kaufmann, California, 1995.
- [Apter 89]
M. Apter. *Reversal Theory: Motivation, Emotion and Personality*. Routledge, London, 1989.
- [Aras *et al.* 10]
H. Aras, M. Krause, A. Haller, and R. Malaka. Webparady: harvesting qa by hc. In *Proceedings of the ACM SIGKDD Workshop on Human Computation*, pages 49–52, 2010.
- [Arbib & Hesse 86]
M. Arbib and M. Hesse. *The Construction of Reality*. Cambridge University Press, Cambridge, 1986.
- [Archambault *et al.* 11]
D. Archambault, D. Greene, P. Cunningham, and N. J. Hurley. ThemeCrowds: Multiresolution summaries of Twitter usage. In *Workshop on Search and Mining User-Generated Contents (SMUC)*, pages 77–84, 2011.

- [ARDA 06]
ARDA. AQUAINT TimeML corpus, 2006.
- [Ardissono & Goy 00]
L. Ardissono and A. Goy. Dynamic generation of adaptive web catalogs. In P. Brusilovsky, O. Stock, and C. Strapparava, editors, *Adaptive Hypermedia and Adaptive Web-Based Systems*, number 1892 in Lecture Notes in Computer Science, pages 5–16, Berlin Heidelberg, 2000. Springer Verlag.
- [Ardissono *et al.* 01]
L. Ardissono, F. Portis, P. Torasso, F. Bellifemine, A. Chiarotto, and A. Difino. Architecture of a System for the Generation of Personalized Electronic Program Guides. In *Workshop on Personalization in Future TV, User Modeling 2001*, Sonthofen, Germany, 2001.
- [Ardissono *et al.* 04]
L. Ardissono, A. Kobsa, and M. Maybury, editors. *Personalized Digital Television*. Kluwer, Dordrecht, 2004.
- [Ardon *et al.* 11]
S. Ardon, A. Bagchi, A. Mahanti, A. Ruhela, A. Seth, R. M. Tripathy, and S. Triukose. Spatio-temporal analysis of topic popularity in twitter. *CoRR*, abs/1111.2904, 2011.
- [Arko *et al.* 06]
R. Arko, K. Ginger, K. Kastens, and J. Weatherly. Using annotations to add value to a digital library for education. *D-Lib Magazine*, 12(5), May 2006.
- [Arnold & Gosling 98]
K. Arnold and J. Gosling. *The Java Programming Language, Second Edition*. Addison-Wesley, Reading, MA, 1998.
- [Arnold 93]
E. Arnold. Special issue on evaluation of MT systems. *Machine Translation*, 8(1-2):1–126, 1993.
- [Arnold 01]
D. Arnold. Computer Graphics and Archaeology: Realism and Symbiosis. SIGGRAPH EG Campfire, 2001. citeseer.nj.nec.com/541996.html.
- [Aronson & Lang 10]
A. R. Aronson and F.-M. Lang. An overview of MetaMap: historical perspective and recent advances. *Journal of the American Medical Informatics Association (JAMIA)*, 17:229–236, 2010.
- [Aronson & Rindflesch 97]
A. Aronson and T. Rindflesch. Query expansion using the UMLS metathesaurus. In D. Masys, editor, *Proceedings of the 1997 AMIA Annual Fall Symposium*, pages 485–489, 1997.
- [Aroyo & Houben 10]
L. Aroyo and G.-J. Houben. User modeling and adaptive semantic web. *Semantic Web*, 1(1,2):105–110, April 2010.
- [ARPA 90]
DARPA Speech and Natural Language Workshop, 1990.
- [ARPA 91]
Defense Advanced Research Projects Agency. *Proceedings of the Third Message Understanding Conference (MUC-3)*. Morgan Kaufmann, California, 1991.
- [ARPA 92a]
DARPA Speech and Natural Language Workshop. Harriman, NY, 1992.

- [ARPA 92b]
 Defense Advanced Research Projects Agency. *Proceedings of the Fourth Message Understanding Conference (MUC-4)*. Morgan Kaufmann, California, 1992.
- [ARPA 93a]
 Advanced Research Projects Agency. *Proceedings of the Fifth Message Understanding Conference (MUC-5)*. Morgan Kaufmann, California, 1993.
- [ARPA 93b]
 Advanced Research Projects Agency. *Proceedings of the TIPSTER Text Program (Phase I)*. Morgan Kaufmann, California, 1993.
- [ARPA 95]
 Defense Advanced Research Projects Agency. *Proceedings of the Sixth Message Understanding Conference (MUC-6)*. Morgan Kaufmann, California, 1995.
- [ARPA 96]
 Advanced Research Projects Agency. *Proceedings of the TIPSTER Text Program (Phase II)*. Morgan Kaufmann, California, 1996.
- [Arranz *et al.* 95]
 M. Arranz, I. Radford, S. Ananiadou, and J.-I. Tsujii. Tools for sublanguage-based semantic knowledge acquisition from corpora. In *Proc. of 11th Conference of the Spanish Society for NLP*, Bilbao, Spain, 1995.
- [Arrington 06]
 M. Arrington. Odeo Releases Twtr. *TechCrunch*, 2006. <http://techcrunch.com/2006/07/15/is-twtr-interesting/>.
- [Artiles *et al.* 07]
 J. Artiles, J. Gonzalo, and S. Sekine. The SemEval-2007 WePS Evaluation: Establishing a benchmark for Web People Search Task. In *Proceedings of Semeval 2007, Association for Computational Linguistics*, 2007.
- [Artola *et al.* 02]
 X. Artola, A. D. de Ilarraza, N. Ezeiza, K. Gojenola, G. Hernández, and A. Soroa. A Class Library for the Integration of NLP Tools: Definition and implementation of an Abstract Data Type Collection for the manipulation of SGML documents in a context of stand-off linguistic annotation. In *Proceedings of LREC 2002 Third International Conference on Language Resources and Evaluation*, pages 1650–1657, Las Palmas, Canary Islands - Spain, 2002.
- [Artstein & Poesio 05]
 R. Artstein and M. Poesio. Kappa3 = Alpha (or Beta). Technical Report CS Technical Report CSM-437, Department of Computer Science, University of Essex, Colchester, UK, 2005.
- [Arunachalam & Sarkar 13]
 R. Arunachalam and S. Sarkar. The new eye of government: Citizen sentiment analysis in social media. In *Sixth International Joint Conference on Natural Language Processing*, page 23, 2013.
- [Asanovic *et al.* 06]
 K. Asanovic, R. Bodik, B. Catanzaro, J. Gebis, P. Husbands, K. Keutzer, D. Patterson, W. Plishker, J. Shalf, S. Williams, and K. Yelick. The landscape of parallel computing research: A view from berkeley. Technical Report UCB/EECS-2006-183, EECS Department, University of California, Berkeley, December 18 2006.
- [Asher 84]
 N. Asher. Lingusitic Understanding and Non-monotonic Reasoning. In *Proceedings of the non-monotonic reasoning workshop*, pages 3–20. AAAI, 1984.

- [Ashish 05]
N. Ashish. Semantic-web technology: Applications at nasa. In Y. Kalfoglou, M. Schorlemmer, A. Sheth, S. Staab, and M. Uschold, editors, *Semantic Interoperability and Integration*, number 04391 in Dagstuhl Seminar Proceedings, 2005.
- [Assel *et al.* 11]
M. Assel, A. Cheptsov, B. Czink, D. Damljanovic, and J. Quesada. MPI Realization of High Performance Search for Querying Large RDF Graphs using Statistical Semantics . In *Proceedings of the 1st Workshop on High-Performance Computing for the Semantic Web, Collocated with the 8th Extended Semantic Web Conference (ESWC 2011)*, Heraklion, Greece, June 2011.
- [Asur & Huberman 10]
S. Asur and B. A. Huberman. Predicting the Future with Social Media. *CoRR*, abs/1003.5, 2010.
- [Aswani & Gaizauskas 09]
N. Aswani and R. Gaizauskas. Evolving a General Framework for Text Alignment: Case Studies with Two South Asian Languages. In *Proceedings of the International Conference on Machine Translation: Twenty-Five Years On*, Cranfield, Bedfordshire, UK, November 2009.
- [Aswani & Gaizauskas 10a]
N. Aswani and R. Gaizauskas. Developing Morphological Analysers for South Asian Languages: Experimenting with the Hindi and Gujarati Languages. In *7th Language Resources and Evaluation Conference (LREC)*, La Valletta, Malta, May 2010. ELRA.
- [Aswani & Gaizauskas 10b]
N. Aswani and R. Gaizauskas. English-Hindi Transliteration using Multiple Similarity Metrics. In *7th Language Resources and Evaluation Conference (LREC)*, La Valletta, Malta, May 2010. ELRA.
- [Aswani *et al.* 05]
N. Aswani, V. Tablan, K. Bontcheva, and H. Cunningham. Indexing and Querying Linguistic Metadata and Document Content. In *Proceedings of Fifth International Conference on Recent Advances in Natural Language Processing (RANLP2005)*, Borovets, Bulgaria, 2005.
- [Aswani *et al.* 06]
N. Aswani, K. Bontcheva, and H. Cunningham. Mining information for instance unification. In *5th International Semantic Web Conference (ISWC2006)*, Athens, Georgia, USA, 2006.
- [Aswani *et al.* 12]
N. Aswani, M. Greenwood, K. Bontcheva, L. Derczynski, J. M. Schneider, H.-U. Krieger, and T. Declerck. Multilingual, ontology-based information extraction from stream media - v1. Technical Report D2.2.1, TrendMiner Project Deliverable, 2012.
- [Aswani *et al.* 13]
N. Aswani, G. Gorrell, K. Bontcheva, and J. Petrak. Multilingual, ontology-based information extraction from stream media - v2. Technical Report D2.3.2, TrendMiner Project Deliverable, 2013. http://www.trendminer-project.eu/images/d2.3.2_final.pdf.
- [Attardi 10]
G. Attardi. Phratris – a phrase annotation game, 2010.
- [Aubert *et al.* 94]
X. Aubert, C. Dugast, H. Ney, and V. Steinbiss. Large vocabulary continuous speech recognition of Wall Street Journal data. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 2, pages 129–132, Adelaide, 1994.
- [Aue & Gamon 05a]
A. Aue and M. Gamon. Automatic identification of sentiment vocabulary: Exploiting low association with known sentiment terms. In *Proceedings of the ACL Workshop on Feature Engineering for Machine Learning in Natural Language Processing*. ACL, 2005.

- [Aue & Gamon. 05b]
 A. Aue and M. Gamon. Customizing sentiment classifiers to new domains: a case study. In *Proc. of the International Conference on Recent Advances in Natural Language Processing*, Borovetz, Bulgaria, 2005.
- [Auer *et al.* 06]
 S. Auer, S. Dietzold, and T. Riechert. OntoWiki—a tool for social, semantic collaboration. In *Proceedings of the Fifth International Semantic Web Conference (ISWC'06)*, 2006.
- [Augenstein 14a]
 I. Augenstein. Joint Information Extraction from the Web using Linked Data. In P. Mika, T. Tudorache, A. Bernstein, C. Welty, C. A. Knoblock, D. Vrandečić, P. T. Groth, N. F. Noy, K. Janowicz, and C. A. Goble, editors, *International Semantic Web Conference (2)*, volume 8797 of *Lecture Notes in Computer Science*, pages 505–512, Heidelberg, Germany, 2014. Springer.
- [Augenstein 14b]
 I. Augenstein. Seed Selection for Distantly Supervised Web-Based Relation Extraction. In D. Maynard, M. van Erp, and B. Davis, editors, *Proceedings of the Third Workshop on Semantic Web and Information Extraction*, pages 17–24, Dublin, Ireland, 2014. Association for Computational Linguistics and Dublin City University.
- [Augenstein 16]
 I. Augenstein. *Web Relation Extraction with Distant Supervision*. Unpublished PhD thesis, University of Sheffield, UK, 2016.
- [Augenstein *et al.* 12]
 I. Augenstein, S. Padó, and S. Rudolph. LODifier: Generating Linked Data from Unstructured Text. In E. Simperl, P. Cimiano, A. Polleres, Ó. Corcho, and V. Presutti, editors, *ESWC*, volume 7295 of *Lecture Notes in Computer Science*, pages 210–224, Heidelberg, Germany, 2012. Springer.
- [Augenstein *et al.* 14]
 I. Augenstein, D. Maynard, and F. Ciravegna. Relation Extraction from the Web using Distant Supervision. In K. Janowicz, S. Schlobach, P. Lambrix, and E. Hyvönen, editors, *EKAW*, volume 8876 of *Lecture Notes in Computer Science*, pages 26–41, Heidelberg, Germany, 2014. Springer.
- [Augenstein *et al.* 16]
 I. Augenstein, T. Rocktäschel, A. Vlachos, and K. Bontcheva. Stance detection with bidirectional conditional encoding. In *EMNLP*, pages 876–885, 2016.
- [Aumueller 05a]
 D. Aumueller. Semantic authoring and retrieval within a Wiki. In *Proceedings of the Second European Semantic Web Conference (ESWC'05)*, 2005.
- [Aumueller 05b]
 S. Aumueller, D. Auer. Towards a semantic wiki experience - desktop integration and interactivity in wikisar. 1st Workshop on The Semantic Desktop, Next Generation Personal Information Management and Collaboration Infrastructure, 2005.
- [Aust 98]
 H. Aust. *Sprachverstehen und Dialogmodellierung in natürlichsprachlichen Informationssystemen*. Ph.D. Thesis, Computer Science Department, RWTH Aachen, 1998.
- [Aust *et al.* 95]
 H. Aust, M. Oerder, F. Seide, and V. Steinbiss. The Philips automatic train timetable information system. *Speech Communication*, 17:249–262, November 1995.
- [Auxerre 86]
 P. Auxerre. MASQUE Modular Answering System for Queries in English - User's Manual. June 1986.

- [Axelrod *et al.* 93]
 B. N. Axelrod, R. S. Goldman, and L. D. Alphas. Validation of the 16-item negative symptom assessment. *Journal of psychiatric research*, 27(3):253–258, September 1993. Cited by 0016.
- [Azar 89]
 S. Azar. *Understanding and Using English Grammar*. Prentice Hall Regents, 1989.
- [Azzam 96]
 S. Azzam. Resolving Anaphors in Embedded Sentences. In *Proceedings of the 34th meetings of the Association for Computational Linguistics, Santa Cruz : California*, 1996.
- [Azzam 5a]
 S. Azzam. *Computation of Ambiguities (Anaphors and PPs) in NL texts. CLAM : The prototype*. Unpublished PhD thesis, Paris Sorbonne University, 1995a.
- [Azzam 5b]
 S. Azzam. Anaphors, PPs and Disambiguation Process for conceptual analysis. In *14th International Joint Conference on Artificial Intelligence (IJCAI'95)*. Morgan Kaufmann, California, 1995b.
- [Azzam *et al.* 97a]
 S. Azzam, K. Humphreys, R. Gaizauskas, H. Cunningham, and Y. Wilks. Using a Language Independent Domain Model for Multilingual Information Extraction. In C. Spryrodopoulos, editor, *Proceedings of the IJCAI-97 Workshop on Multilinguality in the Software Industry: the AI Contribution (MULSAIC-97)*, 1997.
- [Azzam *et al.* 97b]
 S. Azzam, K. Humphreys, R. Gaizauskas, H. Cunningham, and Y. Wilks. A Design for Multilingual Information Extraction (poster). In *IJCAI-97*, 1997.
- [Azzam *et al.* 98a]
 S. Azzam, K. Humphreys, and R. Gaizauskas. Coreference Resolution in a Multilingual Information Extraction System. In *Linguistic Coreference Workshop, in First International Conference on Language Resources and Evaluation, Granada, Spain*, 1998.
- [Azzam *et al.* 98b]
 S. Azzam, K. Humphreys, and R. Gaizauskas. Evaluating a Focus-Based Approach to Anaphora Resolution. In *Proceedings of COLING-ACL'98*, pages 74–78, 1998.
- [Azzam *et al.* 99a]
 S. Azzam, K. Humphreys, and R. Gaizauskas. Using a Language Independent Domain Model for Multilingual Information Extraction. *Journal of Applied Artificial Intelligence*, 1999.
- [Azzam *et al.* 99b]
 S. Azzam, K. Humphreys., and R. Gaizauskas. Using coreference chains for text summarization. In *Proceedings of the ACL '99 Workshop on Coreference and its Applications*, Maryland, 1999.
- [Azzam *et al.* 99c]
 S. Azzam, K. Humphreys, R. Gaizauskas, H. Cunningham, and Y. Wilks. Using a Language Independent Domain Model for Multilingual Information Extraction. *Applied Artificial Intelligence*, 13(7), 1999. Special Issue on Multilinguality in the Software Industry: the AI Contribution (MULSAIC-97).
- [Backstrom *et al.* 10]
 L. Backstrom, E. Sun, and C. Marlow. Find me if you can: Improving geographical prediction with social and spatial proximity. In *Proceedings of the 19th International Conference on World Wide Web*, pages 61–70. ACM, 2010.
- [Baddiel 97]
 I. Baddiel. *Ultimate Football*. Dorling Kindersley, 1997.

- [Bagga & Baldwin 98]
A. Bagga and B. Baldwin. Entity-Based Cross-Document Coreferencing Using the Vector Space Model. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and the 17th International Conference on Computational Linguistics (COLING-ACL'98)*, pages 79–85, 1998.
- [Bagga & Biermann 00]
A. Bagga and A. W. Biermann. A methodology for cross-document coreference. In *Proceedings of the Fifth Joint Conference on Information Sciences (JCIS 2000)*, pages 207–210, 2000.
- [Bagga 98a]
A. Bagga. *Coreference, Cross-document Coreference and Information Extraction Methodologies*. Unpublished PhD thesis, Duke University, USA, 1998.
- [Bagga 98b]
A. Bagga. Analysing the complexity of a domain with respect to an Information Extraction task. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)*. http://www.itl.nist.gov/iaui/894.02/related_projects/muc/index.html, 1998.
- [Bagga et al. 01]
A. Bagga, B. Baldwin, and G. Ramesh. Methodology for cross-document coreference over degraded data sources. In Angelova et al. [Angelova et al. 01], pages 15–21.
- [Bahl & Jelinek 75]
L. R. Bahl and F. Jelinek. Decoding for channels with insertions, deletions and substitutions with applications to speech recognition. *IEEE Transactions on Information Theory*, 21:404–411, 1975.
- [Bahl & Jelinek 88]
L. R. Bahl and F. Jelinek. Apparatus and method for determining a likely word sequence from labels generated by an acoustic processor. US Patent 4,748,670, May 1988.
- [Bahl et al. 83]
L. R. Bahl, F. Jelinek, and R. L. Mercer. A maximum likelihood approach to continuous speech recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-5:179–190, 1983.
- [Bahl et al. 86]
L. R. Bahl, P. F. Brown, P. V. de Souza, and R. L. Mercer. Maximum mutual information estimation of hidden Markov model parameters for speech recognition. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 49–52, Tokyo, 1986.
- [Bahl et al. 87]
L. R. Bahl, P. F. Brown, P. V. de Souza, and R. L. Mercer. Speech recognition with continuous parameter hidden Markov models. *Computer Speech and Language*, 2:219–234, 1987.
- [Bahl et al. 89]
L. R. Bahl, R. Bakis, J. Bellegarda, P. F. Brown, D. Burshtein, S. K. Das, P. V. de Souza, P. S. Gopalakrishnan, F. Jelinek, D. Kanevsky, R. L. Mercer, A. J. Nadas, D. Nahamoo, and M. A. Picheny. Large vocabulary natural language continuous speech recognition. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 465–467, Glasgow, 1989.
- [Bahl et al. 91]
L. R. Bahl, S. Das, P. V. de Souza, M. Epstein, R. L. Mercer, B. Merialdo, D. Nahamoo, M. A. Picheny, and J. Powell. Automatic phonetic baseform determination. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 173–176, Toronto, 1991.
- [Bahl et al. 92]
L. R. Bahl, P. S. Gopalakrishnan, D. S. Kanevsky, and D. Nahamoo. A fast admissible method for identifying a short list of candidate words. *Computer Speech and Language*, 6:215–224, 1992.

- [Bahl *et al.* 93]
L. R. Bahl, S. V. De Gennaro, P. S. Gopalakrishnan, and R. L. Mercer. A fast approximate acoustic match for large vocabulary speech recognition. *IEEE Transactions on Signal Processing*, 1:59–67, 1993.
- [Bailey 06]
B. Bailey. Getting the Complete Picture with Usability Testing. Usability Updates Newsletter, U.S. Department of Health and Human Services, March 2006.
- [Baker *et al.* 02]
P. Baker, A. Hardie, T. McEnery, H. Cunningham, and R. Gaizauskas. EMILLE, A 67-Million Word Corpus of Indic Languages: Data Collection, Mark-up and Harmonisation. In *Proceedings of 3rd Language Resources and Evaluation Conference (LREC'2002)*, pages 819–825, 2002.
- [Balasubramanya *et al.* 06]
M. Balasubramanya, M. Higgins, P. Lucas, J. Senn, and D. Widdows. Collaborative annotation that lasts forever: Using peer-to-peer technology for disseminating corpora and language resources. In *Fifth International Conference on Language Resources and Evaluation (LREC 2006)*, Genoa, Italy, May 2006. ELRA.
- [Baldonado *et al.* 97]
M. Baldonado, C.-C. K. Chang, L. Gravano, and A. Paepcke. Metadata for Digital Libraries: Architecture and Design Rationale. In *Proceedings of the Second ACM International Conference on Digital Libraries*, pages 47–56, 1997. citeseer.nj.nec.com/baldonado97metadata.html.
- [Baldwin & Lui 10]
T. Baldwin and M. Lui. Language identification: The long and the short of the matter. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 229–237, Los Angeles, California, June 2010.
- [Baldwin *et al.* 13]
T. Baldwin, P. Cook, M. Lui, A. MacKinlay, and L. Wang. How noisy social media text, how different social media sources. In *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, pages 356–364. ACL, 2013.
- [Baldwin *et al.* 15a]
T. Baldwin, B. Han, M. M. C. de Marneffe, Y.-B. Kim, A. Ritter, and W. Xu. Findings of the 2015 Workshop on Noisy User-generated Text. In *Proceedings of the Workshop on Noisy User-generated Text (WNUT 2015)*. Association for Computational Linguistics, 2015.
- [Baldwin *et al.* 15b]
T. Baldwin, Y.-B. Kim, M. C. de Marneffe, A. Ritter, B. Han, and W. Xu. Shared Tasks of the 2015 Workshop on Noisy User-generated Text: Twitter Lexical Normalization and Named Entity Recognition. In W. Xu, B. Han, and A. Ritter, editors, *Proceedings of the Workshop on Noisy User-generated Text*, pages 126–135, Beijing, China, 2015. Association for Computational Linguistics.
- [Ballim & Wilks 91a]
A. Ballim and Y. Wilks. *Artificial Believers*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1991.
- [Ballim & Wilks 91b]
A. Ballim and Y. Wilks. Beliefs, Stereotypes and Dynamic agent Modelling. *User Modeling and user-adapted interaction*, 1:33–65, 1991.
- [Ballim 95]
A. Ballim. Abstract Data Types for MULTTEXT Tool I/O. LRE 62-050 Deliverable 1.2.1, 1995.

- [Balog *et al.* 06]
 K. Balog, G. Mishne, and M. De Rijke. Why are they excited?: identifying and explaining spikes in blog mood levels. In *Proceedings of the Eleventh Conference of the European Chapter of the Association for Computational Linguistics: Posters & Demonstrations*, pages 207–210. Association for Computational Linguistics, 2006.
- [Balog *et al.* 10]
 K. Balog, P. Serdyukov, and A. P. de Vries. Overview of the trec 2010 entity track. In E. M. Voorhees and L. P. Buckland, editors, *Proceedings of The Nineteenth Text REtrieval Conference, TREC 2010, Gaithersburg, Maryland, USA, November 16-19, 2010*, 2010.
- [Balog *et al.* 12]
 K. Balog, Y. Fang, M. de Rijke, P. Serdyukov, and L. Si. Expertise retrieval. *Foundations and Trends in Information Retrieval*, 6(2–3):127–256, 2012.
- [Bamman & Smith 15]
 D. Bamman and N. A. Smith. Contextualized sarcasm detection on twitter. In *Ninth International AAAI Conference on Web and Social Media*, 2015.
- [Banerjee *et al.* 87]
 J. Banerjee, W. Kim, H. Kim, and H. F. Korth. Semantics and implementation of schema evolution in object-oriented databases. In *Proceedings of SIGMOD*, 1987.
- [Banerjee *et al.* 01]
 P. K. Banerjee, S. D. Myers, A. Baker, M. Andersen, and S. Bichet. High performance drug discovery - an operating model for a new era. Executive briefing, Accenture, 2001.
- [Banko & Etzioni 08]
 M. Banko and O. Etzioni. The tradeoffs between open and traditional relation extraction. In *Proceedings of ACL-08*, 2008.
- [Bansal & Koudas 07]
 N. Bansal and N. Koudas. Blogscope: Spatio-temporal analysis of the blogosphere. In *Proceedings of the 16th International Conference on World Wide Web, WWW '07*, pages 1269–1270, 2007.
- [Bao & Honavar 03]
 J. Bao and V. Honavar. Collaborative Ontology Building with Wiki@nt - A Multi-agent Based Ontology Building Environment . In *EON 2004 Workshop*, 2003.
- [Bao & Honavar 04]
 J. Bao and V. Honavar. Collaborative ontology building with wiki@nt - a multi-agent based ontology building environment. Technical report, TR-343, Computer Science, Iowa State University, 2004.
- [Barbieri *et al.* 01]
 M. Barbieri, M. Ceccarelli, G. Mekenkamp, and J. Nesvadba. A Personal TV Receiver with Storage and Retrieval Capabilities. In *Workshop on Personalization in Future TV, User Modeling 2001*, Sonthofen, Germany, 2001.
- [Barbu & Mitkov 01]
 C. Barbu and R. Mitkov. Evaluation tool for rule-based anaphora resolution methods. In *Proceedings of ACL'01*, Toulouse, France, 2001.
- [Barga *et al.* 11a]
 R. Barga, D. Gannon, and D. Reed. The client and the cloud: Democratizing research computing. *IEEE Internet Computing*, 15:72–75, 2011.
- [Barga *et al.* 11b]
 R. Barga, D. Gannon, and D. Reed. The client and the cloud: Democratizing research computing. *IEEE Internet Computing*, 15(1):72–75, 2011.

- [Barnard & Botha 93]
E. Barnard and E. C. Botha. Back-propagation uses prior information efficiently. *IEEE Transactions on Neural Networks*, 4:794–802, 1993.
- [Barnickel *et al.* 09]
T. Barnickel, J. Weston, R. Collobert, H.-W. Mewes, and V. Stümpflen. Large scale application of neural network based semantic role labeling for automated relation extraction from biomedical texts. *PLoS ONE*, 4(7):e6393, 07 2009.
- [Barr & Feigenbaum 81]
A. Barr and E. Feigenbaum. *The Handbook of Artificial Intelligence Vol. I*. Pitman, London, 1981.
- [Barwise & Perry 85]
J. Barwise and J. Perry. Shifting Situations and Shaken Attitudes. *Linguistics and Philosophy*, 8(1), 1985.
- [Barzilay & Mckeown 05]
R. Barzilay and K. R. Mckeown. Sentence fusion for multidocument news summarization. *Computational Linguistics*, 31:297–328, 2005.
- [Basave *et al.* 14]
A. E. C. Basave, G. Rizzo, A. Varga, M. Rowe, M. Stankovic, and A.-S. Dadzie. Making sense of microposts (#microposts2014) named entity extraction & linking challenge. In *4th Workshop on Making Sense of Microposts (#Microposts2014)*, 2014.
- [Basili *et al.* 93]
R. Basili, M. T. Pazienza, and P. Velardi. Acquisition of selectional patterns. *Journal of Machine Translation*, 8:175–201, 1993.
- [Basili *et al.* 95]
R. Basili, M. Della Rocca, M. Pazienza, and P. Velardi. Contexts and categories: Tuning a general purpose verb classification to sublanguages. In *Proc. of International Conference on Recent Advances in Natural Language Processing*, Tzigov Chark, Bulgaria, 1995.
- [Basili *et al.* 96]
R. Basili, M. T. Pazienza, and P. Velardi. An Empirical Symbolic Approach to NLP. *Artificial Intelligence*, 1996. (in press).
- [Basili *et al.* 97a]
R. Basili, M. Della Rocca, and M. Pazienza. Towards a bootstrapping framework for corpus semantic tagging. In *Proceedings of the SIGLEX Workshop "Tagging Text with Lexical Semantics: What, why and how?"*, Washington, D.C., April 1997. ANLP.
- [Basili *et al.* 97b]
R. Basili, M. Pazienza, P. Velardi, R. Catizone, R. Collier, M. Stevenson, Y. Wilks, O. Amsaldi, A. Luk, B. Vauthey, and J. Grandchamp. Extracting case relations from corpora. ECRAN Deliverable 2.4 version 1, 1997.
- [Basili *et al.* 98]
R. Basili, R. Catizone, M. Pazienza, M. Stevenson, P. Velardi, M. Vindigni, and Y. Wilks. An empirical approach to lexical tuning. In *First International Conference on Language Resources and Evaluation (LREC) Workshop on Adapting Lexical and Corpus Resources to Sublanguages and Applications*, Granada, Spain, 1998.
- [Basili *et al.* 00]
R. Basili, M. T. Pazienza, and M. Vindign. Corpus-driven learning of event recognition rules. In *ECAI Workshop on Machine Learning for Information Extraction*, Berlin, 2000.

- [Basili *et al.* 07]
 R. Basili, R. Petitti, and D. Saracino. Lsa-based automatic acquisition of semantic image descriptions. In B. Falciديو, M. Spagnuolo, Y. Avrithis, I. Kompatsiaris, and P. Buitelaar, editors, *Semantic Multimedia*, volume 4816 of *Lecture Notes in Computer Science*, pages 41–55. Springer Berlin Heidelberg, 2007.
- [Bass *et al.* 97]
 L. Bass, P. Clements, and R. Kazman. *Software Architecture in Practice*. Addison-Wesley, Reading, MA, 1997.
- [Bast *et al.* 07]
 H. Bast, A. Chitea, F. Suchanek, and I. Weber. Ester: efficient search on text, entities, and relations. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '07, pages 671–678. ACM, 2007.
- [Bast *et al.* 12a]
 H. Bast, F. Bäurle, B. Buchhold, and E. Haussmann. Broccoli: Semantic full-text search at your fingertips. *CoRR*, abs/1207.2615, 2012.
- [Bast *et al.* 12b]
 H. Bast, F. Bäurle, B. Buchhold, and E. Haussmann. A case for semantic full-text search. In *Proceedings of the 1st Joint International Workshop on Entity-Oriented and Semantic Search*, JIWES '12, pages 4:1–4:3. ACM, 2012.
- [Bateman 90]
 J. A. Bateman. Upper Modeling: organizing knowledge for natural language processing. In *5th. International Workshop on Natural Language Generation, 3-6 June 1990*, Pittsburgh, PA., 1990.
- [Bateman 93]
 J. A. Bateman. Ontology construction and natural language. In *Proceedings of the International Workshop on Formal Ontology*, 1993.
- [Bateman 97]
 J. A. Bateman. Enabling technology for multilingual natural language generation: the kpml development environment. *Journal of Natural Language Engineering*, 3:15 – 55, 1997.
- [Bateman *et al.* 95a]
 J. Bateman, B. Magnini, and G. Fabris. The generalized upper model knowledge base: Organization and use. In *Proceedings of the Conference on Knowledge Representation and Sharing*, 1995.
- [Bateman *et al.* 95b]
 J. A. Bateman, B. Magnini, and G. Fabris. The Generalized Upper Model Knowledge Base: Organization and Use. In *Towards Very Large Knowledge Bases*, pages 60–72, 1995.
- [Bates *et al.* 94]
 M. Bates, R. Bobrow, R. Ingria, S. Peters, and D. Stallard. Advances in BBN’s Spoken Language System. In *Proceedings of the Spoken Language Technology Workshop*, pages 43–47, March 1994.
- [Batista & Figueira 17]
 F. Batista and Á. Figueira. The complementary nature of different nlp toolkits for named entity recognition in social media. In *Progress in Artificial Intelligence: 18th EPIA Conference on Artificial Intelligence, EPIA 2017, Porto, Portugal, September 5-8, 2017, Proceedings*, pages 803–814, 2017.
- [Baud *et al.* 92]
 R. Baud, A.-M. Rassinoux, and J.-R. Scherrer. Natural language processing and medical records. In *Proc. of Medinfo '92*, pages 1362–1367, 1992.

- [Baum *et al.* 70]
L. E. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics*, 41:164–171, 1970.
- [Baumgartner *et al.* 05]
R. Baumgartner, O. Frölich, G. Gottlob, P. Harz, M. Herzog, and P. Lehmann. Web data extraction for business intelligence: the lixto approach. In *Proc. of BTW 2005*, 2005.
- [Baumgartner *et al.* 07]
R. Baumgartner, O. Froelich, and G. Gottlob. The Lixto Systems Applications in Business Intelligence and Semantic Web. In *Proceedings of the European Semantic Web Conference (ESWC'07)*, pages 16–26, 2007.
- [Baxendale 58]
P. B. Baxendale. Machine-made index for technical literaturean experiment. *IBM Journal of Research and Development*, 2(4):354–361, 1958.
- [Beale & Jackson 90]
R. Beale and T. Jackson. *Neural Computing: An Introduction*. 1990.
- [Beale *et al.* 95]
S. Beale, S. Nirenburg, and K. Mahesh. Semantic Analysis in the Mikrokosmos Machine Translation Project. In *Proceedings of the Second Symposium on Natural Language Processing (SNLP-95)*, 1995.
- [Bearman & Trant 05]
D. Bearman and J. Trant. Social terminology enhancement through vernacular engagement exploring collaborative annotation to encourage interaction with museum collections. *D-Lib Magazine*, 11(9), September 2005.
- [Beaudoin 08]
C. Beaudoin. Explaining the relationship between internet use and interpersonal trust: Taking into account motivation and information overload. *Journal of Computer Mediated Communication*, 13:550—568, 2008.
- [Bechhofer *et al.* 04]
S. Bechhofer, F. van Harmelen, J. Hendler, I. Horrocks, D. L. McGuinness, P. F. Patel-Schneider, and L. A. Stein. OWL web ontology language reference. W3C recommendation, W3C, <http://www.w3.org/>, Feb 2004.
- [Bechtel 88]
W. Bechtel. *Philosophy of Science: An Overview for Cognitive Science*. Lawrence Erlbaum, Hillsdale, NJ, 1988.
- [Beck 00]
K. Beck. *eXtreme Programming eXplained*. Addison-Wesley, Upper Saddle River, NJ, USA, 2000.
- [Becker *et al.* 02]
M. Becker, W. Drozdzyński, H. Krieger, J. Piskorski, U. Schafer, and F. Xu. Sprout – shallow processing with typed feature structures and unification. In *Proceedings of the International Conference on NLP (ICON 2002)*, Mumbai, India, 2002.
- [Becker *et al.* 10]
H. Becker, M. Naaman, and L. Gravano. Learning similarity metrics for event identification in social media. In *Proceedings of the Third International Conference on Web Search and Web Data Mining*, pages 291–300, 2010.

- [Becker *et al.* 11a]
H. Becker, M. Naaman, and L. Gravano. Selecting Quality Twitter Content for Events. In *Proceedings of the Fifth International Conference on Weblogs and Social Media (ICWSM)*, 2011.
- [Becker *et al.* 11b]
H. Becker, M. Naaman, and L. Gravano. Selecting Quality Twitter Content for Events. In *ICWSM*, 2011.
- [Becker *et al.* 11c]
H. Becker, M. Naaman, and L. Gravano. Beyond Trending Topics: Real-World Event Identification on Twitter. In *Proceedings of the Fifth International Conference on Weblogs and Social Media (ICWSM)*, 2011.
- [Behlendorf *et al.* 99]
B. Behlendorf, S. Bradner, J. Hamerly, K. Mckusick, T. O’Reilly, T. Paquin, B. Perens, E. Raymond, R. Stallman, M. Tiemann, L. Torvalds, P. Vixie, L. Wall, and B. Young. *Open Sources: Voices from the Open Source Revolution*. O’Reilly, Sebastopol, California, USA, 1st edition, January 1999.
- [Bekkerman & McCallum 05]
R. Bekkerman and A. McCallum. Disambiguating web appearances of people in a social network. In *Proceedings of WWW-05, the 14th International World Wide Web Conference*, 2005.
- [Belhajjame *et al.* 08]
K. Belhajjame, S. M. Embury, N. W. Paton, R. Stevens, and C. A. Goble. Automatic annotation of web services based on workflow definitions. *ACM Trans. Web*, 2(2):1–34, 2008.
- [Bellegarda & Nahamoo 90]
J. R. Bellegarda and D. Nahamoo. Tied mixture continuous parameter modeling for speech recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38:2033–2045, 1990.
- [Bellegarda *et al.* 96]
J. Bellegarda, J. Butzberger, Y. Chow, N. Coccaro, and D. Naik. A novel word clustering algorithm based on latent semantic analysis. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 172–175, Atlanta GA, 1996.
- [Bellman 57]
R. Bellman. *Dynamic Programming*. Princeton University Press, 1957.
- [Benbrahim & Ahmad 95]
M. Benbrahim and K. Ahmad. Text Summarisation: the Role of Lexical Cohesion Analysis. *The New Review of Document & Text Management*, pages 321–335, 1995.
- [Bender *et al.* 03a]
O. Bender, F. Och, and H. Ney. Maximum entropy models for named entity recognition. In *CoNLL-2003, 7th Conference on Computational Natural Language Learning*, pages 148–151, Edmonton, Canada, May 2003.
- [Bender *et al.* 03b]
O. Bender, K. Macherey, F.-J. Och, and H. Ney. Comparison of Alignment Templates and Maximum Entropy Models for Natural Language Understanding. In *EACL*, pages 11–18, Budapest, Hungary, April 2003.
- [Bender *et al.* 03c]
O. Bender, F. J. Och, and H. Ney. Maximum entropy models for named entity recognition. In W. Daelemans and M. Osborne, editors, *Proceedings of CoNLL-2003*, pages 148–151. Edmonton, Canada, 2003.

- [Bennacef *et al.* 94]
 S. K. Bennacef, H. Bonnea-Maynard, J. L. Gauvain, L. F. Lamel, and W. Minker. A Spoken Language System for Information Retrieval. In *Int. Conf. on Spoken Language Processing (ICSLP)*, pages 1271–1274, September 1994.
- [Beraldin *et al.* 01]
 J.-A. Beraldin, C. Atzeni, G. Guidi, M. Pieraccini, and S. Lazzari. Establishing a Digital 3D Imaging Laboratory for Heritage Applications: First Trials. In *Proceedings of the Italy-Canada 2001 Workshop of on 3D Digital Imaging and Modeling Applications*, Padova, Italy, April 2001. citeseer.nj.nec.com/489618.html.
- [Beraldin *et al.* 02]
 J.-A. Beraldin, M. Picard, S. El-Hakim, G. Godin, C. Latouche, V. Valzano, and A. Bandiera. Exploring a byzantine crypt through a highresolution texture mapped 3D model: combining range data and photogrammetry. In *Proceedings of the International Workshop on Scanning for Cultural Heritage Recording - Complementing or Replacing Photogrammetry*, Corfu, Greece, September 2002.
- [Berardi *et al.* 11]
 G. Berardi, A. Esuli, D. Marcheggiani, and F. Sebastiani. ISTI@ TREC Microblog Track 2011: Exploring the Use of Hashtag Segmentation and Text Quality Ranking. In *TREC*, 2011.
- [Berck & Russel 06]
 P. Berck and A. Russel. Annex – a web-based framework for exploiting annotated media resources. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, 2006.
- [Bergamaschi *et al.* 10]
 S. Bergamaschi, F. Guerra, and B. Leiba. Guest editors’ introduction: Information overload. *IEEE Internet Computing*, 14:10–13, 2010.
- [Berger 97]
 A. Berger. Improved iterative scaling: A gentle introduction. 1997. <http://www.cs.cmu.edu/afs/cs/user/aberger/www/ps/scaling.ps>.
- [Berger *et al.* 96]
 A. L. Berger, S. A. Della Pietra, and V. J. Della Pietra. A maximum entropy approach to natural language processing. *Computational Linguistics*, 22(1):39–72, March 1996.
- [Bergmeir & Benítez 12]
 C. Bergmeir and J. M. Benítez. On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191:192–213, 2012.
- [Berkovsky *et al.* 11]
 S. Berkovsky, J. Freyne, S. Kimani, and G. Smith. Selecting items of relevance in social network feeds. In J. Konstan, R. Conejo, J. Marzo, and N. Oliver, editors, *User Modeling, Adaption and Personalization*, volume 6787 of *Lecture Notes in Computer Science*, pages 329–334. Springer Berlin Heidelberg, 2011.
- [Berland & Charniak 99]
 M. Berland and E. Charniak. Finding parts in very large corpora. In *Proceedings of ACL-99*, pages 57–64, College Park, MD, 1999.
- [Bernardini *et al.* 02]
 F. Bernardini, H. Rushmeier, I. M. Martin, J. Mittleman, and G. Taubin. Building a Digital Model of Michelangelo’s Florentine Pietà. *IEEE Computer Graphics and Applications*, 22(1):59–67, 2002. citeseer.nj.nec.com/bernardini02building.html.
- [Berners-Lee 99]
 T. Berners-Lee. *Weaving the Web*. Orion Business Books, 1999.

- [Berners-Lee 00]
T. Berners-Lee. Semantic Web - XML2000. Technical report, W3C Consortium, <http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html>, 2000.
- [Berners-Lee *et al.* 94]
T. Berners-Lee, R. Cailliau, A. Luotonen, H. Nielsen, and A. Secret. The World-Wide Web. In *Communications of the ACM* 37,8, pages 76–82, 1994.
- [Berners-Lee *et al.* 99]
T. Berners-Lee, D. Connolly, and R. Swick. Web Architecture: Describing and Exchanging Data. Technical report, W3C Consortium, <http://www.w3.org/1999/04/WebData>, 1999.
- [Berners-Lee *et al.* 01]
T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. *Scientific American*, pages 35–43, May 2001.
- [Berners-Lee *et al.* 06]
T. Berners-Lee, W. Hall, J. A. Hendler, K. O’Hara, N. Shadbolt, and D. J. Weitzner. A framework for web science. *Foundations and Trends in Web Science*, 1(1):1–130, 2006.
- [Berni & Butini 02]
M. Berni and F. Butini. Mesmuses: Metaphors for science museums. In *Proceedings of Electronic Imaging in the Visual Arts (EVA2002)*, Florence, Italy, 2002.
- [Bernstein & Kaufmann 06]
A. Bernstein and E. Kaufmann. GINO – A Guided Input Natural Language Ontology Editor. In *5th International Semantic Web Conference (ISWC2006)* [ISW06].
- [Bernstein *et al.* 05]
A. Bernstein, E. Kaufmann, A. Göring, and C. Kiefer. Querying Ontologies: A Controlled English Interface for End-users. In *4th International Semantic Web Conference (ISWC)*, pages 112–126, November 2005.
- [Bernstein *et al.* 10a]
M. S. Bernstein, B. Suh, L. Hong, J. Chen, S. Kairam, and E. H. Chi. EDDI: Interactive topic-based browsing of social status streams. In *Proceedings of the 23rd ACM Symposium on User Interface Software and Technology (UIST)*, pages 303–312, 2010.
- [Bernstein *et al.* 10b]
M. S. Bernstein, G. Little, R. C. Miller, B. Hartmann, M. S. Ackerman, D. R. Karger, D. Crowell, and K. Panovich. SoyLent: A word processor with a crowd inside. In *Proc. of the 23rd ACM Symposium on User Interface Software and Technology*, 2010.
- [Berry *et al.* 95]
M. W. Berry, S. T. Dumais, and G. W. O’Brien. Using linear algebra for intelligent information retrieval. *SIAM Review*, 37:573–595, 1995.
- [Berwick 85]
R. Berwick. *The Acquisition of Syntactic Knowledge*. MIT Press, Cambridge MA, 1985.
- [Bethard 13a]
S. Bethard. ClearTK-TimeML: A minimalist approach to TempEval 2013. In *Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, volume 2, pages 10–14, 2013.
- [Bethard 13b]
S. Bethard. A synchronous context free grammar for time normalization. In *EMNLP*, pages 821–826, 2013.

- [Bethard *et al.* 15]
S. Bethard, L. Derczynski, J. Pustejovsky, and M. Verhagen. SemEval-2015 Task 6: Clinical TempEval. In *Proceedings of the 9th International Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2015.
- [Beyer *et al.* 84]
T. Beyer, J. Carroll, and L. e. Miller. *Talking Minds: The Study of Language in the Cognitive Sciences*. MIT Press, Cambridge, MA, USA, 1984.
- [Bhagdev *et al.* 08]
R. Bhagdev, S. Chapman, F. Ciravegna, V. Lanfranchi, and D. Petrelli. Hybrid Search: Effectively Combining Keywords and Semantic Searches. In *Proceedings of the 5th European Semantic Web Conference*, 2008.
- [Bharat *et al.* 98]
K. Bharat, A. Broder, M. Henzinger, P. Kumar, and S. Venkatasubramanian. The connectivity server: fast access to linkage information on the web. In *7th International WWW Conference*, Brisbane, Australia, 1998. <http://www7.scu.edu.au/programme/fullpapers/1938/com1938.htm>.
- [Bian *et al.* 13]
J. Bian, Y. Yang, and T. Chua. Multimedia summarization for trending topics in microblogs. In Q. He, A. Iyengar, W. Nejdl, J. Pei, and R. Rastogi, editors, *22nd ACM International Conference on Information and Knowledge Management, CIKM'13, San Francisco, CA, USA, October 27 - November 1, 2013*, pages 1807–1812. ACM, 2013.
- [Bianchini *et al.* 08]
D. Bianchini, V. De Antonellis, and M. Melchiori. Flexible semantic-based service matchmaking and discovery. *World Wide Web*, 11(2):227–251, 2008.
- [Bibel 82]
W. Bibel. *Automated Theorem Proving*. Vieweg, Braunschweig, 1982.
- [Biber 89]
D. Biber. A typology of English texts. *Linguistics*, 27:3–43, 1989.
- [Biber 95]
D. Biber. *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge University Press, Cambridge, UK, 1995.
- [Biermann & Feldman 72]
A. Biermann and J. Feldman. A Survey of Results in Grammatical Inference. In S. Watanabe, editor, *Frontiers of Pattern Recognition*. Academic Press, New York/London, 1972.
- [Biewald 12]
L. Biewald. Massive multiplayer human computation for fun, money, and survival. In *Current Trends in Web Engineering*, pages 171–176. Springer, 2012.
- [Bikakis *et al.* 10]
N. Bikakis, G. Giannopoulos, T. Dalamagas, and T. Sellis. Integrating keywords and semantics on document annotation and search. In R. Meersman, T. Dillon, and P. Herrero, editors, *On the Move to Meaningful Internet Systems*, volume 6427, pages 921–938. Springer, 2010.
- [Bikel 02]
D. M. Bikel. Design of a multi-lingual, parallel-processing statistical parsing engine. In *Proceedings of the second international conference on Human Language Technology Research, HLT '02*, pages 178–182, San Francisco, CA, USA, 2002. Morgan Kaufmann Publishers Inc.

- [Bikel *et al.* 97]
 D. Bikel, S. Miller, R. Schwartz, and R. Weischedel. Nymble: a High-Performance Learning Name-finder. In *Proceedings of the Fifth conference on Applied Natural Language Processing*, 1997.
- [Bikel *et al.* 99]
 D. Bikel, R. Schwartz, and R. Weischedel. An Algorithm that Learns What’s in a Name. *Machine Learning, Special Issue on Natural Language Learning*, 34(1-3), Feb. 1999.
- [Bingham & Mannila 01]
 E. Bingham and H. Mannila. Random projection in dimensionality reduction: applications to image and text data. In *Knowledge Discovery and Data Mining*, pages 245–250, 2001.
- [Binot & Jensen 87]
 J. Binot and K. Jensen. A Semantic Expert Using an Online Standard Dictionary. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence (IJCAI-87)*, pages 709–714, Milan, Italy, 1987.
- [Bird & Liberman 98]
 S. Bird and M. Liberman. Towards a Formal Framework for Linguistic Annotation. In *Proceedings of the ICLSP*, Sydney, December 1998.
- [Bird & Liberman 99a]
 S. Bird and M. Liberman. A Formal Framework for Linguistic Annotation. Technical Report MS-CIS-99-01, Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA, 1999. <http://xxx.lanl.gov/abs/cs.CL/9903003>.
- [Bird & Liberman 99b]
 S. Bird and M. Liberman. Annotation graphs as a framework for multidimensional linguistic data analysis. In *Towards Standards and Tools for Discourse Tagging, Proceedings of the Workshop ACL-99*, pages 1–10, 1999.
- [Bird & Simons 01]
 S. Bird and G. Simons. The olac metadata set and controlled vocabularies. In *Proceedings of the ACL 2001 Workshop on Sharing Tools and Resources*, pages 27–38, 2001.
- [Bird *et al.* 00a]
 S. Bird, P. Buneman, and W. Tan. Towards a query language for annotation graphs. In *Proceedings of the Second International Conference on Language Resources and Evaluation*, Athens, 2000.
- [Bird *et al.* 00b]
 S. Bird, D. Day, J. Garofolo, J. Henderson, C. Laprun, and M. Liberman. ATLAS: A flexible and extensible architecture for linguistic annotation. In *Proceedings of the second International Conference on Language Resources and Evaluation (LREC 2000), 30 May – 2 Jun 2000*, pages 1699–1706, Athens, 2000.
- [Bird *et al.* 02]
 S. Bird, H. Uszkoreit, and G. Simons. The open language archives community. In *Proceedings of 3rd Language Resources and Evaluation Conference (LREC’2002)*, Gran Canaria, Spain, 2002.
- [Bird *et al.* 06]
 S. Bird, Y. Chen, S. Davidson, H. Lee, and Y. Zheng. Designing and evaluating an xpath dialect for linguistic queries. In *Data Engineering, 2006. ICDE ’06. Proceedings of the 22nd International Conference on*, pages 52–52, 2006.
- [Bird *et al.* 09]
 S. Bird, E. Klein, and E. Loper. *Natural language processing with Python*. " O’Reilly Media, Inc.", 2009.

- [Bizer *et al.* 09a]
 C. Bizer, T. Heath, and T. Berners-Lee. Linked Data – The Story So Far. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 5(3):1–22, Mar 2009.
- [Bizer *et al.* 09b]
 C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann. DBpedia – a crystallization point for the web of data. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 7:154–165, 2009.
- [Black & Cunningham 93]
 W. Black and H. Cunningham. The COBALT Language Analysis Module. Technical Report C/CCL/04/93/a, Centre for Computational Linguistics, UMIST, Manchester, 1993.
- [Black & Rinaldi 00]
 W. Black and F. Rinaldi. Facile pre-processor v3.0 - a user guide. Technical report, Department of Language Engineering, UMIST, 2000.
- [Black 90]
 W. J. Black. Knowledge based abstracting. *Online Review*, 14(5):327–340, 1990.
- [Black *et al.* 92]
 W. Black, N. Underwood, and H. Cunningham. Dialogue Management in a Pragmatics-Based Natural Language Information System. In T. McEnery and C. Paice, editors, *British Computer Society 14th Information Retrieval Colloquium*, Lancaster, 1992. Springer-Verlag.
- [Black *et al.* 93a]
 E. Black, R. Garside, and G. Leech, editors. *Statistically-driven computer grammars of English*. Number 8 in Language and Computers. Amsterdam, 1993.
- [Black *et al.* 93b]
 W. Black, H. Cunningham, B. Geurts, and G. Rentier. The PLUS Natural Language Engine. Technical Report ESPRIT Project P5254 (PLUS) Deliverable 2, Centre for Computational Linguistics, UMIST, Manchester, 1993.
- [Black *et al.* 98]
 W. Black, F. Rinaldi, and D. Mowatt. Facile: Description of the named entity system used for muc-7. In *Proceedings of the 7th MUC*, 1998.
- [Black *et al.* 99]
 B. Black, J. McNaught, F. Rinaldi, M. Ferraro, L. Gilardoni, S. Mazza, G. Zarri, A. Brasher, and A. Persidis. Detailed specification of the text extraction and concept recognition components of the concerto architecture. Deliverable 6, version 1.2, CONCERTO Consortium, 1999.
- [Blair & Maron 90]
 D. Blair and M. Maron. Full-text information retrieval: Further analysis and classification. *Information Processing and Management*, 26(3):437–477, 1990.
- [Blanco *et al.* 11]
 R. Blanco, P. Mika, and S. Vigna. Effective and efficient entity search in rdf data. In *Proceedings of the 10th International Conference on The Semantic Web, ISWC’11*, pages 83–97. Springer-Verlag, 2011.
- [Blanco *et al.* ss]
 R. Blanco, H. Halpin, D. M. Herzig, P. Mika, J. Pound, H. S. Thompson, and T. Tran. Repeatable and reliable semantic search evaluation. *Web Semantics: Science, Services and Agents on the World Wide Web*, In Press.

- [Blaschke *et al.* 99]
C. Blaschke, M. A. Andrade, C. Ouzounis, and A. Valencia. In *Proc Int Conf Intell Syst Mol Biol.*, pages 60–67, 1999.
- [Blei *et al.* 03]
D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3:993–1022, March 2003.
- [Blitzer *et al.* 07]
J. Blitzer, M. Dredze, and F. Pereira. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In *Annual Meeting-Association For Computational Linguistics*, page 440, 2007.
- [Blois 88]
M. S. Blois. Medicine and the nature of vertical reasoning. *New England Journal of Medicine*, 318(13):847–851, 1988.
- [Blum & Mitchell 98]
A. Blum and T. Mitchell. Combining labeled and unlabeled data with co-training. In *Proceedings of the Workshop on Computational Learning Theory*, 1998.
- [Blumer *et al.* 86]
A. Blumer, A. Ehrenfeucht, D. Haussler, and M. Warmuth. Occam’s Razor. Working paper ucsc-cr1-86-2, University of California Santa Cruz, Santa Cruz CA, 1986.
- [Blythe *et al.* 08]
J. Blythe, D. Kapoor, C. Knoblock, K. Lerman, and S. Minton. Information integration for the masses. *Journal of Universal Computer Science*, 14(11):1811–1837, 2008.
- [Boag *et al.* 10]
S. Boag, D. Chamberlin, M. F. Fernandez, D. Florescu, J. Robie, and J. Simeon. Xquery 1.0: An xml query language (second edition). Technical report, 14 December 2010.
- [Bobrow & Winograd 77]
D. Bobrow and T. Winograd. An Overview of KRL, a Knowledge Representation Language. *Cognitive Science* 1, pages 3–46, 1977.
- [Bod 95]
R. Bod. *Enriching Linguistics with Statistics: Performance Models of Natural Language*. Institute for Logic, Language and Computation, University of Amsterdam, 1995.
- [Bod 96]
R. Bod. Two Questions about Data-Oriented Parsing. In *Proceedings of the Fourth Workshop on Very Large Corpora*, Copenhagen, 1996.
- [Boden 90]
M. Boden, editor. *The Philosophy of Artificial Intelligence*. Oxford University Press, Oxford, 1990.
- [Bogers & van denBosch 11]
T. Bogers and A. van den Bosch. Fusing recommendations for social bookmarking websites. *International Journal of Electronic Commerce*, 15(3):33–75, Spring 2011.
- [Boguraev & Briscoe 87]
B. Boguraev and T. Briscoe. Large Lexicons for Natural Language Processing: Exploring the Grammar Coding System of LDOCE. *Computational Linguistics*, 13, 1987.
- [Boguraev & Briscoe 89]
B. Boguraev and T. Briscoe, editors. *Computational Lexicography for Natural Language Processing*. Longman, London, 1989.

- [Boguraev & Kennedy 97]
 B. Boguraev and C. Kennedy. Technical terminology for domain specification and content characterisation. In M. Pazienza, editor, *Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology*, Lecture Notes in Artificial Intelligence No. 1299. Springer, 1997.
- [Boguraev & Kennedy 99]
 B. Boguraev and C. Kennedy. Applications of term identification technology: domain description and content characterisation. *Natural Language Engineering*, 5(1):17–44, 1999.
- [Boguraev 03]
 B. K. Boguraev. Annotation-based finite state processing in a large-scale nlp architecture. In *Proceedings of the International Conference (RANLP-2003)*, pages 61–80, Borovets, Bulgaria, 2003.
- [Boguraev et al. 87]
 B. K. Boguraev, T. Briscoe, J. Carroll, D. Carter, and C. Grover. The Derivation of a Grammatically Indexed Lexicon from the Longman Dictionary of Contemporary English. In *Proceedings of the 25th Annual Meeting of the ACL*, pages 193–200, Stanford, CA, 1987.
- [Boguraev et al. 95]
 B. Boguraev, R. Garigliano, and J. Tait. Editorial. *Natural Language Engineering.*, 1, Part 1., 1995.
- [Boitet & Seligman 94]
 C. Boitet and M. Seligman. The “Whiteboard” Architecture: A Way to Integrate Heterogeneous Components of NLP Systems. In *Proceedings of COLING '94*, pages 426–430, Kyoto, Japan, 1994.
- [Boiy & Moens 09]
 E. Boiy and M.-F. Moens. A machine learning approach to sentiment analysis in multilingual web texts. *Information Retrieval*, 12(5):526–558, 2009.
- [Boiy et al. 07]
 E. Boiy, P. Hens, K. Deschacht, and M.-F. Moens. Automatic sentiment analysis of on-line text. In *Proc. of the 11th International Conference on Electronic Publishing*, Vienna, Austria, 2007.
- [Boldi & Vigna 05a]
 P. Boldi and S. Vigna. Compressed perfect embedded skip lists for quick inverted-index lookups. In M. Consens and G. Navarro, editors, *String Processing and Information Retrieval*, volume 3772 of *Lecture Notes in Computer Science*, pages 25–28. Springer Berlin / Heidelberg, 2005. 10.1007/11575832_3.
- [Boldi & Vigna 05b]
 P. Boldi and S. Vigna. MG4J at TREC 2005. In E. M. Voorhees and L. P. Buckland, editors, *Proceedings of the Fourteenth Text REtrieval Conference (TREC 2005), 15–18 November 2005*, volume 500 of *Special Publications*, pages 266–271. NIST, 2005. <http://mg4j.dsi.unimi.it/>.
- [Boldi & Vigna 06]
 P. Boldi and S. Vigna. Efficient lazy algorithms for minimal-interval semantics. In F. Crestani, P. Ferragina, and M. Sanderson, editors, *String Processing and Information Retrieval*, volume 4209 of *Lecture Notes in Computer Science*, pages 134–149. Springer Berlin / Heidelberg, 2006. 10.1007/11880561_12.
- [Bollacker et al. 08]
 K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor. Freebase: a collaboratively created graph database for structuring human knowledge. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pages 1247–1250. ACM, 2008.
- [Bollegara et al. 06]
 D. Bollegara, Y. Matsuo, and M. Ishizuka. Disambiguating Personal Names on the Web using Automatically Extracted Key Phrases. In *Proc. 17th European Conference on Artificial Intelligence (ECAI-06)*, 2006.

- [Bollen & Mao 11]
J. Bollen and H. Mao. Twitter mood as a stock market predictor. *IEEE Computer*, 44(10):91–94, 2011.
- [Bollen *et al.* 09]
J. Bollen, A. Pepe, and H. Mao. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. <http://arxiv.org/abs/0911.1583>, 2009.
- [Bono & Ficorilli 92]
G. Bono and P. Ficorilli. Natural Language Restatement of Queries Expressed in a Graphical Language. In *ER'92- 11th International Conference on the Entity-Relational Approach. LNCS 645*, pages 357–274. Springer-Verlag, 1992.
- [Bontcheva & Cunningham 03]
K. Bontcheva and H. Cunningham. The semantic web: A new opportunity and challenge for human language technology. In H. Cunningham, Y. Ding, and A. Kiryakov, editors, *Proceedings of Workshop on Human Language Technology for the Semantic Web and Web Services, 2nd International Semantic Web Conference*, Sanibel Island, Florida, Oct 2003. <http://www.gate.ac.uk/sale/iswc03/iswc03.pdf>.
- [Bontcheva & Cunningham 11a]
K. Bontcheva and H. Cunningham. Semantic annotation and retrieval: Manual, semi-automatic and automatic generation. In J. Domingue, D. Fensel, and J. A. Hendler, editors, *Handbook of Semantic Web Technologies*. Springer, 2011.
- [Bontcheva & Cunningham 11b]
K. Bontcheva and H. Cunningham. Semantic annotations and retrieval: Manual, semiautomatic, and automatic generation. In J. Domingue, D. Fensel, and J. Hendler, editors, *Handbook of Semantic Web Technologies*, pages 77–116. Springer, 2011.
- [Bontcheva & Davis 09]
K. Bontcheva and B. Davis. Natural Language Generation from Ontologies. In J. Davies, M. Grobelnik, and D. Mladenic, editors, *Semantic Knowledge Management*, pages 117–130. 2009.
- [Bontcheva & Derczynski 13]
K. Bontcheva and L. Derczynski. TwitIE: An open-source information extraction pipeline for microblog text: slides. <http://dx.doi.org/10.6084/m9.figshare.1003767>, 2013.
- [Bontcheva & Dimitrova 04]
K. Bontcheva and V. Dimitrova. Examining the Use of Conceptual Graphs in Adaptive Web-Based Systems that Aid Terminology Learning. *International Journal on Artificial Intelligence Tools – Special issue on AI Techniques in Web-Based Educational Systems*, 13(2), 2004.
- [Bontcheva & Rout 14]
K. Bontcheva and D. Rout. Making sense of social media through semantics: A survey. *Semantic Web - Interoperability, Usability, Applicability*, 5(5):373–403, 2014.
- [Bontcheva & Sabou 06]
K. Bontcheva and M. Sabou. Learning Ontologies from Software Artifacts: Exploring and Combining Multiple Sources. In *Workshop on Semantic Web Enabled Software Engineering (SWESE)*, Athens, G.A., USA, November 2006.
- [Bontcheva & Wilks 01]
K. Bontcheva and Y. Wilks. Dealing with Dependencies between Content Planning and Surface Realisation in a Pipeline Generation Architecture. In *Proceedings of the International Joint Conference in Artificial Intelligence (IJCAI'2001)*, Seattle, USA, August 2001.

- [Bontcheva & Wilks 04a]
K. Bontcheva and Y. Wilks. Automatic Report Generation from Ontologies: the MIAKT approach. In *Nineth International Conference on Applications of Natural Language to Information Systems (NLDB'2004)*, Manchester, UK, 2004.
- [Bontcheva & Wilks 04b]
K. Bontcheva and Y. Wilks. Tailoring Automatically Generated Hypertext. *User Modeling and User-Adapted Interaction*, 2004. Special issue on Language-Based Interaction.
- [Bontcheva 97]
K. Bontcheva. Generation of multilingual explanations from conceptual graphs. In R. Mitkov and N. Nicolov, editors, *Recent Advances in Natural Language Processing: Selected Papers from RANLP'95*, volume 136 of *Current Issues in Linguistic Theory (CILT)*, pages 365 – 376. John Benjamins, Amsterdam/Philadelphia, 1997.
- [Bontcheva 01a]
K. Bontcheva. *Generating Adaptive Hypertext*. Unpublished PhD thesis, University of Sheffield, UK, 2001.
- [Bontcheva 01b]
K. Bontcheva. Tailoring the Content of Dynamically Generated Explanations. In M. Bauer, P. Gmytrasiewicz, and J. Vassileva, editors, *User Modelling 2001*, volume 2109 of *Lecture Notes in Artificial Intelligence*. Springer, Berlin, 2001.
- [Bontcheva 02]
K. Bontcheva. Adaptivity, adaptability, and reading behaviour: Some results from the evaluation of a dynamic hypertext system. In P. De Bra, P. Brusilovsky, and R. Conejo, editors, *Adaptive Hypermedia and Adaptive Web-Based Systems*, volume 2347 of *Lecture Notes in Computer Science*. Springer Verlag, Berlin Heidelberg, 2002.
- [Bontcheva 03]
K. Bontcheva. Reuse and Problems in the Evaluation of NLG systems. In *Proceedings of EACL'03 Workshop on Evaluation Initiatives*, Budapest, Hungary, 2003.
- [Bontcheva 04a]
K. Bontcheva. Open-source Tools for Creation, Maintenance, and Storage of Lexical Resources for Language Generation from Ontologies. In *Proceedings of 4th Language Resources and Evaluation Conference (LREC'04)*, 2004.
- [Bontcheva 04b]
K. Bontcheva. Open-source Tools for Creation, Maintenance, and Storage of Lexical Resources for Language Generation from Ontologies. In *Proceedings of 4th Language Resources and Evaluation Conference (LREC'04)*, 2004. <http://www.dcs.shef.ac.uk/kalina/papers/bontcheva-lrec04.pdf>.
- [Bontcheva 05]
K. Bontcheva. Generating Tailored Textual Summaries from Ontologies. In *Second European Semantic Web Conference (ESWC'2005)*, 2005.
- [Bontcheva et al. 00]
K. Bontcheva, H. Brugman, A. Russel, P. Wittenburg, and H. Cunningham. An Experiment in Unifying Audio-Visual and Textual Infrastructures for Language Processing R&D. In *Proceedings of the Workshop on Using Toolsets and Architectures To Build NLP Systems at COLING-2000*, Luxembourg, 2000. <http://gate.ac.uk/>.
- [Bontcheva et al. 01]
K. Bontcheva, C. Brewster, F. Ciravegna, H. Cunningham, L. Guthrie, R. Gaizauskas, and Y. Wilks.

Using HLT for Acquiring, Retrieving and Publishing Knowledge in AKT: Position Paper. In *Workshop on Human Language Technology and Knowledge Management*, Toulouse, France, July 2001. <http://www.elsnet.org/ac12001-hlt+km.html>.

[Bontcheva *et al.* 02a]

K. Bontcheva, H. Cunningham, V. Tablan, D. Maynard, and O. Hamza. Using GATE as an Environment for Teaching NLP. In *Proceedings of the ACL Workshop on Effective Tools and Methodologies in Teaching NLP*, 2002. <http://gate.ac.uk/sale/ac102/gate4teaching.pdf>.

[Bontcheva *et al.* 02b]

K. Bontcheva, H. Cunningham, V. Tablan, D. Maynard, and H. Saggion. Developing Reusable and Robust Language Processing Components for Information Systems using GATE. In *Proceedings of the 3rd International Workshop on Natural Language and Information Systems (NLIS'2002)*, Aix-en-Provence, France, 2002. IEEE Computer Society Press. <http://gate.ac.uk/sale/nlis/nlis.ps>.

[Bontcheva *et al.* 02c]

K. Bontcheva, M. Dimitrov, D. Maynard, V. Tablan, and H. Cunningham. Shallow Methods for Named Entity Coreference Resolution. In *Chaînes de références et résolveurs d'anaphores, workshop TALN 2002*, Nancy, France, 2002. <http://gate.ac.uk/sale/taln02/taln-ws-coref.pdf>.

[Bontcheva *et al.* 02d]

K. Bontcheva, D. Maynard, H. Cunningham, and H. Saggion. Using Human Language Technology for Automatic Annotation and Indexing of Digital Library Content. In *Proceedings of the 6th European Conference on Research and Advanced Technology for Digital Libraries (ECDL'2002)*, Rome, Italy, 2002. <http://gate.ac.uk/sale/ecdl02/ecdl.pdf>.

[Bontcheva *et al.* 02e]

K. Bontcheva, D. Maynard, H. Cunningham, and H. Saggion. Using human language technology for automatic annotation and indexing of digital library content. In M. Agosti and C. Thanos, editors, *Research and Advanced Technology for Digital Libraries, Springer Lecture Notes in Computer Science, 2458*. Springer, Berlin/New York, 2002.

[Bontcheva *et al.* 03a]

K. Bontcheva, D. Maynard, V. Tablan, and H. Cunningham. GATE: A Unicode-based infrastructure supporting multilingual information extraction. In *Proceedings of Workshop on Information Extraction for Slavonic and other Central and Eastern European Languages (IESL'03)*, Borovets, Bulgaria, 2003.

[Bontcheva *et al.* 03b]

K. Bontcheva, A. Kiryakov, H. Cunningham, B. Popov, and M. Dimitrov. Semantic web enabled, open source language technology. In *EACL workshop on Language Technology and the Semantic Web: NLP and XML*, Budapest, Hungary, 2003. <http://gate.ac.uk/sale/eac103-semweb/bontcheva-et-al-final.pdf>.

[Bontcheva *et al.* 04]

K. Bontcheva, V. Tablan, D. Maynard, and H. Cunningham. Evolving GATE to Meet New Challenges in Language Engineering. *Natural Language Engineering*, 10(3/4):349—373, 2004.

[Bontcheva *et al.* 06a]

K. Bontcheva, H. Cunningham, A. Kiryakov, and V. Tablan. Semantic Annotation and Human Language Technology. In J. Davies, R. Studer, and P. Warren, editors, *Semantic Web Technology: Trends and Research*. John Wiley and Sons, 2006.

[Bontcheva *et al.* 06b]

K. Bontcheva, J. Davies, A. Duke, T. Glover, N. Kings, and I. Thurlow. Semantic Information Access. In J. Davies, R. Studer, and P. Warren, editors, *Semantic Web Technologies*. John Wiley and Sons, 2006.

- [Bontcheva *et al.* 07a]
 K. Bontcheva, D. Damljanovic, N. Aswani, M. Agatonovic, J. Sun, and F. Amardeilh. Key concept identification and clustering of similar content. Technical Report D3.1, TAO Project Deliverable, 2007. <http://www.tao-project.eu/resources/publicdeliverables/d3-1.pdf>.
- [Bontcheva *et al.* 07b]
 K. Bontcheva, I. Roberts, M. Agatonovic, J. Nioche, and J. Sun. Case study 1: Requirement analysis and application of tao methodology in data intensive applications. Technical Report D6.1, TAO Project Deliverable, 2007. <http://www.tao-project.eu/resources/publicdeliverables/d6-1.pdf>.
- [Bontcheva *et al.* 09]
 K. Bontcheva, B. Davis, A. Funk, Y. Li, and T. Wang. Human Language Technologies. In J. Davies, M. Grobelnik, and D. Mladenic, editors, *Semantic Knowledge Management*, pages 37–49. 2009.
- [Bontcheva *et al.* 10]
 K. Bontcheva, H. Cunningham, I. Roberts, and V. Tablan. Web-based collaborative corpus annotation: Requirements and a framework implementation. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, 17–23 May 2010*, pages 20–27, Valletta, Malta, May 2010.
- [Bontcheva *et al.* 12]
 K. Bontcheva, J. Kieniewicz, N. Aswani, M. Wallis, and S. Andrews. User feedback report on the envilod semantic search interface. Technical Report <http://gate.ac.uk/projects/envilod/EnviLOD-user-feedback-report.pdf>, EnviLOD project deliverable, 2012.
- [Bontcheva *et al.* 13a]
 K. Bontcheva, H. Cunningham, I. Roberts, A. Roberts, V. Tablan, N. Aswani, and G. Gorrell. GATE Teamware: A Web-based, Collaborative Text Annotation Framework. *Language Resources and Evaluation*, 47:1007–1029, 2013.
- [Bontcheva *et al.* 13b]
 K. Bontcheva, L. Derczynski, A. Funk, M. A. Greenwood, D. Maynard, and N. Aswani. TwitIE: An Open-Source Information Extraction Pipeline for Microblog Text. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing*. Association for Computational Linguistics, 2013.
- [Bontcheva *et al.* 13c]
 K. Bontcheva, G. Gorrell, and B. Wessels. Social media and information overload: Survey results. Technical Report 1306.0813 [cs.SI], arXiv, 2013. <http://arxiv.org/abs/1306.0813>.
- [Bontcheva *et al.* 14a]
 K. Bontcheva, L. Derczynski, and I. Roberts. Crowdsourcing named entity recognition and entity linking corpora. *The Handbook of Linguistic Annotation (to appear)*, 2014.
- [Bontcheva *et al.* 14b]
 K. Bontcheva, I. Roberts, L. Derczynski, and D. Rout. The GATE Crowdsourcing Plugin: Crowdsourcing Annotated Corpora Made Easy. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. Association for Computational Linguistics, 2014.
- [Bontcheva *et al.* 14c]
 K. Bontcheva, V. Tablan, and H. Cunningham. Semantic search over documents and ontologies. In *Bridging Between Information Retrieval and Databases*, volume 8173, pages 31–53. Springer Verlag, 2014.
- [Bontcheva *et al.* 15]
 K. Bontcheva, J. Kieniewicz, S. Andrews, and M. Wallis. Semantic Enrichment and Search: A Case Study on Environmental Science Literature. *D-Lib Magazine*, 21(1/2), 2015.

- [Booch 94]
G. Booch. *Object-Oriented Analysis and Design 2nd Edn.* Benjamin/Cummings, 1994.
- [Booch *et al.* 99]
G. Booch, J. Rumbaugh, and I. Jacobson. *The Unified Modelling Language User Guide.* Addison-Wesley, Reading, MA, 1999.
- [Boolos & Jeffrey 80]
G. Boolos and R. Jeffrey. *Computability and Logic 2nd ed.* Cambridge University Press, Cambridge, 1980.
- [Borgo *et al.* 01]
R. Borgo, P. Cignoni, and R. Scopigno. An Easy-to-use Visualization System for Huge Cultural Heritage Meshes. In *Proceedings of the VAST'01 Conference*, Greece, November 2001. cite-seer.nj.nec.com/509560.html.
- [Borthwick *et al.* 98]
A. Borthwick, J. Sterling, E. Agichtein, and R. Grishman. Description of the Mene named entity system as used in MUC-7. In *Proceedings of the MUC-7 Conference*, NYU, 1998.
- [Bos 08]
J. Bos. Wide-coverage semantic analysis with boxer. In *Proceedings of the 2008 Conference on Semantics in Text Processing*, pages 277–286. Association for Computational Linguistics, 2008.
- [Bos *et al.* 98]
J. Bos, C. Rupp, B. Buschbeck-Wolf, and M. Dorna. Managing information at linguistic interfaces. In *Proceedings of the 36th ACL and the 17th COLING (ACL-COLING '98)*, pages 160–166, Montreal, 1998.
- [Bosma & Vossen 10]
W. Bosma and P. Vossen. Bootstrapping language-neutral term extraction. In *7th Language Resources and Evaluation Conference (LREC)*, Valletta, Malta, 2010.
- [Bossard 93]
S. Bossard. DGXIII study ML77-A. In *Language Industries Atlas*. IOS Press, 1993.
- [Botley & McEnery 97]
S. Botley and T. McEnery. *Discourse Anaphora and Anaphor Resolution*. University College London Press, 1997.
- [Bouma 09]
G. Bouma. Normalized (pointwise) mutual information in collocation extraction. In *From Form to Meaning: Processing Texts Automatically, Proceedings of the Biennial GSCL Conference 2009*, volume Normalized, pages 31–40, Tübingen, 2009.
- [Bourigault & Habert 98]
D. Bourigault and B. Habert. Evaluation of terminology extractors: Principles and experiments. In *Proc. of 1st International Conference on Language Resources and Evaluation (LREC)*, pages 299–305, Granada, Spain, 1998.
- [Bourigault 92]
D. Bourigault. Surface grammatical analysis for the extraction of terminological noun phrases. In *Proc. of 14th International Conference on Computational Linguistics (COLING)*, pages 977–981, Nantes, France, 1992.
- [Bourke *et al.* 13]
S. Bourke, M. P. O'Mahony, R. Rafter, and B. Smyth. Ranking in information streams. In J. Kim, J. Nichols, and P. A. Szekely, editors, *18th International Conference on Intelligent User Interfaces*,

IUI '13, Santa Monica, CA, USA, March 19-22, 2013, Companion Volume, pages 99–100. ACM, 2013.

[Bourlard & Morgan 90]

H. Bourlard and N. Morgan. A continuous speech recognition system embedding MLP into HMM. In D. S. Touretzky, editor, *Advances in Neural Information Processing Systems*, volume 2, pages 413–416. Morgan Kaufmann, California, San Mateo CA, 1990.

[Bourlard & Morgan 91]

H. Bourlard and N. Morgan. Merging multilayer perceptrons and hidden Markov models: some experiments in continuous speech recognition. In E. Gelenbe, editor, *Neural Networks: Advances and Applications*, pages 215–239. North-Holland, 1991.

[Bourlard & Morgan 93]

H. Bourlard and N. Morgan. Continuous speech recognition by connectionist statistical methods. *IEEE Transactions on Neural Networks*, 4:893–909, 1993.

[Bourlard & Morgan 94]

H. Bourlard and N. Morgan. *Connectionist Speech Recognition—A Hybrid Approach*. Kluwer Academic, Amsterdam, 1994.

[Bourlard & Wellekens 89a]

H. Bourlard and C. J. Wellekens. Links between Markov models and multi-layer perceptrons. In D. S. Touretzky, editor, *Advances in Neural Information Processing Systems*, volume 1, pages 502–510. Morgan Kaufmann, California, San Mateo CA, 1989.

[Bourlard & Wellekens 89b]

H. Bourlard and C. J. Wellekens. Speech dynamics and recurrent neural networks. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 33–36, Glasgow, 1989.

[Bourlard & Wellekens 89c]

H. Bourlard and C. J. Wellekens. Speech pattern discrimination and multilayer perceptrons. *Computer Speech and Language*, 3:1–19, 1989.

[Bourlard & Wellekens 90]

H. Bourlard and C. J. Wellekens. Links between Markov models and multilayer perceptrons. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-12:1167–1178, 1990.

[Bourlard *et al.* 85]

H. Bourlard, Y. Kamp, H. Ney, and C. J. Wellekens. Speaker-dependent connected speech recognition via dynamic programming and statistical methods. In M. R. Schroeder, editor, *Speech and Speaker Recognition*, volume 12 of *Bibliotheca Phonetica*, pages 115–148. Karger, Basel, 1985.

[Bourlard *et al.* 91]

H. Bourlard, N. Morgan, and C. Wooters. Connectionist approaches to the use of Markov models for speech recognition. In R. P. Lippmann, J. E. Moody, and D. S. Touretzky, editors, *Advances in Neural Information Processing Systems*, volume 3, pages 213–219. Morgan Kaufmann, California, San Mateo CA, 1991.

[Bourlard *et al.* 92]

H. Bourlard, N. Morgan, C. Wooters, and S. Renals. CDNN: A context dependent neural network for continuous speech recognition. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 2, pages 349–352, San Francisco, 1992.

[Bourlard *et al.* 94]

H. Bourlard, B. D’hoore, and J.-M. Boite. Optimizing recognition and rejection performance in wordspotting systems. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 1, pages 373–376, Adelaide, 1994.

- [Bousfield 07]
D. Bousfield. Never a truer word said in jest: A pragmastylistic analysis of impoliteness as banter in Henry IV, Part I. *Contemporary Stylistics*, pages 195–208, 2007.
- [Bowden *et al.* 96]
P. R. Bowden, P. Halstead, and T. Rose. Extracting conceptual knowledge from text using explicit relation markers. In *Proc. 9th European Knowledge Acquisition Workshop (EKAW-96)*, pages 147–162, Nottingham, UK, 1996.
- [Bowden *et al.* 98]
P. R. Bowden, L. Evett, and P. Halstead. Automatic acronym acquisition in a knowledge extraction program. In *Proc. of 1st Workshop on Computational Terminology (Computerm '98)*, pages 43–49, Montreal, Canada, 1998. COLING-ACL.
- [Bowker 95]
L. Bowker. *A multidimensional approach to classification in Terminology: Working with a computational framework*. Unpublished PhD thesis, University of Manchester, England, 1995.
- [Bowker 97]
L. Bowker. Multidimensional classification of concepts and terms. In S. Wright and G. Budin, editors, *Handbook of Terminology Management*, volume 1, pages 133–143. John Benjamins, Amsterdam, 1997.
- [Bowman *et al.* 97]
C. Bowman, D. Michaud, and H. Suonuuti. Do's and don'ts of terminology management. In S. Wright and G. Budin, editors, *Handbook of Terminology Management*, volume 1: Basic Aspects of Terminology Management, chapter 2.2.4, pages 215–217. John Benjamins, Amsterdam, 1997.
- [Boyd *et al.* 10]
D. Boyd, S. Golder, and G. Lotan. Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on System Sciences*, pages 1–10. IEEE, 2010.
- [Boyer & Moore 79]
R. Boyer and J. Moore. *A Computational Logic*. Academic Press, 1979.
- [Brachman & Levesque 84]
R. Brachman and H. Levesque. The tractability of subsumption in frame-based description languages. *Proceedings of the Ninth Conference on Artificial Intelligence*, 1984.
- [Brachman & Levesque 85]
R. Brachman and H. Levesque. *Readings in Knowledge Representation*. Morgan Kaufmann, California, Los Altos, CA, 1985.
- [Brachman & Schmolze 85]
R. Brachman and J. Schmolze. An overview of the KL-ONE knowledge representation system. *Cognitive Science*, (9):171 – 216, 1985.
- [Braden-Hader 93]
L. Braden-Hader. Sense disambiguation using online dictionaries. In K. Jenon, G. Heidhorn, and S. Richardson, editors, *Natural Language Processing: the PLNLP Approach*, pages 247–261. Kluwer Academic Publishers, Dordrecht, 1993.
- [Brady & Berwick 83]
M. Brady and R. Berwick, editors. *Computational Models of Discourse*. MIT Press, Cambridge, MA, 1983.

- [Branavan *et al.* 09]
 S. R. K. Branavan, H. Chen, J. Eisenstein, and R. Barzilay. Learning document-level semantic properties from free-text annotations. *Journal of Artificial Intelligence Research*, 34(1):569–603, April 2009.
- [Brand 94]
 S. Brand. *How Buildings Learn*. Penguin, London, 1994.
- [Brandow *et al.* 95]
 R. Brandow, K. Mitze, and L. F. Rau. Automatic condensation of electronic publications by sentence selection. *Information Processing and Management*, 31(5):675–685, September 1995.
- [Brants 00]
 T. Brants. Tnt: a statistical part-of-speech tagger. In *Proceedings of the sixth conference on Applied Natural Language Processing, ANLP '00*, pages 224–231, 2000.
- [Bratko 90]
 I. Bratko. *Prolog Programming for Artificial Intelligence*. Addison-Wesley, Reading, MA, second edition, 1990.
- [Breiman 01]
 L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.
- [Breiman *et al.* 84]
 L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen. *Classification and regression trees*. CRC press, 1984.
- [Breu *et al.* 97]
 M. Breu, A. Brüggemann-Klein, C. Haber, and R. Weber. The MeDoc Distributed Electronic Library – Accounting and Security Aspects. In *Proceedings of the ICC/IFIP Conference on Electronic Publishing, 1997*. citeseer.nj.nec.com/article/breu97medoc.html.
- [Brew & McKelvie 96]
 C. Brew and D. McKelvie. Word pair extraction for lexicography. In *Proc. of the Second International Conference on New Methods in Language Processing*, Ankara, 1996.
- [Brew *et al.* 99]
 C. Brew, D. McKelvie, R. Tobin, H. Thompson, and A. Mikheev. *The XML Library LT XML version 1.1 User documentation and reference guide*. Language Technology Group, Edinburgh, 1999. <http://www.ltg.ed.ac.uk/>.
- [Brew *et al.* 10]
 A. Brew, D. Greene, and P. Cunningham. Using crowdsourcing and active learning to track sentiment in online media. In *Proceedings of the 19th European Conference on Artificial Intelligence*, pages 145–150, 2010.
- [Brewka 91]
 G. Brewka. *Nonmonotonic Reasoning: Logical Foundations of Commonsense*. Cambridge University Press, Cambridge, 1991.
- [Brewster *et al.* 02]
 C. Brewster, F. Ciravegna, and Y. Wilks. User-centred ontology learning for knowledge management. In *Proceedings of the 7th International Conference on Applications of Natural Language to Information Systems*, volume 2553 of *Lecture Notes in Computer Sciences*, Springer Verlag, Stockholm, June 2002. Springer-Verlag.

- [Brewster *et al.* 04]
 C. Brewster, H. Alani, S. Dasmahapatra, and Y. Wilks. Data Driven Ontology Evaluation. In *Proceedings of the Language Resources and Evaluation Conference (LREC 2004)*, Lisbon, Portugal, 2004.
- [Brickley & Guha 04]
 D. Brickley and R. V. Guha. RDF vocabulary description language 1.0: RDF schema. W3C recommendation, W3C, 2004. Available at <http://www.w3.org/TR/rdf-schema/>.
- [Bridle & Cox 91]
 J. S. Bridle and S. J. Cox. RecNorm: Simultaneous normalisation and classification applied to speech recognition. In R. P. Lippmann, J. E. Moody, and D. S. Touretzky, editors, *Advances in Neural Information Processing Systems*, volume 3, pages 234–240. Morgan Kaufmann, California, San Mateo CA, 1991.
- [Bridle & Dodd 91]
 J. S. Bridle and L. Dodd. An Alphanet approach to optimising input transformations for continuous speech recognition. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 277–280, Toronto, 1991.
- [Bridle & Sedgwick 77]
 J. S. Bridle and N. C. Sedgwick. A method for segmenting acoustic patterns, with applications to automatic speech recognition. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 656–660, Hartford CT, 1977.
- [Bridle 90a]
 J. S. Bridle. Alpha-nets: a recurrent neural network architecture with a hidden Markov model interpretation. *Speech Communication*, 9:83–92, 1990.
- [Bridle 90b]
 J. S. Bridle. Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimation of parameters. In D. S. Touretzky, editor, *Advances in Neural Information Processing Systems*, volume 2, pages 211–217. Morgan Kaufmann, California, San Mateo CA, 1990.
- [Brill 92a]
 E. Brill. A simple rule-based part-of-speech tagger. In *Proceedings of the Third Conference on Applied Natural Language Processing*, Trento, Italy, 1992.
- [Brill 92b]
 E. Brill. A simple rule-based part of speech tagger. In *Proceedings of the DARPA Speech and Natural Language Workshop*. Harriman, NY, 1992.
- [Brill 92c]
 E. Brill. A simple rule-based part of speech tagger. In *Proc. of 3rd Conference of Applied Natural Language Processing*, 1992.
- [Brill 94a]
 E. Brill. Some Advances in Transformation-Based Part of Speech Tagging. In *Proceedings of the Twelfth National Conference on AI (AAAI-94)*, Seattle, Washington, 1994.
- [Brill 94b]
 E. Brill. Some advances in transformation-based part of speech tagging. In *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI '94)*, pages 722–727, 1994.
- [Brill 94c]
 E. Brill. *A Corpus-Based Approach to Language Learning*. Unpublished PhD thesis, Dept. of Computer and Information Science, University of Pennsylvania, 1994.

- [Brill 94d]
E. Brill. Some advances in transformation-based part of speech tagging. In *Proceedings of the 12th National Conference on Artificial Intelligence (AAAI-94)*, pages 722–727, Seattle, WA, July 1994.
- [Brill 95a]
E. Brill. Transformation-Based Error-Driven Learning and Natural Language. *Computational Linguistics*, 21(4), December 1995.
- [Brill 95b]
E. Brill. Transformation-Based Error-Driven Learning and Natural Language Processing: A Case Study in Part-of-speech Tagging. *Computational Linguistics*, 21(4):543–565, 1995.
- [Brill 95c]
E. Brill. Transformation-based error-driven learning and natural language processing: A case study in part of speech tagging. *Computational Linguistics*, 1995.
- [Brin & Page 98]
S. Brin and L. Page. The anatomy of a large-scale hypertextual (web) search engine. In *Proc. 7th International World Wide Web Conference (WWW7)*, 1998.
- [Brin *et al.* 97]
S. Brin, R. Motwani, and C. Silverstein. Beyond market baskets: Generalizing association rules to correlations. In *ACM SIGMOD International Conference on Management of Data*, 1997.
- [Briscoe *et al.* 06]
T. Briscoe, J. Carroll, and R. Watson. The second release of the rasp system. In *Proceedings of the COLING/ACL on Interactive presentation sessions*, pages 77–80, 2006.
- [Brockmans *et al.* 08]
S. Brockmans, M. Erdmann, and W. Schoch. Hybrid matchmaker and Service-Finder ontologies (alpha release). Deliverable D4.2, Service-Finder Consortium, 2008.
- [Brockmans *et al.* 09]
S. Brockmans, I. Celino, D. Cerizza, E. D. Valle, M. Erdmann, A. Funk, H. Lausen, W. Schoch, N. Steinmetz, and A. Turati. Service-Finder: First steps toward the realization of web service discovery at web scale. In *Interoperability through Semantic Data and Service Integration Workshop (ISDSI 2009) at SEBD*, Genova, Italy, June 2009.
- [Brodie *et al.* 84]
M. Brodie, J. Mylopoulos, and J. Schmidt. *On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases, and Programming Languages*. Springer-Verlag, New York, 1984.
- [Broeder & Wittenburg 01]
D. Broeder and P. Wittenburg. Multimedia language resources. In *Proceedings of the ACL 2001 Workshop on Sharing Tools and Resources*, pages 47–51, 2001.
- [Broekstra & Kampman 03]
J. Broekstra and A. Kampman. SeRQL: A Second Generation RDF Query Language. In *In Proceedings of the SWAD-Europe Workshop on Semantic Web Storage and Retrieval*, pages 13–14, 2003.
- [Broekstra *et al.* 02]
J. Broekstra, A. Kampman, and F. Van Harmelen. Sesame: A generic architecture for storing and querying RDF and RDF Schema. In *The Semantic Web—ISWC 2002*. Springer, 2002.
- [Broglia *et al.* 95]
R. Broglia, J. Callan, W. Croft, and D. Nachbar. Document Retrieval and Routing Using the INQUERY System. In D. Harman, editor, *Overview of the Third retrieval Conference (TREC-3)*., 1995.

- [Brooke 96]
J. Brooke. SUS: a “Quick and Dirty” Usability Scale. In Jordan et al. [Jordan *et al.* 96].
- [Brown & *et al.* 90]
P. Brown et al. Class-based n-gram models of natural language. In *Proceedings of the IBM NLP Workshop*. Paris, France, 1990.
- [Brown 89]
A. Brown. *Database Support for Software Engineering*. Kogan Page, 1989.
- [Brown *et al.* 90]
P. Brown, J. Cocke, S. D. Pietra, V. D. Pietra, F. Jelinek, J. Lafferty, R. Mercer, and P. Roossin. A Statistical Approach to Machine Translation. *Computational Linguistics*, 16(June):79–85, 1990.
- [Brown *et al.* 91a]
P. F. Brown, S. A. Di Pietra, V. . Di Pietra, and R. L. Mercer. Word sense disambiguation using statistical methods. In *Proceedings of ACL-91*, 1991.
- [Brown *et al.* 91b]
P. F. Brown, A. Della Pietra, V. J. Della Pietra, and R. L. Mercer. A Statistical Approach to Sense Disambiguation in Machine Translation. In *Proceedings of the Speech and Natural Language Workshop*, pages 146–151, February 1991.
- [Brown *et al.* 92]
P. Brown, V. Della Pietra, P. de Souza, J. Lai, and R. Mercer. Class-based n-gram models of natural language. *Computational Linguistics*, 18:467–479, 1992.
- [Brown *et al.* 93]
P. F. Brown, S. A. Della Pietra, V. J. Della Pietra, and R. L. Mercer. The Mathematics of Machine Translation: Parameter Estimation. *Computational Linguistics*, 19(2):263–311, 1993.
- [Browne *et al.* 05]
A. C. Browne, G. Divita, C. Lu, L. McCreedy, and D. Nace. Lexical systems; a report to the board of scientific counselors. Technical report, National Library of Medicine, Bethesda, MD, 2005.
- [Brucato *et al.* 13]
M. Brucato, L. Derczynski, H. Llorens, K. Bontcheva, and C. S. Jensen. Recognising and interpreting named temporal expressions. In *RANLP*, pages 113–121, 2013.
- [Bruce & Guthrie 92a]
R. Bruce and L. Guthrie. An Automatically Generated Semantic Hierarchy. In *Proceedings of COLING '92*, Nantes, 1992.
- [Bruce & Guthrie 92b]
R. Bruce and L. Guthrie. Genus Disambiguation: A Study in Weighted Preference. In *Proceedings of COLING-92*, pages 1187–1191, Nantes, France, 1992.
- [Bruce & Wiebe 94a]
R. Bruce and J. Wiebe. Word-sense disambiguation using decomposable models. In *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, Las Cruces, New Mexico, 1994.
- [Bruce & Wiebe 94b]
R. Bruce and J. Wiebe. Word-sense disambiguation using decomposable models. In *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, pages 139–145, Las Cruces, New Mexico, 1994.

- [Bruce & Wiebe 98]
 R. Bruce and J. Wiebe. Word sense distinguishability and inter-coder agreement. In *Proceedings of 3rd Empirical Methods in Natural Language Processing (EMNLP-98)*. Association for Computational Linguistics SIGDAT, pages 53–60, Granada, Spain, 1998.
- [Bruce & Wiebe 99]
 R. Bruce and J. M. Wiebe. Recognizing Subjectivity: A Case Study of Manual Tagging. *Natural Language Engineering*, 1(1):1–16, 1999.
- [Brughman *et al.* 98]
 H. Brughman, A. Russel, P. Wittenburg, and R. Piepenbrock. Corpus-based research using the Internet. In *First International Conference on Language Resources and Evaluation (LREC) Workshop on Distributing and Accessing Linguistic Resources*, Granada, Spain, 1998.
- [Brugman & Russel 04]
 H. Brugman and A. Russel. Annotating multi-media/multi-modal resources with elan. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation*, 2004.
- [Brugman *et al.* 98a]
 H. Brugman, A. Russel, P. Wittenburg, and R. Piepenbrock. Corpus-based Research using the Internet. In *Workshop on Distributing and Accessing Linguistic Resources*, pages 8–15, Granada, Spain, 1998. <http://www.dcs.shef.ac.uk/~hamish/dalr/>.
- [Brugman *et al.* 98b]
 H. Brugman, H. Russel, and P. Wittenburg. An infrastructure for collaboratively building and using multimedia corpora in the humaniora. In *Proceedings of the ED-MEDIA/ED-TELECOM Conference*, Freiburg, 1998.
- [Brugman *et al.* 99]
 H. Brugman, K. Bontcheva, P. Wittenburg, and H. Cunningham. Integrating Multimedia and Textual Software Architectures for Language Technology. Technical report MPI-TG-99-1, Max-Planck Institute for Psycholinguistics, Nijmegen, Netherlands, 1999.
- [Brugman *et al.* 04]
 H. Brugman, O. Crasborn, and A. Russel. Collaborative annotation of sign language data with peer-to-peer technology. In *Proceedings of LREC*, 2004.
- [Bruijn & Polleres 04]
 J. Bruijn and A. Polleres. Towards An Ontology Mapping Specification Language For the Semantic Web. Technical report, Digital Enterprise Research Institute, 2004.
- [Brush 02]
 A. Brush. *Annotating Digital Documents for Asynchronous Collaboration*. Unpublished PhD thesis, Department of Computer Science and Engineering, University of Washington, Seattle, 2002.
- [Brusilovsky 96]
 P. Brusilovsky. Methods and techniques of adaptive hypermedia. *User Modelling and User-Adapted Interaction*, 6(2-3):87–129, 1996. Special issue on Adaptive Hypertext and Hypermedia.
- [Bruza & Woods 08]
 P. D. Bruza and J. Woods. Quantum collapse in semantic space: Interpreting natural language argumentation. In *Proceedings of the Second Quantum Interaction Symposium (QI-2008)*. College Publications, 2008.
- [Bruza *et al.* 06]
 P. Bruza, D. Widdows, and J. Woods. A quantum logic of down below. In K. Engesser, D. Gabbay, and D. Lehmann, editors, *Handbook of Quantum Logic, Quantum Structure and Quantum Computation*, volume 2. Stanford: CSLI Publications, 2006.

- [Bry & Schwertel 06]
F. Bry and U. Schwertel, editors. *REWERSE Annual Meeting 2006*, March 2006.
- [Buchanan *et al.* 95]
B. G. Buchanan, J. D. Moore, D. E. Forsythe, G. Cerenini, S. Ohlsson, and G. Banks. An intelligent interactive system for delivering individualized information to patients. *Artificial Intelligence in Medicine*, 7:117 — 154, 1995.
- [Buckley & Salton 09]
C. Buckley and G. Salton. Term-weighting approaches in automatic text retrieval. *Information Processing and Management*, 24(5):513–523, 2009.
- [Buckley 98]
C. Buckley. TIPSTER Advanced Query (DN2). TIPSTER programme working paper, 1998.
- [Buitelaar & Cimiano 08a]
P. Buitelaar and P. Cimiano, editors. *Bridging the Gap between Text and Knowledge - Selected Contributions to Ontology Learning and Population from Text*. IOS Press, 2008.
- [Buitelaar & Cimiano 08b]
P. Buitelaar and P. Cimiano. *Ontology learning and population: bridging the gap between text and knowledge*, volume 167. Ios Press, 2008.
- [Buitelaar & Declerck 03]
P. Buitelaar and T. Declerck. Linguistic annotation for the semantic web. In S. Handschuh and S. Staab, editors, *Annotation for the Semantic Web*. IOS Press, 2003.
- [Buitelaar *et al.* 03]
P. Buitelaar, T. Declerck, N. Calzolari, and A. Lenci. Towards a language infrastructure for the semantic web. In H. Cunningham, Y. Ding, and A. Kiryakov, editors, *Proceedings of Workshop on Human Language Technology for the Semantic Web and Web Services, 2nd International Semantic Web Conference*, Sanibel Island, Florida, Oct 2003.
- [Buitelaar *et al.* 04]
P. Buitelaar, D. Olejnik, and M. Sintek. A Protege Plug-in for Ontology Extraction from Text Based on Linguistic Analysis. In *Proceedings of the 1st European Semantic Web Symposium*, 2004.
- [Buitelaar *et al.* 05]
P. Buitelaar, P. Cimiano, and B. Magnini. *Ontology learning from text: Methods, applications and evaluation*. IOS Press, 2005.
- [Buitelaar *et al.* 09]
P. Buitelaar, P. Cimiano, P. Haase, and M. Sintek. Towards Linguistically Grounded Ontologies. In *Proceedings of the European Semantic Web Conference (ESWC'09), LNCS 5554*, pages 111–125, 2009.
- [Bundy 83]
A. Bundy. *The Computer Modelling of Mathematical Reasoning*. Academic Press, London, 1983.
- [Bundy *et al.* 85]
A. Bundy, L. Byrd, and C. Mellish. Special-Purpose, But Domain-Independent, Inference Mechanisms. In L. Steels and J. Campbell, editors, *Progress in Artificial Intelligence*. Ellis Horwood, Chichester, 1985.
- [Buneman *et al.* 98]
P. Buneman, A. Deutsch, W. Fan, H. Liefke, A. Sahuguet, and W. Tan. Beyond XML Query Languages. In *In Proceedings of the Query Language Workshop (QL'98)*, 1998.

- [Bunescu & Pasca 06]
R. C. Bunescu and M. Pasca. Using encyclopedic knowledge for named entity disambiguation. In *EACL*. The Association for Computer Linguistics, 2006.
- [Burger *et al.* 11]
J. Burger, J. Henderson, G. Kim, and G. Zarrella. Discriminating Gender on Twitter. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP '11*, pages 1301–1309, 2011.
- [Burges 98]
C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.
- [Burgess & Lund 97]
C. Burgess and K. Lund. Modelling parsing constraints with high dimensional context space. *Language and Cognitive Processes*, 12:1–34, 1997.
- [Burman *et al.* 11]
A. Burman, A. Jayapal, S. Kannan, M. Kavilikatta, A. Alhelbawy, L. Derczynski, and R. Gaizauskas. USFD at KBP 2011: Entity linking, slot filling and temporal bounding. In *Proceedings of the Text Analysis Conference (TAC'11)*, 2011.
- [Burnage & Dunlop 92]
G. Burnage and D. Dunlop. Encoding the British National Corpus. In *Proceedings of the 13th International Conference on English Language Research on Computerised Corpora*, 1992.
- [Burnage 90a]
G. Burnage. CELEX, A Guide for Users. Technical report, CELEX Centre for Lexical Information, Nijmegen, The Netherlands, 1990.
- [Burnage 90b]
G. Burnage. *CELEX: A Guide for users*. Centre for Lexical Information, Nijmegen, 1990.
- [Burnard 95]
L. Burnard. Users Reference Guide for the British National Corpus. <http://info.ox.ac.uk/bnc/>, May 1995.
- [Burnett *et al.* 87]
M. Burnett, M. Baker, C. Bohus, P. Carlson, S. Yang, and van Zee P. Scaling Up Visual Languages. *IEEE Computer*, 28(3):45–54, 1987.
- [Burton 76]
R. Burton. Semantic Grammar. An Engineering Technique for Constructing Natural Language Understanding Systems. Technical Report 3353, BBN, Cambridge Massachusetts, 1976.
- [Busemann & Horacek 98]
S. Busemann and H. Horacek. A Flexible Shallow Approach to Text Generation. In *Proceedings of the 9th International Workshop on Natural Language Generation*, 1998.
- [Busemann 99]
S. Busemann. Constraint-Based Techniques for Interfacing Software Modules. In *Proceedings of the AISB'99 Workshop on Reference Architectures and Data Standards for NLP*, Edinburgh, April 1999. The Society for the Study of Artificial Intelligence and Simulation of Behaviour.
- [Cadilhac *et al.* 10]
A. Cadilhac, F. Benamara, and N. Aussenac-Gilles. Ontolexical resources for feature based opinion mining: a case-study. In *23rd International Conference on Computational Linguistics*, page 77, 2010.

- [Caesius a]
Caesius. WebQL User's Guide: Introduction to WebQL. <http://www.webql.com>.
- [Caesius b]
Caesius. WebQL User's Guide: Introduction to WebQL. <http://www.webql.com>.
- [Cahill & Reape 98]
L. Cahill and M. Reape. Component Tasks in Applied NLG Systems. RAGS deliverable, <http://www.tri.brighton.ac.uk/projects/rags/>, 1998.
- [Cahill & Reape 99]
L. Cahill and M. Reape. Component tasks in applied NLG systems. Technical Report ITRI-99-05, ITRI, University of Brighton, 1999. Obtainable at <http://www.itri.brighton.ac.uk/projects/rags/>.
- [Cahill *et al.* 92a]
L. J. Cahill, R. Gaizauskas, and R. Evans. POETIC: A Fully-Implemented NL System for Understanding Traffic Reports. In *Fully-Implemented Natural Language Understanding Systems: Proceedings of the Trento Workshop*, number IWBS Report No. 236, pages 86–99. IBM Institute for Knowledge Based Systems, Heidelberg, March 1992.
- [Cahill *et al.* 92b]
L. Cahill, R. Gaizauskas, and R. Evans. POETIC: a fully-implemented NL system for understanding traffic reports. In *Fully Implemented Natural Language Understanding Systems: Proceedings of the Trento Workshop, March 30, 1992 (IWBS Report No. 236)*, pages 86–99, IBM Institute for Knowledge Based Systems, Heidelberg, 1992.
- [Cahill *et al.* 99a]
L. Cahill, C. Doran, R. Evans, C. Mellish, D. Paiva, M. Reape, D. Scott, and N. Tipper. Towards a Reference Architecture for Natural Language Generation Systems. Technical Report ITRI-99-14; HCRC/TR-102, University of Edinburgh and Information Technology Research Institute, Edinburgh and Brighton, 1999.
- [Cahill *et al.* 99b]
L. Cahill, C. Doran, R. Evans, D. Paiva, D. Scott, C. Mellish, and M. Reape. Achieving Theory-Neutrality in Reference Architectures for NLP: To What Extent is it Possible/Desirable? In *Proceedings of the AISB'99 Workshop on Reference Architectures and Data Standards for NLP*, Edinburgh, April 1999. The Society for the Study of Artificial Intelligence and Simulation of Behaviour.
- [Cahill *et al.* 01]
L. Cahill, J. Carroll, R. Evans, D. Paiva, R. Power, D. Scott, and K. van Deemter. From RAGS to RICHES: exploiting the potential of a flexible generation architecture. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL'01)*, pages 98–105, Toulouse, France, 2001. The Society for the Study of Artificial Intelligence and Simulation of Behaviour.
- [Cahoon & McKinley 96]
B. Cahoon and K. McKinley. Performance Evaluation of a Distributed Architecture for Information Retrieval. In *Proceedings of SIGIR '96*, pages 110–118, Zurich, 1996.
- [Calder 96]
J. Calder. Statistical techniques. In Sapsford and Jupp [Sapsford & Jupp 96], chapter 9.
- [Califf & Mooney 97]
M. Califf and R. Mooney. Relational learning of pattern-match rules for information extraction. *Working Papers of the ACL-97 Workshop in Natural Language Learning*, pages 9–15, 1997.
- [Califf 98]
M. E. Califf. *Relational Learning Techniques for Natural Language Information Extraction*. Unpublished PhD thesis, University of Texas at Austin, 1998.

- [Callaway 02]
C. B. Callaway. Tools for large-scale generation. In *Proceedings of the ACL-02 Demonstrations Session*, pages 118–119, Philadelphia, July 2002.
- [Callison-Burch & Dredze 10]
C. Callison-Burch and M. Dredze. Creating speech and language data with Amazon’s Mechanical Turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, pages 1–12, 2010.
- [Callison-Burch 09]
C. Callison-Burch. Fast, cheap, and creative: Evaluating translation quality using Amazon’s Mechanical Turk. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 286–295, 2009.
- [Calzolari *et al.* 01]
N. Calzolari, A. Lenci, and A. Zampolli. International standards for multilingual resource sharing: The isle computational lexicon working group. In *Proceedings of the ACL 2001 Workshop on Sharing Tools and Resources*, pages 39–46, 2001.
- [Calzolari *et al.* 02]
N. Calzolari, A. Zampolli, and A. Lenci. Towards a standard for a multilingual lexical entry: The eagles/isle initiative. In *CICLing 2002*, pages 264–279, 2002.
- [Campbell *et al.* 98]
K. Campbell, D. E. Oliver, K. A. Spackmann, and E. H. Shortliffe. Representing thoughts, words and things in the UMLS. *Journal of the American Medical Informatics Association*, 5(5):421–431, 1998.
- [Campbell *et al.* 00]
C. Campbell, N. Cristianini, and A. Smola. Query Learning with Large Margin Classifiers. In *Proceedings of the Seventeenth International Conference on Machine Learning (ICML-00)*, 2000.
- [Campione *et al.* 98]
M. Campione, K. Walrath, A. Huml, and the Tutuorial Team. *The Java Tutorial Continued: The Rest of the JDK*. Addison-Wesley, Reading, MA, 1998.
- [Cancedda *et al.* 03]
N. Cancedda, N. Cesa-Bianchi, A. Conconi, C. Gentile, C. Goutte, T. Graepel, Y. Li, J. Renders, and J. Shawe-Taylor. Kernel methods for document filtering. In E. M. Voorhees and L. P. Buckland, editors, *Proceedings of The Eleventh Text Retrieval Conference (TREC 2002)*. The NIST, 2003.
- [Cann 93]
R. Cann. *Formal Semantics: An Introduction*. Cambridge University Press, 1993.
- [Cano Basave *et al.* 13]
A. Cano Basave, A. Varga, M. Rowe, M. Stankovic, and A.-S. Dadzie. Making sense of microposts (#msm2013) concept extraction challenge. *CEUR Workshop Proceedings*, 1019:1–15, 2013.
- [Caporaso *et al.* 07]
J. G. Caporaso, W. A. B. Jr., D. A. Randolph, K. B. Cohen, , and L. Hunter. MutationFinder: A high-performance system for extracting point mutation mentions from text. *Bioinformatics*, 23(14):1862–1865, 2007.
- [Caracciolo *et al.* 09]
C. Caracciolo, J. Heguiabehere, V. Presutti, and A. Gangemi. Initial network of fisheries ontologies. Technical Report D7.2.3, NeOn Project Deliverable, 2009.

- [Carbonell & Goldstein 98a]
 J. G. Carbonell and J. Goldstein. The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *Research and Development in Information Retrieval*, pages 335–336, 1998.
- [Carbonell & Goldstein 98b]
 J. G. Carbonell and J. Goldstein. The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *SIGIR*, pages 335–336, 1998.
- [Carbonell *et al.* 83]
 J. Carbonell, R. Michalski, and T. Mitchell. An Overview of Machine Learning. In J. Carbonell, R. Michalski, and T. Mitchell, editors, *Machine Learning: An Artificial Intelligence Approach*, pages 3–23. Tioga Pub. Co., Palo Alto, CA, 1983.
- [Card *et al.* 15]
 D. Card, A. E. Boydston, J. H. Gross, P. Resnik, and N. A. Smith. The media frames corpus: Annotations of frames across issues. In *ACL (2)*, pages 438–444. The Association for Computer Linguistics, 2015.
- [Cardie & Wagstaff 99]
 C. Cardie and K. Wagstaff. Noun phrase coreference as clustering. In *Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Very Large Corpora*, pages 82–89. ACL, 1999.
- [Cardie 97]
 C. Cardie. Empirical Methods in Information Extraction. *AI Magazine*, 18(4), 1997.
- [Carenini & Cheung 08]
 G. Carenini and J. C. K. Cheung. Extractive vs. NLG-based abstractive summarization of evaluative text: the effect of corpus controversiality. In *Proceedings of the Fifth International Natural Language Generation Conference, INLG '08*, pages 33–41, 2008.
- [Carenini & Moore 94]
 G. Carenini and J. Moore. Using the UMLS semantic network as a basis for constructing a terminological knowledge base: a preliminary report. In *Proc. of SCAMC '94*, pages 725–729, 1994.
- [Carenini *et al.* 05]
 G. Carenini, R. Ng, and E. Zwart. Extracting knowledge from evaluative text. In *Proceedings of the 3rd international conference on Knowledge capture*, pages 11–18. ACM, 2005.
- [Carletta 96]
 J. Carletta. Assessing agreement on classification tasks: the Kappa statistic. *Computational Linguistics*, 22(2):249–254, 1996.
- [Carletta *et al.* 02]
 J. Carletta, D. McKelvie, A. Isard, A. Mengel, M. Klein, and M. B. Møller. A generic approach to software support for linguistic annotation using XML. In G. Sampson and D. McCarthy, editors, *Readings in Corpus Linguistics*. London and NY: Continuum International, 2002.
- [Carletta *et al.* 03]
 J. Carletta, S. Evert, U. Heid, J. Kilgour, J. Robertson, and H. Voormann. The NITE XML Toolkit: flexible annotation for multi-modal language data. *Behavior Research Methods, Instruments, and Computers*, 35(3), 2003. Special issue on Measuring Behavior.
- [Carletta *et al.* 05]
 J. Carletta, S. Evert, U. Heid, and J. Kilgour. The nite xml toolkit: Data model and query language. *Language Resources and Evaluation*, 39(4):313–334, 2005.

- [Carlos *et al.* 11]
C. Carlos, M. Marcelo, and P. Barbara. Information credibility on twitter. In *Proceedings of the 20th International Conference on World Wide Web, WWW '11*, pages 675–684, 2011.
- [Carlos *et al.* 13]
C. Carlos, M. Marcelo, and P. Barbara. Predicting information credibility in time-sensitive social media. *Internet Research*, 23(5):560–588, 2013.
- [Carlson & Nirenburg 90]
L. Carlson and S. Nirenburg. World modeling for NLP. Technical Report CMU-CMT-90-121, Carnegie Mellon University, Center for Machine Translation, 1990.
- [Carlson & Nirenburg 92]
L. Carlson and S. Nirenburg. Practical world modeling for NLP applications. In *Proceedings of the Third Conference on Applied Natural Language Processing*, pages 235–236. Association for Computational Linguistics, 1992.
- [Carmona *et al.* 98]
J. Carmona, S. Cervell, L. Màrquez, M. A. Martí, L. Padró, R. Placer, H. Rodríguez, M. Taulé, and J. Turmo. An environment for morphosyntactic processing of unrestricted spanish text. In *Proc. of LREC-98*, Granada, Spain, 1998.
- [Carpenter & Penn 92]
B. Carpenter and G. Penn. ALE 2.0 User’s Guide. Technical report, Carnegie Mellon University Laboratory for Computational Linguistics, Pittsburgh, 1992.
- [Carreras & Padró 02]
X. Carreras and L. Padró. A Flexible Distributed Architecture for Natural Language Analyzers. In *Proceedings of the LREC 2002 Third International Conference on Language Resources and Evaluation*, pages 1813–1817, Las Palmas, Canary Islands - Spain, 2002.
- [Carreras *et al.* 03]
X. Carreras, L. Màrquez, and L. Padró. Learning a perceptron-based named entity chunker via online recognition feedback. In *Proceedings of CoNLL-2003*, pages 156–159. Edmonton, Canada, 2003.
- [Carter 87]
D. Carter. *Interpreting Anaphors in natural language texts*. Chichester : Ellis Horwood, 1987.
- [Carter *et al.* 13a]
S. Carter, W. Weerkamp, and E. Tsagkias. Microblog language identification: Overcoming the limitations of short, unedited and idiomatic text. *Language Resources and Evaluation Journal*, 2013.
- [Carter *et al.* 13b]
S. Carter, W. Weerkamp, and E. Tsagkias. Microblog language identification: Overcoming the limitations of short, unedited and idiomatic text. *Language Resources and Evaluation*, 2013.
- [Carthy & Smeaton 00]
J. Carthy and A. Smeaton. The design of a topic tracking system. In *Proceedings of the 22nd Annual Colloquium on Information Retrieval Research*, Cambridge, UK, 2000.
- [Cassidy & Bird 00]
S. Cassidy and S. Bird. Querying databases of annotated speech. In *Eleventh Australasian Database Conference*, Australian National University, Canberra, 2000.
- [Cassidy & Harrington 01]
S. Cassidy and J. Harrington. Multi-level annotation in the Emu speech database management system. *Speech Communication*, 33:61–77, 2001.

- [Cassidy 02]
S. Cassidy. Xquery as an annotation query language: a use case analysis. In *Proceedings of 3rd Language Resources and Evaluation Conference (LREC'2002)*, Gran Canaria, Spain, 2002.
- [Caterpillar Corporation 74]
Caterpillar Corporation. *Dictionary for Caterpillar Fundamental English*. Caterpillar Corporation, 1974.
- [Catizone *et al.*]
R. Catizone, R. Collier, and M. Stevenson. Verification of Consistency of Resources. Deliverable for ECRANecran project [?].
- [Cattuto *et al.* 08]
C. Cattuto, D. Benz, A. Hotho, and G. Stumme. Semantic grounding of tag relatedness in social bookmarking systems. In *Proceedings of the 7th International Conference on The Semantic Web*, pages 615–631, 2008.
- [Cavazza & Zweigenbaum 92]
M. Cavazza and P. Zweigenbaum. Extracting Implicit Information from Free Text Technical Reports. *Information Processing and Management*, 28(5), 1992.
- [Cavazza & Zweigenbaum 94]
M. Cavazza and P. Zweigenbaum. Lexical Semantics: Dictionary or Encyclopaedia? In P. Saint-Dizier and E. Viegas, editors, *Computational Lexical Semantics*. Cambridge University Press, Cambridge, 1994.
- [Cavnar & Trenkle 94]
W. Cavnar and J. Trenkle. N-gram-based text categorization. In *Proceedings of the Annual Symposium on Document Analysis and Information Retrieval*, pages 161–175, 1994.
- [CEC 96]
CEC. *Telematics Applications Programme, Language Engineering Project Directory*. European Commission Directorate General XIII, Luxembourg, 1996.
- [Ceccaroni & Kendall 03]
L. Ceccaroni and E. Kendall. A graphical environment for ontology development. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, 2003.
- [Cederberg & Widdows 03]
S. Cederberg and D. Widdows. Using LSA and noun coordination information to improve the precision and recall of automatic hyponymy extraction. In *Proceedings of the 7th conference on Natural language learning at HLT-NAACL*, pages 111–118, Morristown, NJ, 2003.
- [Celino *et al.* 11]
I. Celino, D. Dell’Aglia, E. Della Valle, Y. Huang, T. Lee, S. Park, and V. Tresp. Making Sense of Location-based Micro-posts Using Stream Reasoning. In *Proceedings of the Making Sense of Microposts Workshop (#MSM2011), collocated with the 8th Extended Semantic Web Conference, Heraklion, Crete, Greece, 2011*.
- [Cen95]
Centre for Computational Linguistics, UMIST, Manchester, England. *User manual for the CCL Conceptual Graph Engine*, preliminary version 0.1 edition, 1995.
- [Center for Digital Democracy 01]
Center for Digital Democracy. TV That Watches You: The Prying Eyes of Interactive Television. <http://www.democraticmedia.org/privacyreport.pdf>, 2001.

- [Cerami 02]
E. Cerami. *Web Services Essentials*. O’Reilly, 2002.
- [Cerbah 08]
F. Cerbah. Mining the content of relational databases to learn ontologies with deeper taxonomies. In *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pages 553–557, 2008.
- [Cesa-Bianchi *et al.* 04]
N. Cesa-Bianchi, C. Gentile, A. Tironi, and L. Zaniboni. Incremental Algorithms for Hierarchical Classification. In *Neural Information Processing Systems*, 2004.
- [Chadwick & Sasse 06]
D. Chadwick and A. Sasse. The virtuous circle of expressing authorization policies. In *Semantic Web Policy Workshop*, Athens, Georgia, USA, 2006.
- [Chai & Biermann 97]
J. Chai and A. Biermann. The use of WordNet in automatic information extraction. In *Proc. of ACL Workshop on Word Sense Disambiguation*, Washington D.C., 1997.
- [Chai & Biermann 99]
J. Y. Chai and A. W. Biermann. The use of word sense disambiguation in an information extraction system. In *AAAI/IAAI*, pages 850–855, 1999.
- [Chai *et al.* 99]
J. Y. Chai, A. W. Biermann, and C. I. Guinn. Two dimensional generalization in information extraction. In *AAAI/IAAI*, pages 431–438, 1999.
- [Chaisorn *et al.* 03]
T. Chaisorn, L. and Chua, C. Koh, Y. Zhao, H. Xu, H. Feng, and Q. Tian. A two-level multi-modal approach for story segmentation of large news video corpus. In *Proceedings of the TRECVID Conference*, Gaithersburg, Washington D.C, 2003.
- [Chakrabarti & Punera 11]
D. Chakrabarti and K. Punera. Event Summarization Using Tweets. In *Proceedings of the Fifth International Conference on Weblogs and Social Media (ICWSM)*, 2011.
- [Chakrabarti 03]
S. Chakrabarti. *Mining the Web – Discovering Knowledge from Hypertext Data*. Morgan Kaufmann Publishers, 2003.
- [Chamberlain *et al.* 09a]
J. Chamberlain, M. Poesio, and U. Kruschwitz. A new life for a dead parrot: Incentive structures in the phrase detectives game. In *Proceedings of WWW*, 2009.
- [Chamberlain *et al.* 09b]
J. Chamberlain, U. Kruschwitz, and M. Poesio. Constructing an anaphorically annotated corpus with non-experts: assessing the quality of collaborative annotations. In *Proceedings of the 2009 Workshop on The People’s Web Meets NLP: Collaboratively Constructed Semantic Resources*, pages 57–62, 2009.
- [Chamberlain *et al.* 12]
J. Chamberlain, U. Kruschwitz, and M. Poesio. Motivations for participation in socially networked collective intelligence systems. *CoRR*, abs/1204.4071, 2012.
- [Champin 01]
P.-A. Champin. Rdf tutorial. <http://www710.univ-lyon1.fr/champin/rdf-tutorial>, April 2001.

- [Chan *et al.* 10]
 J. Chan, C. Hayes, and E. Daly. Decomposing discussion forums using common user roles. In *Proceedings of WebSci10: Extending the Frontiers of Society On-Line*, 2010.
- [Chandler & Sweller 92]
 P. Chandler and J. Sweller. The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62(2):233–246, 1992.
- [Chandler 97]
 D. Chandler. An Introduction to Genre Theory. <http://www.aber.ac.uk/media/Documents/intgenre/intgenre.htm> 1997.
- [Chandrasekar & Bangalore 98]
 R. Chandrasekar and S. Bangalore. Knowing a word by the company it keeps using local information in a maximum entropy model for word sense disambiguation. Submitted for publication, 1998.
- [Chang & Lee 73]
 C.-L. Chang and R. C.-T. Lee. *Symbolic Logic and Mechanical Theorem Proving*. Academic Press, London, 1973.
- [Chang & Lin 11a]
 C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 2011.
- [Chang & Lin 11b]
 C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines. *ACM TIST*, 2(3), 2011.
- [Chang & Manning 12]
 A. X. Chang and C. D. Manning. Sutime: A library for recognizing and normalizing time expressions. In *LREC*, pages 3735–3740, 2012.
- [Chang *et al.* 06]
 C. Chang, M. Kayed, M. Girgis, and K. Shaalan. A survey of web information extraction systems. *IEEE transactions on knowledge and data engineering*, pages 1411–1428, 2006.
- [Chapelle *et al.* 00]
 O. Chapelle, J. Weston, L. Bottou, and V. Vapnik. Vicinal risk minimization. In *NIPS*, pages 416–422, 2000.
- [Charlett *et al.* 95]
 D. Charlett, R. Garland, and N. Marr. How Damaging is Negative Word of Mouth? *Marketing Bulletin*, 6:42–50, 1995.
- [Charniak & Goldman 88]
 E. Charniak and R. Goldman. A Logic for Semantic Interpretation. In *Proceedings of the 26th Conference of the Association for Computational Linguistics*, pages 87–94, Buffalo, June 7-10 1988.
- [Charniak & McDermott 85]
 E. Charniak and D. McDermott. *Introduction to Artificial Intelligence*. Addison-Wesley, Reading MA, 1985.
- [Charniak 93]
 E. Charniak. *Statistical Language Learning*. MIT Press, Cambridge MA, 1993.
- [Charniak 96]
 E. Charniak. Tree-bank Grammars. Technical Report CS-96-02, Department of Computer Science, Brown University, 1996.

- [Charniak 00]
E. Charniak. A maximum-entropy-inspired parser. In *Proceedings of the 1st North American chapter of the Association for Computational Linguistics conference*, pages 132–139. Morgan Kaufmann Publishers Inc., 2000.
- [Chebotko *et al.* 04]
A. Chebotko, Y. Deng, S. Lu, F. Fotouhi, A. Aristar, H. Brugman, A. Klassmann, H. Sloetjes, A. Russel, and P. Wittenburg. Ontoelan: an ontology-based linguistic multimedia annotator. In *Proceedings of IEEE Sixth International Symposium on Multimedia Software Engineering*, pages 329–336, December 2004.
- [Cheeseman & Stutz 96]
P. Cheeseman and J. Stutz. Bayesian Classification (AutoClass): Theory and results. In U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, editors, *Advances in Knowledge Discovery and Data Mining*, pages 153–180. AAAI Press/MIT Press, Menlo Park CA, 1996.
- [Chen & etal. 97]
H. Chen and et al. Description of the NTU system used for MET2. In *Proceedings of the 7th Message Understanding Conference*, 1997.
- [Chen & Martin 07]
Y. Chen and J. Martin. Cu-comsem: Exploring rich features for unsupervised web personal named disambiguation. In *Proceedings of SemEval 2007, Association for Computational Linguistics*, pages 125–128, 2007.
- [Chen & Sharp 04]
H. Chen and B. Sharp. Content-rich biological network constructed by mining pubmed abstracts. *BMC Bioinformatics*, 5(1):147, 2004.
- [Chen & Withgott 92]
F. R. Chen and M. Withgott. The Use of Emphasis to Automatically Summarize a Spoken Discourse. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 1, pages 229–232, San Francisco, 1992.
- [Chen & Witte 07]
G. Chen and J. Witte. Enriching the analysis of genomewide association studies with hierarchical modeling. *The American Journal of Human Genetics*, 81(2), August 2007.
- [Chen 03]
A. Chen. Chinese Word Segmentation Using Minimal Linguistic Knowledge. In *Proceedings of the 2nd SIGHAN Workshop on Chinese Language Processing*, 2003.
- [Chen *et al.* 10a]
D. Chen, N. Schneider, D. Das, and N. A. Smith. Semafor: Frame argument resolution with log-linear models. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pages 264–267. Association for Computational Linguistics, 2010.
- [Chen *et al.* 10b]
J. Chen, R. Nairn, L. Nelson, M. Bernstein, and E. Chi. Short and tweet: Experiments on recommending content from information streams. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems, CHI '10*, pages 1185–1194, 2010.
- [Chen *et al.* 11]
J. Chen, R. Nairn, and E. Chi. Speak little and well: Recommending conversations in online social streams. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems, CHI '11*, pages 217–226, 2011.

- [Chen *et al.* 12a]
 K. Chen, T. Chen, G. Zheng, O. Jin, E. Yao, and Y. Yu. Collaborative personalized tweet recommendation. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 661–670, New York, NY, USA, 2012. ACM.
- [Chen *et al.* 12b]
 L. Chen, W. Wang, M. Nagarajan, S. Wang, and A. Sheth. Extracting diverse sentiment expressions with target-dependent polarity from twitter. 2012.
- [Cheng & Evans 09]
 A. Cheng and M. Evans. An In-Depth Look Inside the Twitter World. Technical report, Sysomos Inc., 2009. <http://www.sysomos.com/insidetwitter/>.
- [Cheng & Xu 08]
 X. Cheng and F. Xu. Fine-grained opinion topic and polarity identification. In *Proceedings of LREC*, volume 8, 2008.
- [Cheng *et al.* 08a]
 G. Cheng, W. Ge, and Y. Qu. Falcons: Searching and Browsing Entities on the Semantic Web. In *Proceedings of WWW2008*, pages 1101–1102, 2008.
- [Cheng *et al.* 08b]
 Y. Cheng, G. Qiu, J. Bu, K. Liu, Y. Han, C. Wang, and C. Chen. Model bloggers' interests based on forgetting mechanism. In *Proceedings of the 17th International Conference on World Wide Web*, WWW '08, pages 1129–1130, 2008.
- [Cheng *et al.* 10]
 Z. Cheng, J. Caverlee, and K. Lee. You are where you tweet: A content-based approach to geolocating twitter users. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, CIKM '10, pages 759–768, New York, NY, USA, 2010. ACM.
- [Cheong & Lee 10]
 M. Cheong and V. Lee. Twittering for earth: A study on the impact of microblogging activism on earth hour 2009 in australia. In *Intelligent Information and Database Systems*, pages 114–123. Springer, 2010.
- [Cheong *et al.* 94]
 T. Cheong, A. Kwang, A. Gunawan, G. Loo, L. Qwun, and S. Leng. A Pragmatic Information Extraction Architecture for the Message Formatting Export (MFE) System. In *Proceedings of the 2nd Singapore Conference on Intelligent Systems (SPICIS '94)*, pages B371–B377, Singapore, 1994.
- [Chernova *et al.* 10]
 S. Chernova, J. Orkin, and C. Breazeal. Crowdsourcing hri through online multiplayer games. In *Dialog with Robots: Papers from the AAAI Fall Symposium (FS-10-05)*, 2010.
- [Cherry & Guo 15]
 C. Cherry and H. Guo. The Unreasonable Effectiveness of Word Representations for Twitter Named Entity Recognition. In R. Mihalcea, J. Chai, and A. Sarkar, editors, *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 735–745, Denver, Colorado, 2015. Association for Computational Linguistics.
- [Chiaramella & Nie 90]
 Y. Chiaramella and J. Nie. A retrieval model based on an extended modal logic and its application to the RIME experiment approach. In *Proceedings of ACM SIGIR*, pages 25–43, Brussels, 1990.

- [Chieu & Ng. 02a]
H. L. Chieu and H. T. Ng. A Maximum Entropy Approach to Information Extraction from Semi-Structured and Free Text. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence*, pages 786–791, 2002.
- [Chieu & Ng 02b]
H. L. Chieu and H. T. Ng. Named entity recognition: A maximum entropy approach using global information. In *Proceedings of the 19th International Conference on Computational Linguistics (COLING'02)*, Taipei, Taiwan, 2002.
- [Chieu & Ng 02c]
H. L. Chieu and H. T. Ng. Teaching a weaker classifier: Named entity recognition on upper case text. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL'02)*, Philadelphia, PA, 2002.
- [Chieu & Ng 03]
H. L. Chieu and H. T. Ng. Named entity recognition with a maximum entropy approach. In W. Daelemans and M. Osborne, editors, *Proceedings of CoNLL-2003*, pages 160–163. Edmonton, Canada, 2003.
- [Chin 84]
D. Chin. An Analysis of Scripts Generated in Writing Between Users and Computer Consultants. *National Computer Conference*, pages 637–642, 1984.
- [Chinchor 92]
N. Chinchor. MUC-4 Evaluation Metrics. In *Proceedings of the Fourth Message Understanding Conference*, pages 22–29, 1992.
- [Chinchor 95]
N. A. Chinchor. Named entity task definition. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, pages 317–332, Columbia, MD, November 1995. version 2.1.
- [Chinchor 98a]
N. A. Chinchor. Overview of MUC-7/MET-2. In *Proceedings of the 7th Message Understanding Conference (MUC7)*, Apr 1998. Available at http://www.muc.saic.com/proceedings/muc_7_toc.html.
- [Chinchor 98b]
N. A. Chinchor. Overview of Proceedings of the Seventh Message Understanding Conference (MUC-7)/MET-2. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)*, page 5 pages, Fairfax, VA, April 1998. http://www.itl.nist.gov/iaui/894.02/related_projects/muc/.
- [Chinchor 98c]
N. A. Chinchor. Proceedings of the Seventh Message Understanding Conference (MUC-7) named entity task definition. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)*, page 21 pages, Fairfax, VA, April 1998. version 3.5, http://www.itl.nist.gov/iaui/894.02/related_projects/muc/.
- [Chinchor *et al.* 93]
N. Chinchor, L. Hirschman, and D. Lewis. Evaluating Message Understanding Systems: An Analysis of the Third Message Understanding Conference (MUC-3). *Computational Linguistics*, 19(3):409–449, 1993.
- [Chinchor *et al.* 99a]
N. Chinchor, E. Brown, L. Ferro, and P. Robinson. Named entity recognition task definition. Technical Report Version 1.4, The MITRE Corporation and SAIC, 1999.
- [Chinchor *et al.* 99b]
N. A. Chinchor, P. Robinson, and E. Brown. HUB-4 named entity task definition. In *Proceedings of the DARPA Broadcast News Workshop*, page 38 pages, Herndon, VA, February 1999.

- [Chklovski & Mihalcea 02]
T. Chklovski and R. Mihalcea. Building a sense tagged corpus with open mind word expert. In *Proceedings of the ACL-02 workshop on Word sense disambiguation: recent successes and future directions - Volume 8*, pages 116–122, 2002.
- [Chklovski 03]
T. Chklovski. Learner: a system for acquiring commonsense knowledge by analogy. In *Proceedings of the 2nd international conference on Knowledge capture*, pages 4–12, 2003.
- [Chklovski 05a]
T. Chklovski. Collecting paraphrase corpora from volunteer contributors. In *In Proceedings of the 3rd Int. Conference on Knowledge Capture*, pages 115–120, 2005.
- [Chklovski 05b]
T. Chklovski. Collecting paraphrase corpora from volunteer contributors. In *Proceedings of the 3rd international conference on Knowledge capture*, pages 115–120, 2005.
- [Chodorow *et al.* 85]
M. S. Chodorow, R. J. Byrd, and G. E. Heidorn. Extracting Semantic Hierarchies from a Large On-Line Dictionary. In *Proceedings of the 23rd Annual Meeting of the ACL*, pages 299–304, Chicago, IL, 1985.
- [Choi *et al.* 05]
Y. Choi, C. Cardie, E. Riloff, and S. Patwardhan. Identifying Sources of Opinions with Conditional Random Fields and Extraction Patterns. In *Proceedings of the 2005 Human Language Technology Conference / Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP-05)*, 2005.
- [Choi *et al.* 12a]
J. Choi, W. B. Croft, and J. Y. Kim. Quality models for microblog retrieval. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management, CIKM '12*, pages 1834–1838, New York, NY, USA, 2012. ACM.
- [Choi *et al.* 12b]
J. Choi, W. B. Croft, and J. Y. Kim. Quality models for microblog retrieval. In *CIKM*, 2012.
- [Chomsky]
N. Chomsky. On Binding. *Linguistic Inquiry*, 11(1):1–46.
- [Chomsky 99]
N. Chomsky. *Profit over People : Neoliberalism and Global Order*. Seven Stories Press, New York, 1st ed. edition, 1999.
- [Choudhury & Breslin 11]
S. Choudhury and J. Breslin. Extracting semantic entities and events from sports tweets. In *Proceedings, 1st Workshop on Making Sense of Microposts (#MSM2011): Big things come in small packages*, pages 22–32, 2011.
- [Choudhury *et al.* 13]
M. D. Choudhury, S. Counts, E. Horvitz, and M. Gamon. Predicting depression via social media. In *Proceedings of ICWSM*. AAAI, July 2013.
- [Chowdhury & Lynch 91]
G. Chowdhury and M. Lynch. Natural Language Processing of the Texts of Chemical Patent Abstracts. In A. Lichnerowicz, editor, *Intelligent Text and Image Handling*, pages 740–753, Amsterdam, 1991. Elsevier.

- [Christ 94]
O. Christ. A Modular and Flexible Architecture for an Integrated Corpus Query System. In *Proceedings of the 3rd Conference on Computational Lexicography and Text Research (COMPLEX '94)*, Budapest, 1994. <http://xxx.lanl.gov/abs/cs.CL/9408005>.
- [Christ 95]
O. Christ. Linking WordNet to a Corpus Query System. In *Proceedings of the Conference on Linguistic Databases*, Groningen, 1995.
- [Christensen *et al.* 01]
E. Christensen, F. Curbera, G. Meredith, and S. Weerawarana. Web Services Description Language (WSDL) 1.1. Technical report, W3C Note, <http://www.w3.org/TR/wsdl>, 2001.
- [Christine *et al.* 98]
K. Christine, E. Adolphson, T. Mitamura, and E. Nyberg. Controlled language for multilingual document production: Experience with caterpillar technical english, 1998.
- [Christoforaki *et al.* 95]
M. Christoforaki, P. Constantopoulos, and M. Doerr. Modelling occurrences in cultural documentation. In *III Convegno Internazionale di Archeologia e Informatica*, Roma, Italy, November 1995.
- [Chua. 02]
T. Chua. Learning Pattern Rules for Chinese Named Entity Extraction. In *Proceedings of AAAI '02*, 2002.
- [Chuang & Yang 00]
W. T. Chuang and J. Yang. Extracting sentence segments for text summarization: A machine learning approach. In *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '00, pages 152–159, New York, NY, USA, 2000. ACM.
- [Chubak & Rafiei 12]
P. Chubak and D. Rafiei. Efficient indexing and querying over syntactically annotated trees. *PVLDB*, 5(11):1316–1327, 2012.
- [Chung *et al.* 03]
W. Chung, H. Chen, and N. J. J.F. Business Intelligence Explorer: A Knowledge Map Framework for Discovering Business Intelligence on the Web. In *Hawaii International Conference on System Sciences*, Los Alamitos, CA, USA, 2003. IEEE Computer Society.
- [Church & Hanks 89]
K. W. Church and P. Hanks. Word association norms, mutual information and lexicography. In *Proc. of 27th Annual Meeting of the Association for Computational Linguistics, ACL'89*, pages 76–83, 1989.
- [Church & Hanks 90a]
K. W. Church and P. Hanks. Word Association Norms, Mutual Information, and Lexicography. *Computational Linguistics*, 16(1):22–29, 1990.
- [Church & Hanks 90b]
K. W. Church and P. Hanks. Word association norms, mutual information and lexicography. *Computational Linguistics*, 16(1):22–29, 1990.
- [Church & Mercer 93a]
K. Church and R. Mercer. Introduction. *Special issue on Computational Linguistics Using Large Corpora, Computational Linguistics*, 19(1), 1993.

- [Church & Mercer 93b]
K. W. Church and R. L. Mercer. Introduction to the special issue on computational linguistics using large corpora. *Computational Linguistics*, 19:1–24, 1993.
- [Church & Patil 82]
K. Church and R. Patil. Coping with Syntactic Ambiguity or How to Put the Block in the Box. *American Journal of Computational Linguistics*, 8(3-4), 1982.
- [Church & Rau 95]
K. Church and L. Rau. Commercial Applications of Natural Language Processing. *Communications of the ACM*, 38(11), November 1995.
- [Church 88]
K. Church. A stochastic parts of speech program and noun phrase parser for unrestricted texts. In *Second Conference on Applied NLP, ACL-88*, 1988.
- [Cieliebak *et al.* 13]
M. Cieliebak, O. Dürr, and F. Uzdilli. Potential and limitations of commercial sentiment detection tools. In *ESSEM@AI*IA*, pages 47–58, 2013.
- [Cimiano & Handschuh 03]
P. Cimiano and S. Handschuh. Ontology-Based Linguistic Annotation. In *Workshop On Linguistic Annotation: Getting The Model Right*, 2003.
- [Cimiano & Voelker 05]
P. Cimiano and J. Voelker. Text2Onto - A Framework for Ontology Learning and Data-driven Change Discovery. In *Proceedings of the 10th International Conference on Applications of Natural Language to Information Systems (NLDB)*, Alicante, Spain, 2005.
- [Cimiano 04a]
P. Cimiano. Orakel: A natural language interface to an f-logic knowledge base. In *9th International Conference on Applications of Natural Language to Information Systems (NLDB'04)*, 2004.
- [Cimiano 04b]
P. Cimiano. Orakel: A natural language interface to an f-logic knowledge base. In F. Mezziane and E. Métais, editors, *Natural Language Processing and Information Systems*, volume 3136 of *Lecture Notes in Computer Science*, pages 149–165. Springer Berlin / Heidelberg, 2004.
- [Cimiano *et al.* 03a]
P. Cimiano, F. Ciravegna, J. Domingue, S. Handschuh, A. Lavelli, S. Staab, and M. Stevenson. Requirements for Information Extraction for Knowledge Management. In *Proceedings of the Knowledge Management and Semantic Annotation Workshop*, 2003.
- [Cimiano *et al.* 03b]
P. Cimiano, S. Staab, and J. Tane. Automatic Acquisition of Taxonomies from Text: FCA meets NLP. In *Proceedings of the ECML/PKDD Workshop on Adaptive Text Extraction and Mining*, pages 10–17, Cavtat-Dubrovnik, Croatia, 2003.
- [Cimiano *et al.* 04a]
P. Cimiano, S. Handschuh, and S. Staab. Towards the Self-Annotating Web. In *Proc. of the 13th International Conference on World Wide Web (WWW'04)*, 2004.
- [Cimiano *et al.* 04b]
P. Cimiano, S. Handschuh, and S. Staab. Towards the Self-Annotating Web. In *Proceedings of WWW'04*, 2004.

- [Cimiano *et al.* 04c]
P. Cimiano, A. Hotho, and S. Staab. Comparing conceptual, divide and agglomerative clustering for learning taxonomies from text. In *Proceedings of the 16th European Conference on Artificial Intelligence, ECAI'2004, including Prestigious Applicants of Intelligent Systems, PAIS 2004*, 2004.
- [Cimiano *et al.* 05]
P. Cimiano, G. Ladwig, and S. Staab. Gimme' The Context: Context-driven automatic semantic annotation with C-PANKOW. In *Proceedings of the 14th World Wide Web Conference*, 2005.
- [Cimiano *et al.* 06]
P. Cimiano, M. Hartung, and E. Ratsch. Learning the appropriate generalization level for relations extracted from the Genia corpus. In *Proc. of the 5th Language Resources and Evaluation Conference (LREC)*, 2006.
- [Cimiano *et al.* 07]
P. Cimiano, P. Haase, and J. Heizmann. Porting Natural Language Interfaces Between Domains: an Experimental User Study with the ORAKEL System. In *IUI '07: Proceedings of the 12th international conference on Intelligent user interfaces*, pages 180–189, New York, NY, USA, 2007. ACM.
- [Cimino *et al.* 89]
J. Cimino, G. Hripcsak, S. Johnson, and P. Clayton. Designing an introspective, controlled medical vocabulary. In L. Kingsland, editor, *Proc. of 14th Annual Symposium on Computer Applications in Medical Care*, pages 513–518. IEEE Computer Society Press, 1989.
- [Cimino *et al.* 90]
J. Cimino, G. Hripcsak, S. Johnson, C. Friedman, D. Fink, and P. Clayton. Umls as a knowledge base - a rule-based expert system approach to controlled medical vocabulary management. In L. Kingsland, editor, *Proc. of 14th Annual Symposium on Computer Applications in Medical Care*, pages 175–179. IEEE Computer Society Press, 1990.
- [Ciravegna & Wilks 03a]
F. Ciravegna and Y. Wilks. Designing Adaptive Information Extraction for the Semantic Web in Amilcare. In S. Handschuh and S. Staab, editors, *Annotation for the Semantic Web*. IOS Press, Amsterdam, 2003.
- [Ciravegna & Wilks 03b]
F. Ciravegna and Y. Wilks. Designing Adaptive Information Extraction for the Semantic Web in Amilcare. In S. Handschuh and S. Staab, editors, *Annotation for the Semantic Web*. IOS Press, Amsterdam, 2003.
- [Ciravegna 00]
F. Ciravegna. Learning to Tag for Information Extraction from Text. In *ECAI Workshop on Machine Learning for Information Extraction*, Berlin, 2000.
- [Ciravegna 01a]
F. Ciravegna. Adaptive information extraction from text by rule induction and generalisation. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI 2001)*, 2001.
- [Ciravegna 01b]
F. Ciravegna. (LP)², an Adaptive Algorithm for Information Extraction from Web-related Texts. In *Proceedings of the IJCAI-2001 Workshop on Adaptive Text Extraction and Mining*, Seattle, 2001.
- [Ciravegna 03]
F. Ciravegna. (LP)², Rule Induction for Information Extraction Using Linguistic Constraints. Technical Report CS-03-07, Department of Computer Science, University of Sheffield, Sheffield, September 2003.

- [Ciravegna *et al.* 99]
 F. Ciravegna, A. Lavelli, N. Mana, J. Matiasek, L. Gilardoni, S. Mazza, M. Ferraro, W. Black, F. Rinaldi, and D. Mowatt. Facile: Classifying texts integrating pattern matching and information extraction. In *Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI99)*, 1999.
- [Ciravegna *et al.* 02a]
 F. Ciravegna, A. Dingli, D. Petrelli, and Y. Wilks. User-System Cooperation in Document Annotation Based on Information Extraction. In *13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02)*, pages 122–137, Siguenza, Spain, 2002.
- [Ciravegna *et al.* 02b]
 F. Ciravegna, A. Dingli, Y. Wilks, and D. Petrelli. Timely and non-intrusive active document annotation via adaptive information extraction. In *Proceedings of ECAI Workshop on Semantic Authoring, Annotation and Knowledge Markup*, Lyon, France, 2002.
- [Ciravegna *et al.* 03]
 F. Ciravegna, A. Dingli, J. Iria, and Y. Wilks. Multi-strategy definition of annotation services in Melita. In H. Cunningham, Y. Ding, and A. Kiryakov, editors, *Proceedings of Workshop on Human Language Technology for the Semantic Web and Web Services, 2nd International Semantic Web Conference*, Sanibel Island, Florida, Oct 2003.
- [Ciravegna *et al.* 04]
 F. Ciravegna, S. Chapman, A. Dingli, and Y. Wilks. Learning to Harvest Information for the Semantic Web. In *Proceedings of the 1st European Semantic Web Symposium*, Heraklion, Greece, May 2004.
- [Civit 00]
 M. Civit. Guía para la anotación morfosintáctica del corpus CLiC-TALP. Technical report, XTRACT Working Paper, Universitat de Barcelona, 2000.
- [Clark & Tarnlund 82]
 K. Clark and S.-A. Tarnlund. *Logic Programming*. Academic Press, New York, 1982.
- [Clark 73]
 E. Clark. What’s in a word? On the child’s acquisition of semantics in his first language. In T. Moore, editor, *Cognitive development and the acquisition of language*. Academic Press, New York, 1973.
- [Clark 77]
 K. Clark. Negation as Failure. In H. Gallaire and J. Minker, editors, *Logic and Databases*. Plenum Press, New York, 1977.
- [Clark *et al.* 08]
 S. Clark, B. Coecke, and M. Sadrzadeh. A compositional distributional model of meaning. In *Proceedings of the Second Quantum Interaction Symposium (QI-2008)*. College Publications, 2008.
- [Clarkson & Rosenfeld 97]
 P. Clarkson and R. Rosenfeld. Statistical Language Modeling using the SMU-Cambridge Toolkit. In *Proceedings of ESCA Eurospeech*, Greece, 1997.
- [Clematide *et al.* 12]
 S. Clematide, S. Gindl, M. Klenner, S. Petrakis, R. Remus, J. Ruppenhofer, U. Waltinger, and M. Wiegand. Mlsa – a multi-layered reference corpus for german sentiment analysis. In *Proceedings of LREC*, 2012.
- [Clements & Northrop 96]
 P. Clements and L. Northrop. Software Architecture: An Executive Overview. Technical Report CMU/SEI-96-TR-003, Software Engineering Institute, Carnegie Mellon University, 1996.

- [Clifton & Cooley 99]
C. Clifton and R. Cooley. TopCat: Data mining for topic identification in a text corpus. In J. M. Zytkow and J. Rauch, editors, *Proceedings of the 3rd European Conference on Principles of Data Mining and Knowledge Discovery (PKDD'99)*, volume 1704 of *Lecture Notes in Computer Science*, pages 174–183, Prague, Czech Republic, Sep 1999.
- [Clocksin & Mellish 84]
W. Clocksin and C. Mellish. *Programming in PROLOG*. Springer-Verlag, Berlin, 1984.
- [Clough *et al.* 02]
P. Clough, R. Gaizauskas, S. Piao, and Y. Wilks. METER: MEasuring TExt Reuse. In *Proceedings of the ACL*. Association for Computational Linguistics, July 2002.
- [Cobuild 99]
C. Cobuild, editor. *English Grammar*. Harper Collins, 1999.
- [Cochard 93]
J. Cochard. Application de techniques linguistiques en indexation et recherche documentaire. *La Tribune des Industries de la Langue*, (11–12), février 1993.
- [Cockburn 97]
A. Cockburn. Structuring Use Cases with Goals. *Journal of Object-Oriented Programming*, Sept-Oct and Nov-Dec, 1997.
- [Codd 70]
F. Codd. A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, 13(6):377–387, 1970.
- [Coden *et al.* 14]
A. Coden, D. Gruhl, N. Lewis, P. N. Mendes, M. Nagarajan, C. Ramakrishnan, and S. Welch. Semantic lexicon expansion for concept-based aspect-aware sentiment analysis. In *Semantic Web Evaluation Challenge*, pages 34–40. Springer, 2014.
- [Cogliano *et al.* 04]
V. J. Cogliano, R. A. Baan, K. Straif, Y. Grosse, M. B. Secretan, F. E. Ghissassi, and P. Kleihues. The science and practice of carcinogen identification and evaluation. *Environmental Health Perspectives*, 112(13):1269–1274, September 2004.
- [Cohen & Hersh 05]
A. M. Cohen and W. R. Hersh. A survey of current work in biomedical text mining. *Briefings in Bioinformatics*, 6(1):57–71, 2005.
- [Cohen & Hunter 08]
K. B. Cohen and L. Hunter. Getting started in text mining. *PLoS Comput Biol*, 4(1):e20, 01 2008.
- [Cohen 89]
M. Cohen. *Phonological Structures for Speech Recognition*. Unpublished PhD thesis, University of California at Berkeley, 1989.
- [Cohen *et al.* 90]
M. Cohen, H. Murveit, J. Bernstein, P. Price, and M. Weintraub. The DECIPHER speech recognition system. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 77–80, Albuquerque, 1990.
- [Cohen *et al.* 92]
M. Cohen, H. Franco, N. Morgan, D. Rumelhart, and V. Abrash. Hybrid neural network/hidden Markov model continuous-speech recognition. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 915–918, Banff, 1992.

- [Cohen *et al.* 03]
S. Cohen, J. Mamou, Y. Kanza, and Y. Sagiv. Xsearch: a semantic search engine for xml. In *Proceedings of the 29th international conference on Very large data bases - Volume 29, VLDB '2003*, pages 45–56. VLDB Endowment, 2003.
- [Cohen *et al.* 08]
K. B. Cohen, M. Palmer, and L. Hunter. Nominalization and alternations in biomedical language. *PLoS ONE*, 3(9):e3158, 09 2008.
- [Cohen *et al.* 09]
T. Cohen, R. Schvaneveldt, and D. Widdows. Reflective random indexing and indirect inference: A scalable method for discovery of implicit connections. *Journal of Biomedical Informatics*, 2009.
- [Cole 96]
R. Cole, editor. *Survey of the State of the Art in Human Language Technology*. Cambridge University Press, 1996.
- [Collier & Takeuchi 02]
N. Collier and K. Takeuchi. PIA-Core: Semantic Annotation through Example-based Learning. In *Proceedings of the LREC 2002 Third International Conference on Language Resources and Evaluation*, pages 1611–1614, Las Palmas, Canary Islands - Spain, 2002.
- [Collier 96]
R. Collier. Automatic Template Creation for Information Extraction. Technical report CS-96-07, Department of Computer Science, University of Sheffield, U.K., September 1996.
- [Collier 98]
R. Collier. *Automatic Template Creation for Information Extraction*. Unpublished PhD thesis, University of Sheffield, UK, 1998.
- [Collier *et al.* 02]
N. Collier, K. Takeuchi, C. Nobata, J. Fukumoto, and N. Ogata. Progress on multi-lingual named entity annotation guidelines using rdf(s). In *Proceedings of the 3rd Language Resources and Evaluation Conference (LREC'2002)*, Gran Canaria, Spain, 2002.
- [Collins 02]
M. Collins. Ranking algorithms for named entity extraction: Boosting and the voted perceptron. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL'02)*, Philadelphia, PA, 2002.
- [Collins 03]
M. Collins. Head-driven statistical models for natural language parsing. *Computational linguistics*, 29(4):589–637, 2003.
- [Collobert *et al.* 06]
R. Collobert, F. Sinz, J. Weston, and L. Bottou. Large Scale Transductive SVMs. *Journal of Machine Learning Research*, 7:1687–1712, 2006.
- [Coltheart *et al.* 79]
M. Coltheart, D. Besner, J. Jonasson, and E. Davelaar. Phonological encoding in a lexical decision task. *Quarterly Journal of Experimental Psychology*, 31:489–507, 1979.
- [Committee 96]
T. A. Committee. TIPSTER text phase II architecture concept. In *Proceedings of a workshop on TIPSTER held at Vienna, Virginia: May 6–8, 1996*, TIPSTER '96, pages 221–247, Stroudsburg, PA, USA, 1996. Association for Computational Linguistics.

- [Condamines & Rebeyrolle 98]
 A. Condamines and J. Rebeyrolle. CTKB: a corpus-based approach to a terminological knowledge base. In *Proc. of 1st Workshop on Computational Terminology (Computerm '98)*, pages 29–35, Montreal, Canada, 1998. COLING-ACL.
- [Connine *et al.* 84]
 C. Connine, F. Ferreira, C. Jones, C. Clifton, and L. Frazier. Verb frame preferences - descriptive norms. *Journal of Psycholinguistic Research*, 13(4):307–319, 1984.
- [Connolly & Sluckin 71]
 T. G. Connolly and W. Sluckin. *An Introduction to Statistics for the Social Sciences*. Macmillan, third edition, 1971.
- [Connolly 97]
 D. Connolly. *XML: Principles, Tools and Techniques*. O'Reilly, Sebastopol, California, 1997.
- [Conrad *et al.* 09]
 J. G. Conrad, J. L. Leidner, F. Schilder, and R. Kondadadi. Query-based opinion summarization for legal blog entries. In *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, pages 167–176, 2009.
- [Conroy *et al.* 06]
 J. M. Conroy, J. D. Schlesinger, D. P. O'leary, and J. Goldstein. Back to basics: Classy 2006. In *Proceedings of the Document Understanding Conference (DUC) 2006*, 2006.
- [Constantopoulos *et al.* 04]
 P. Constantopoulos, M. Doerr, M. Theodoridou, and M. Tzobanakis. On information organization in annotation systems. In G. Grieser and Y. Tanaka, editors, *Intuitive Human Interfaces for Organizing and Accessing Intellectual Assets*, pages 189–200, Dagstuhl, Germany, march 2004. Springer.
- [Cooke 96]
 G. Cooke. The SemanTag project. Further details at <http://www.rt66.com/gcooke/SemanTag>, 1996.
- [Cooper 94]
 I. Cooper. Indexing the World. Technical Report 10-94*, University of Kent, Computing Laboratory, University of Kent, Canterbury, UK, July 1994.
- [Cooper *et al.* 10]
 S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popović, and F. players. Predicting protein structures with a multiplayer online game. 2010.
- [Cootes *et al.* 95]
 T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models – their training and application. *Comput. Vis. Image Underst.*, 61(1):38–59, January 1995.
- [Cootes *et al.* 01]
 T. Cootes, G. Edwards, and C. Taylor. Active appearance models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(6):681–685, 2001.
- [Coppersmith *et al.* 14]
 G. Coppersmith, M. Dredze, and C. Harman. Quantifying mental health signals in Twitter. In *Association for Computational Linguistics Workshop of Computational Linguistics and Clinical Psychology*, 2014.
- [Corcho 06]
 O. Corcho. Ontology based document annotation trends and open research problems. *Int. J. Metadata Semant. Ontologies*, 1(1):47–57, 2006.

- [Cormen *et al.* 90]
 T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. MIT Press, Cambridge MA, 1990.
- [Cornolti *et al.* 13]
 M. Cornolti, P. Ferragina, and M. Ciaramita. A framework for benchmarking entity-annotation systems. In *Proceedings of the 22nd International Conference on World Wide Web, WWW '13*, pages 249–260, 2013.
- [Corston-Oliver 01]
 S. Corston-Oliver. Text Compaction for Display on Very Small Screens. In *Proceedings of the Workshop on Automatic Summarization (NAAACL 2001)*, pages 1–8, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA, 2001.
- [Costa & Branco 12]
 F. Costa and A. Branco. TimeBankPT: A TimeML Annotated Corpus of Portuguese. In *LREC*, pages 3727–3734, 2012.
- [Costa & Melucci 10]
 A. Costa and M. Melucci. An information retrieval model based on discrete fourier transform. In H. Cunningham, A. Hanbury, and S. Rüger, editors, *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference)*. LNCS volume number: 6107, Lecture Notes in Computer Science, Vienna, Austria, May 2010. Springer.
- [Costagliola *et al.* 95a]
 G. Costagliola, A. De Lucia, S. Orefice, and G. Tortora. Efficient Parsing of Data-Flow Graphs. In *Proceedings of 7th International Conference on Software Engineering and Knowledge Engineering, Rockville, Maryland, U.S.A., June 1995*, pages 226–233. Knowledge Systems Institute, 1995.
- [Costagliola *et al.* 95b]
 G. Costagliola, G. Tortora, S. Orefice, and A. De Lucia. Automatic Generation of Visual Programming Environments. *IEEE Computer*, 28(3):56–66, 1995.
- [Côté 93]
 R. Côté, editor. *Systematized Nomenclature of Medicine - SNOMED International*. College of American Pathologists, 1993.
- [Courant & Robin 85]
 M. Courant and S. Robin. Classified Advertisement Analysis in the Context of an Expert System in Ad Matching. In P. S. Dizier and V. Dahl, editors, *Natural Language Understanding and Logic Programming*, pages 33–47. North-Holland, 1985.
- [Couto *et al.* 06]
 F. Couto, M. Silva, V. Lee, E. Dimmer, E. Camon, R. Apweiler, H. Kirsch, and D. Rebholz-Schuhmann. Goannotator: linking protein go annotations to evidence text. *Journal of Biomedical Discovery and Collaboration*, 1(1):19, 2006.
- [Cowie & Lehnert 96]
 J. Cowie and W. Lehnert. Information Extraction. *Communications of the ACM*, 39(1):80–91, 1996.
- [Cowie *et al.* 92a]
 J. Cowie, J. Guthrie, and L. Guthrie. Lexical disambiguation using simulated annealing. In *Proc. of DARPA Workshop on Speech and Natural Language*, pages 238–242, New York, 1992.
- [Cowie *et al.* 92b]
 J. Cowie, L. Guthrie, and J. Guthrie. Lexical Disambiguation Using Simulated Annealing. In *Proceedings of COLING-92*, Nantes, France, 1992.

- [Cowie *et al.* 92c]
J. Cowie, L. Guthrie, and J. Guthrie. Lexical Disambiguation Using Simulated Annealing. In *Proceedings of the 14th International Conference on Computational Linguistics (COLING-92)*, pages 359–365, Nantes, France, 1992.
- [Cowie *et al.* 93a]
J. Cowie, L. Guthrie, W. Jin, W. Odgen, J. Pustejovsky, R. Wanf, T. Wakao, S. Waterman, and Y. Wilks. CRL/Brandeis: The Diderot System. In *Proceedings of Tipster Text Program (Phase I)*. Morgan Kaufmann, California, 1993.
- [Cowie *et al.* 93b]
J. Cowie, T. Wakao, W. Jin, J. Pustejovsky, and S. Waterman. The Diderot information extraction system. In *Proceedings of the First Conference of the Pacific Association for Computational Linguistics (PACLING 93)*. Vancouver, Canada, 1993.
- [Cowling 96]
T. Cowling. Personal Communication, February 1996. Dept. Computer Science, University of Sheffield., 1996.
- [Cox 90]
S. Cox. Hidden Markov Models for Automatic Speech Recognition: Theory and Application. In *Speech and Language Processing*. Chapman and Hall, London, 1990.
- [Crammer & Singer 01]
K. Crammer and Y. Singer. On the Algorithmic Implementation of Multi-class Kernel-based Vector Machines. *Journal of Machine Learning Research*, 2:265–292, 2001.
- [Crane 96]
G. Crane. Building a digital library: the perseus project as a case study in the humanities. In *Proceedings of the first ACM international conference on Digital libraries*, pages 3–10. ACM Press, 1996.
- [Crane 02]
G. Crane. Cultural Heritage Digital Libraries: Needs and Components. In *Proceedings of the 6th European Conference on Digital Libraries*, pages 626–637, Rome, Italy, September 2002. Springer-Verlag Berlin Heidelberg 2002.
- [Crane 06]
G. Crane. What do you do with a million books? *D-Lib Magazine*, 12(3), March 2006.
- [Crane *et al.* 01a]
G. Crane, R. F. Chavez, A. Mahoney, T. L. Milbank, J. A. Rydberg-Cox, D. A. Smith, and C. E. Wulfman. Drudgery and deep thought. *Communications of the ACM*, 44(5):34–40, 2001.
- [Crane *et al.* 01b]
G. Crane, D. A. Smith, and C. E. Wulfman. Building a hypertextual digital library in the humanities: a case study on london. In *Proceedings of the first ACM/IEEE-CS joint conference on Digital libraries*, pages 426–434. ACM Press, 2001.
- [Crane *et al.* 05]
G. Crane, K. Bontcheva, J. A. Rydberg-Cox, and C. E. Wulfman. Emerging language technologies and the rediscovery of the past: a research agenda. *International Journal on Digital Libraries*, 5(4):309–316, 2005.
- [Craven & Kumlien 99]
M. Craven and J. Kumlien. Constructing biological knowledge bases by extracting information from text sources. In *ISMB-99*, 1999.

- [Cregan *et al.* 07]
A. Cregan, R. Schwitter, T. Meyer, et al. Sydney owl syntax-towards a controlled natural language syntax for owl 1.1. In *OWLED*, volume 258, 2007.
- [Crescenzi *et al.* 01]
V. Crescenzi, G. Mecca, P. Merialdo, et al. Roadrunner: Towards automatic data extraction from large web sites. In *Proceedings of the international conference on very large data bases*, pages 109–118. Citeseer, 2001.
- [Crimi *et al.* 90]
C. Crimi, A. Guercio, G. Pacini, G. Tortora, and M. Tucci. Automating Visual Language Generation. *IEEE Transactions on Software Engineering*, 16(10):1122–1135, 1990.
- [Cristianini & Shawe-Taylor 00]
N. Cristianini and J. Shawe-Taylor. *An introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, 2000.
- [Cristianini 06]
N. Cristianini. *Introduction to Computational Genomics: A Case Studies Approach*. Cambridge University Press, 2006.
- [Cristianini *et al.* 02]
N. Cristianini, J. Shawe-Taylor, and H. Lodhi. Latent semantic kernels. *Journal of Intelligent Information System*, 18(2/3):127–152, 2002.
- [Croft *et al.* 09]
W. B. Croft, D. Metzler, and T. Strohman. *Search Engines – Information Retrieval in Practice*. Pearson Education, Boston, MA, 2009.
- [Crooks *et al.* 12]
Crooks, Croitoru, Stefanidis, and Radzikowski. #Earthquake: Twitter as a Distributed Sensor System. *Transactions on Geographic Information Systems*, 17(1):124, 2012.
- [Crookshank 23]
F. Crookshank. The importance of a theory of signs and a critique of language in the study of medicine. In C. Ogden and I. Richards, editors, *The Meaning of Meaning*, pages 337–355. Routledge & Kegan Paul Ltd., 1923. supplementary essay.
- [Crouch *et al.* 95]
R. Crouch, R. Gaizauskas, and K. Netter. Interim Report of the Study Group on Assessment and Evaluation. Technical report, EAGLES project, Language Engineering Programme, European Commission, 1995.
- [Crowe & Jack 87]
A. Crowe and M. A. Jack. A globally optimising formant tracker using generalised centroids. *Electronics Letters*, 23:1019–1020, 1987.
- [Crowe 96]
J. Crowe. *Constraint Based Event Recognition for Information Extraction*. Unpublished PhD thesis, University of Edinburgh, 1996.
- [Cruz *et al.* 89]
I. Cruz, A. Mendelzon, and P. Wood. G+: Recursive Queries without Recursion. In *Proceedings of the 2nd Expert Database Systems Conference*, pages 645–666. Benjamin-Cummings, 1989.
- [Crystal 91]
D. Crystal. *A Dictionary of Linguistics and Phonetics. 3rd Edn.* Blackwell Publishers, Oxford, 1991.

- [Cubranic *et al.* 02]
D. Cubranic, G. Murphy, and K. Booth. Hipikat: A Developer’s Recommender. In *OOPSLA*, 2002.
- [Cucerzan 07]
S. Cucerzan. Large-scale named entity disambiguation based on wikipedia data. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 708–716, 2007.
- [Cuddon 80]
J. Cuddon. *The Macmillan Dictionary of Sports and Games*. McMillan Reference Books, 1980.
- [Cumby & Roth 03]
C. Cumby and D. Roth. On Kernel Methods for Relational Learning. In *Proceedings of the 10th International Conference on Machine Learning (ICML-2003)*, pages 107–114, 2003.
- [Cunliffe *et al.* 97]
D. Cunliffe, C. Taylor, and D. Tudhope. Query-based Navigation in Semantically Indexed Hypermedia. In *Proceedings of the UK Conference on Hypertext*, pages 87–95, 245–246, 1997.
- [Cunningham & Bontcheva 03]
H. Cunningham and K. Bontcheva. Software Architectures for Language Engineering: a Critical Review. Research Memorandum CS–03–09, Department of Computer Science, University of Sheffield, 2003.
- [Cunningham & Bontcheva 05]
H. Cunningham and K. Bontcheva. Computational Language Systems, Architectures. *Encyclopedia of Language and Linguistics, 2nd Edition*, pages 733–752, 2005.
- [Cunningham & Scott 04a]
H. Cunningham and D. Scott. Introduction to the Special Issue on Software Architecture for Language Engineering. *Natural Language Engineering*, 2004. <http://gate.ac.uk/sale/jnle-sale/intro/intro-main.pdf>.
- [Cunningham & Scott 04b]
H. Cunningham and D. Scott, editors. *Special Issue of Natural Language Engineering on Software Architecture for Language Engineering*. Cambridge University Press, 2004.
- [Cunningham 84]
J. Cunningham. Comprehension by Model-Building as a Basis for an Expert System. Cognitive studies research paper 25, University of Sussex, 1984.
- [Cunningham 91]
H. Cunningham. Automatic speech recognition. MSc thesis, University of Manchester Institute of Science and Technology, 1991.
- [Cunningham 92]
H. Cunningham. Artificial Neural Networks for Pragmatics. In *Workshop on Abduction, Beliefs and Context (ABC-92)*, Alghero, Sardinia, 1992. ESPRIT Project P-5254 (PLUS).
- [Cunningham 93]
H. Cunningham. Managing Code Development. Technical Report 93/10, Centre for Computational Linguistics, UMIST, Manchester, 1993.
- [Cunningham 94]
H. Cunningham. Support Software for Language Engineering Research. Technical Report 94/05, Centre for Computational Linguistics, UMIST, Manchester, 1994.

- [Cunningham 96]
H. Cunningham. AVENTINUS, ECRAN and GATE – Information Extraction and Language Engineering. In *Proceedings of the Workshop on Multi-lingual Access to Textual Information (MAIN-96)*, Volterra, Italy, July 1996.
- [Cunningham 97]
H. Cunningham. Information Extraction – a User Guide. Research Memorandum CS-97-02, Department of Computer Science, University of Sheffield, January 1997. <http://xxx.lanl.gov/abs/cs.CL/9702006>.
- [Cunningham 98]
H. Cunningham. GATE to infrastructure. *ELNews*, 7(4), 1998.
- [Cunningham 99a]
H. Cunningham. A Definition and Short History of Language Engineering. *Journal of Natural Language Engineering*, 5(1):1–16, 1999.
- [Cunningham 99b]
H. Cunningham. Information extraction - a user guide. Technical Report CS-99-07, University of Sheffield, UK, Apr 1999.
- [Cunningham 99c]
H. Cunningham. Information Extraction: a User Guide (revised version). Research Memorandum CS-99-07, Department of Computer Science, University of Sheffield, May 1999.
- [Cunningham 99d]
H. Cunningham. JAPE: a Java Annotation Patterns Engine. Research Memorandum CS-99-06, Department of Computer Science, University of Sheffield, May 1999.
- [Cunningham 00]
H. Cunningham. *Software Architecture for Language Engineering*. Unpublished PhD thesis, Department of Computer Science, University of Sheffield, Sheffield, UK, 2000. <http://gate.ac.uk/sale/thesis/>.
- [Cunningham 01]
H. Cunningham. A General Architecture for Text Engineering. *Computing and the Humanities*, May 2001.
- [Cunningham 02a]
H. Cunningham. GATE, a General Architecture for Text Engineering. *Computers and the Humanities*, 36:223–254, 2002.
- [Cunningham 02b]
H. Cunningham. GATE, a General Architecture for Text Engineering. *Computers and the Humanities*, 36:223–254, 2002.
- [Cunningham 05]
H. Cunningham. Information Extraction, Automatic. *Encyclopedia of Language and Linguistics, 2nd Edition*, pages 665–677, dec 2005.
- [Cunningham 06]
G. Cunningham. Sumerian. *Encyclopedia of Language and Linguistics*, 12:271, 2006.
- [Cunningham 12]
H. Cunningham. Agile Research. *ArXiv e-prints*, <http://arxiv.org/abs/1202.0652v1>, February 2012.
- [Cunningham *et al.* 93]
H. Cunningham, N. Underwood, and W. Black. The Surface-level COBALT Analysis Module. LRE Project 61-011 Technical Report C/CCL/11/93/b, Centre for Computational Linguistics, UMIST, Manchester, 1993.

- [Cunningham *et al.* 94a]
H. Cunningham, M. Freeman, and W. Black. Software Reuse, Object-Oriented Frameworks and Natural Language Processing. In *New Methods in Language Processing (NeMLaP-1)*, 14-16 September 1994, pages 357–367, Manchester, 1994. UCL Press.
- [Cunningham *et al.* 94b]
H. Cunningham, N. Underwood, and W. Black. Efficient Shallow Parsing for Text Categorisation. Technical Report 94/02, Centre for Computational Linguistics, UMIST, Manchester, 1994.
- [Cunningham *et al.* 94c]
H. Cunningham, N. Underwood, and W. Black. Reusable Components for CRISTAL. Technical Report LRE Project 62-059 Deliverable 1, Centre for Computational Linguistics, UMIST, Manchester, 1994.
- [Cunningham *et al.* 95]
H. Cunningham, R. Gaizauskas, and Y. Wilks. A General Architecture for Text Engineering (GATE) – a new approach to Language Engineering R&D. Technical Report CS-95-21, Department of Computer Science, University of Sheffield, 1995. <http://xxx.lanl.gov/abs/cs.CL/9601009>.
- [Cunningham *et al.* 96a]
H. Cunningham, S. Azzam, and Y. Wilks. Domain Modelling for AVENTINUS (WP 4.2). LE project LE1-2238 AVENTINUS internal technical report, University of Sheffield, U.K., 1996.
- [Cunningham *et al.* 96b]
H. Cunningham, S. Azzam, Y. Wilks, K. Humphreys, and R. Gaizauskas. AVENTINUS Domain Model Specifications (WP 4.1/T12). LE project LE1-2238 AVENTINUS internal technical report, University of Sheffield, U.K., 1996.
- [Cunningham *et al.* 96c]
H. Cunningham, S. Azzam, Y. Wilks, K. Humphreys, and R. Gaizauskas. AVENTINUS Information Extraction Specifications (WP 4.1/T6). LE project LE1-2238 AVENTINUS internal technical report, University of Sheffield, U.K., 1996.
- [Cunningham *et al.* 96d]
H. Cunningham, K. Humphreys, R. Gaizauskas, and M. Stower. CREOLE Developer’s Manual. Technical report, Department of Computer Science, University of Sheffield, 1996. <http://www.dcs.shef.ac.uk/nlp/gate>.
- [Cunningham *et al.* 96e]
H. Cunningham, K. Humphreys, R. Gaizauskas, and Y. Wilks. TIPSTER-Compatible Projects at Sheffield. In *Advances in Text Processing, TIPSTER Program Phase II*. DARPA, Morgan Kaufmann, California, 1996.
- [Cunningham *et al.* 96f]
H. Cunningham, Y. Wilks, and R. Gaizauskas. GATE – a General Architecture for Text Engineering. In *Proceedings of the 16th Conference on Computational Linguistics (COLING-96)*, Copenhagen, August 1996. ftp://ftp.dcs.shef.ac.uk/home/hamish/auto_papers/Cun96b.ps.
- [Cunningham *et al.* 96g]
H. Cunningham, Y. Wilks, and R. Gaizauskas. Software Infrastructure for Language Engineering. In *Proceedings of the AISB Workshop on Language Engineering for Document Analysis and Recognition*, Brighton, U.K., April 1996.
- [Cunningham *et al.* 96h]
H. Cunningham, Y. Wilks, and R. Gaizauskas. New Methods, Current Trends and Software Infrastructure for NLP. In *Proceedings of the Conference on New Methods in Natural Language Processing (NeMLaP-2)*, Bilkent University, Turkey, September 1996. ftp://ftp.dcs.shef.ac.uk/home/hamish/auto_papers/Cun96c.ps.

- [Cunningham *et al.* 97a]
H. Cunningham, K. Humphreys, R. Gaizauskas, and Y. Wilks. GATE – a TIPSTER-based General Architecture for Text Engineering. In *Proceedings of the TIPSTER Text Program (Phase III) 6 Month Workshop*. DARPA, Morgan Kaufmann, California, May 1997. ftp://ftp.dcs.shef.ac.uk/home/hamish/auto_papers/Cun97e.ps.
- [Cunningham *et al.* 97b]
H. Cunningham, K. Humphreys, R. Gaizauskas, and Y. Wilks. Software Infrastructure for Natural Language Processing. In *Proceedings of the 5th Conference on Applied Natural Language Processing (ANLP-97)*, March 1997. ftp://ftp.dcs.shef.ac.uk/home/hamish/auto_papers/Cun97a.ps.gz.
- [Cunningham *et al.* 98a]
H. Cunningham, W. Peters, C. McCauley, K. Bontcheva, and Y. Wilks. A Level Playing Field for Language Resource Evaluation. In *Workshop on Distributing and Accessing Lexical Resources at Conference on Language Resources Evaluation, Granada, Spain, 1998*. <http://www.dcs.shef.ac.uk/hamish/dalr>.
- [Cunningham *et al.* 98b]
H. Cunningham, M. Stevenson, and Y. Wilks. Implementing a Sense Tagger within a General Architecture for Language Engineering. In *Proceedings of the Third Conference on New Methods in Language Engineering (NeMLaP-3)*, pages 59–72, Sydney, Australia, 1998.
- [Cunningham *et al.* 99]
H. Cunningham, R. Gaizauskas, K. Humphreys, and Y. Wilks. Experience with a Language Engineering Architecture: Three Years of GATE. In *Proceedings of the AISB'99 Workshop on Reference Architectures and Data Standards for NLP*, Edinburgh, April 1999. The Society for the Study of Artificial Intelligence and Simulation of Behaviour. <http://www.dcs.shef.ac.uk/hamish/GateAisb99.html>.
- [Cunningham *et al.* 00a]
H. Cunningham, K. Bontcheva, W. Peters, and Y. Wilks. Uniform language resource access and distribution in the context of a General Architecture for Text Engineering (GATE). In *Proceedings of the Workshop on Ontologies and Language Resources (OntoLex'2000)*, Sozopol, Bulgaria, September 2000. <http://gate.ac.uk/sale/ontolex/ontolex.ps>.
- [Cunningham *et al.* 00b]
H. Cunningham, K. Bontcheva, V. Tablan, and Y. Wilks. Software Infrastructure for Language Resources: a Taxonomy of Previous Work and a Requirements Analysis. In *Proceedings of the second International Conference on Language Resources and Evaluation (LREC 2000), 30 May – 2 Jun 2000*, pages 815–824, Athens, Greece, 2000.
- [Cunningham *et al.* 00c]
H. Cunningham, K. Bontcheva, V. Tablan, and Y. Wilks. Software Infrastructure for Language Resources: a Taxonomy of Previous Work and a Requirements Analysis. In *Proceedings of the 2nd International Conference on Language Resources and Evaluation (LREC-2)*, Athens, 2000.
- [Cunningham *et al.* 00d]
H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan, and Y. Wilks. Experience of using GATE for NLP R&D. In *Proceedings of the Workshop on Using Toolsets and Architectures To Build NLP Systems at COLING-2000*, Luxembourg, 2000. <http://gate.ac.uk/>.
- [Cunningham *et al.* 00e]
H. Cunningham, D. Maynard, and V. Tablan. JAPE: a Java Annotation Patterns Engine (Second Edition). Research Memorandum CS-00-10, Department of Computer Science, University of Sheffield, Sheffield, UK, November 2000.

- [Cunningham *et al.* 01]
P. Cunningham, R. Bergmann, S. Schmitt, R. Traphöner, S. Breen, and B. Smyth. Websell: Intelligent sales assistants for the world wide web. *KI - Zeitschrift für Künstliche Intelligenz, Schwerpunktthema: Electronic Commerce*, 1, 2001.
- [Cunningham *et al.* 02a]
H. Cunningham, K. Bontcheva, D. Maynard, and V. Tablan. GATE - A New Release. *ELSNews*, 11(1), 2002. <http://www.elsnet.org/publications/elsnews/11.1.pdf>.
- [Cunningham *et al.* 02b]
H. Cunningham, D. Maynard, K. Bontcheva, and V. Tablan. GATE: an Architecture for Development of Robust HLT Applications. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pages 168–175, 2002.
- [Cunningham *et al.* 02c]
H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan, and C. Ursu. *The GATE User Guide*. <http://gate.ac.uk/>, 2002.
- [Cunningham *et al.* 02d]
H. Cunningham, D. Maynard, K. Bontcheva, and V. Tablan. GATE: an Architecture for Development of Robust HLT Applications. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, 7–12 July 2002, ACL '02*, pages 168–175, Stroudsburg, PA, USA, 2002. Association for Computational Linguistics.
- [Cunningham *et al.* 02e]
H. Cunningham, D. Maynard, K. Bontcheva, and V. Tablan. Gate: an architecture for development of robust hlt applications. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, ACL '02*, pages 168–175, Stroudsburg, PA, USA, 2002. Association for Computational Linguistics.
- [Cunningham *et al.* 03a]
H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan, C. Ursu, and M. Dimitrov. Developing language processing components with GATE (a user guide). <http://gate.ac.uk>, Feb 2003.
- [Cunningham *et al.* 03b]
H. Cunningham, V. Tablan, K. Bontcheva, and M. Dimitrov. Language Engineering Tools for Collaborative Corpus Annotation. In *Proceedings of Corpus Linguistics 2003*, Lancaster, UK, 2003. <http://gate.ac.uk/sale/cl03/distrib-ollie-cl03.doc>.
- [Cunningham *et al.* 05a]
H. Cunningham, K. Bontcheva, and Y. Li. Knowledge Management and Human Language: Crossing the Chasm. *Journal of Knowledge Management*, 9(5):108–131, 2005.
- [Cunningham *et al.* 05b]
H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan, C. Ursu, M. Dimitrov, M. Dowman, N. Aswani, and I. Roberts. *Developing Language Processing Components with GATE Version 3 (a User Guide)*. <http://gate.ac.uk/>, 2005.
- [Cunningham *et al.* 06]
H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan, C. Ursu, M. Dimitrov, M. Dowman, N. Aswani, and I. Roberts. *Developing Language Processing Components with GATE Version 3.1 (a User Guide)*. <http://gate.ac.uk/>, 2006.
- [Cunningham *et al.* 09]
H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan, M. Dimitrov, M. Dowman, N. Aswani, I. Roberts, Y. Li, and A. Funk. *Developing Language Processing Components with GATE Version 5.0 (a User Guide)*. <http://gate.ac.uk/>, 2009.

- [Cunningham *et al.* 10a]
H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan, M. Dimitrov, M. Dowman, N. Aswani, I. Roberts, Y. Li, and A. Funk. *Developing Language Processing Components with GATE Version 6.0 (a User Guide)*. <http://gate.ac.uk/>, 2010.
- [Cunningham *et al.* 10b]
H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan, M. Dimitrov, M. Dowman, N. Aswani, I. Roberts, Y. Li, A. Funk, G. Gorrell, J. Petrak, H. Saggion, D. Damljanovic, and A. Roberts. *Developing Language Processing Components with GATE Version 5.0 (a User Guide)*. The University of Sheffield, 2010.
- [Cunningham *et al.* 10c]
H. Cunningham, A. Hanbury, and S. Rüger, editors. *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference, 31 May 2010)*. LNCS volume number: 6107, Lecture Notes in Computer Science, Vienna, Austria, May 2010. Springer.
- [Cunningham *et al.* 10d]
H. Cunningham, A. Hanbury, and S. Rüger. Scaling up high-value retrieval to medium-volume data. In H. Cunningham, A. Hanbury, and S. Rüger, editors, *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference, 31 May 2010)*. LNCS volume number: 6107, Lecture Notes in Computer Science, pages 1–5, Vienna, Austria, May 2010. Springer.
- [Cunningham *et al.* 11a]
H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan, N. Aswani, and I. Roberts. *Text Processing with GATE (Version 6)*. 2011.
- [Cunningham *et al.* 11b]
H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan, N. Aswani, I. Roberts, G. Gorrell, A. Funk, A. Roberts, D. Damljanovic, T. Heitz, M. Greenwood, H. Saggion, J. Petrak, Y. Li, and W. Peters. *Text Processing with GATE (Version 6)*. 2011.
- [Cunningham *et al.* 11c]
H. Cunningham, N. Fuhr, and B. Stein, editors. *Challenges in Document Mining – Report from Dagstuhl Seminar 11171*, Dagstuhl Reports, Dagstuhl, Germany, May 2011. Leibniz-Zentrum für Informatik.
- [Cunningham *et al.* 11d]
H. Cunningham, N. Fuhr, and B. M. Stein. Challenges in Document Mining (Dagstuhl Seminar 11171). *Dagstuhl Reports*, 1(4):65–99, 2011.
- [Cunningham *et al.* 11e]
H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan, N. Aswani, I. Roberts, G. Gorrell, A. Funk, A. Roberts, D. Damljanovic, T. Heitz, M. A. Greenwood, H. Saggion, J. Petrak, Y. Li, and W. Peters. *Text Processing with GATE (Version 6)*. The University of Sheffield, 2011.
- [Cunningham *et al.* 11f]
H. Cunningham, V. Tablan, I. Roberts, M. A. Greenwood, and N. Aswani. Information Extraction and Semantic Annotation for Multi-Paradigm Information Management. In M. Lupu, K. Mayer, J. Tait, and A. J. Trippe, editors, *Current Challenges in Patent Information Retrieval*, volume 29 of *The Information Retrieval Series*, pages 307–327. Springer Berlin Heidelberg, 2011.
- [Cunningham *et al.* 13a]
H. Cunningham, V. Tablan, A. Roberts, and K. Bontcheva. Getting more out of biomedical documents with GATE’s full lifecycle open source text analytics. *PLoS Computational Biology*, 9(2):e1002854, 02 2013.

- [Cunningham *et al.* 13b]
H. Cunningham, V. Tablan, A. Roberts, and K. Bontcheva. Getting more out of biomedical documents with gate's full lifecycle open source text analytics. *PLoS Computational Biology*, 9(2), 2013.
- [Curran & Clark 03]
J. R. Curran and S. Clark. Language independent ner using a maximum entropy tagger. In W. Daelemans and M. Osborne, editors, *Proceedings of CoNLL-2003*, pages 164–167. Edmonton, Canada, 2003.
- [Cutting *et al.* 91]
D. Cutting, J. Pedersen, and P.-K. Halvorsen. An Object-Oriented Architecture for Text Retrieval. In *Proceedings of RIAO '91*, pages 285–298, Barcelona, 1991.
- [Cutting *et al.* 92]
D. R. Cutting, J. O. Pedersen, D. Karger, and J. W. Tukey. Scatter/gather: A cluster-based approach to browsing large document collections. In *Proceedings of the Fifteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 318–329, 1992.
- [Daelemans & (eds.) 02]
W. Daelemans and M. O. (eds.). *Proceedings of CoNLL-2002*. 2002.
- [Daelemans & Hoste 02]
W. Daelemans and V. Hoste. Evaluation of Machine Learning Methods for Natural Language Processing Tasks. In *Proceedings of the LREC 2002 Third International Conference on Language Resources and Evaluation*, pages 755–760, Las Palmas, Canary Islands - Spain, 2002.
- [Daelemans & Osborne 03]
W. Daelemans and M. Osborne, editors. *CoNLL-2003, 7th Conference on Computational Natural Language Learning*, Edmonton, Canada, May 2003.
- [Daelemans *et al.* 98]
W. Daelemans, J. Zavrel, K. van der Sloot, and A. van den Bosch. Timbl: Tilburg memory based learner version 1.0. Technical Report Technical Report 98-03, ILK, 1998.
- [Dagan *et al.* 91]
I. Dagan, A. Itai, and U. Schwall. Two languages are more informative than one. In *Proc. of 29th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 130–137, Berkeley, California, 1991.
- [Dagan *et al.* 94]
I. Dagan, F. Pereira, and L. Lee. Similarity-based estimation of word co-occurrence probabilities. In *Proc. of the 32nd Annual Meeting of the Association for Computational Linguistics (ACL'94)*, pages 272–278, 1994.
- [Dagon & Itai 94]
I. Dagon and A. Itai. Word sense disambiguation using a second language monolingual corpus. *Computational Linguistics*, 20, 1994.
- [Dahab & Belz 10]
K. Dahab and A. Belz. A game-based approach to transcribing images of text. In *7th International Conference on Language Resources and Evaluation*, 2010.
- [Dahdul *et al.* 10]
W. M. Dahdul, J. P. Balhoff, J. Engeman, T. Grande, E. J. Hilton, C. Kothari, H. Lapp, J. G. Lundberg, P. E. Midford, T. J. Vision, M. Westerfield, and P. M. Mabee. Evolutionary characters, phenotypes and ontologies: Curating data from the systematic biology literature. *PLoS ONE*, 5(5):e10708, 05 2010.

- [Dahlgren 88a]
K. Dahlgren. *Naive Semantics for Natural Language Understanding*. Kluwer Academic, Amsterdam, Boston, 1988.
- [Dahlgren 88b]
K. Dahlgren. *Naive semantics for natural language understanding*. Kluwer, Boston, USA, 1988.
- [Dahlgren *et al.* 90]
K. Dahlgren, J. McDowell, and E. Stabler, Jr. Knowledge Representation for Commonsense Reasoning with Text. *Computational Linguistics*, 15(3):149–170, 1990.
- [Dai 11]
W. U. M. N. S. W. U. W. D. S. S. W. U. Dai, Peng (Seattle. Decision-theoretic control of crowd-sourced workflows, December 2011.
- [Dai *et al.* 10]
P. Dai, Mausam, and D. Weld. Decision-theoretic control of crowd-sourced workflows. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*, 2010.
- [Daiber *et al.* 13]
J. Daiber, M. Jakob, C. Hokamp, and P. N. Mendes. Improving efficiency and accuracy in multilingual entity extraction. In *Proc. of the 9th Int. Conf. on Semantic Systems, I-SEMANTICS '13*, pages 121–124, New York, NY, USA, 2013.
- [Daille 96]
B. Daille. Study and implementation of combined techniques for automatic extraction of terminology. *The Balancing Act: Combining Symbolic and Statistical Approaches to Language*, 1:49–66, 1996.
- [Daille *et al.* 94]
B. Daille, E. Gaussier, and J. Langé. Towards automatic extraction of monolingual and bilingual terminology. In *Proc. of 15th International Conference on Computational Linguistics (COLING)*, pages 515–521, 1994.
- [Dalby *et al.* 89]
J. Dalby, A. Crowe, and A. M. Sutherland. Formant-based vowel classification in continuous speech. In *Proceedings European Conference on Speech Communication and Technology*, volume 1, pages 244–247, 1989.
- [Dale *et al.* 92]
R. Dale, E. H. Hovy, D. Rösner, and O. Stock. *Aspects of Automated Natural Language Generation*. Lecture Notes in Artificial Intelligence, 587. Springer Verlag, Berlin, April 1992.
- [Dale *et al.* 98]
R. Dale, J. Oberlander, M. Milosavljevic, and A. Knott. Integrating natural language generation and hypertext to produce dynamic documents. *Interacting with Computers*, 11:109–135, 1998.
- [Dale *et al.* 00]
R. Dale, H. Moisl, and H. Somers, editors. *A Handbook of Natural Language Processing*. Dekker, New York, 2000.
- [Dalli 02]
A. Dalli. Creation and evaluation of extensible language resources for maltese. In *Proceedings of 3rd Language Resources and Evaluation Conference (LREC'2002)*, Gran Canaria, Spain, 2002.
- [Dalli *et al.* 04a]
A. Dalli, V. Tablan, K. Bontcheva, Y. Wilks, D. Broeder, H. Brugman, and P. Wittenburg. Web services architecture for language resources. In *Fourth International Conference on Language Resources and Evaluation (LREC'2004)*, Lisbon, Portugal, 2004.

- [Dalli *et al.* 04b]
 A. Dalli, Y. Xia, and Y. Wilks. Fasil email summarisation system. In *Proceedings of the 20th International Conference on Computational Linguistics*, COLING '04, Stroudsburg, PA, USA, 2004. Association for Computational Linguistics.
- [Damerau 93]
 F. J. Damerau. Generating and evaluating domain-oriented multi-word terms from texts. *Information Processing & Management*, 29(4):433–447, 1993.
- [Damljjanovic & Bontcheva 08]
 D. Damljjanovic and K. Bontcheva. Enhanced Semantic Access to Software Artefacts. In *Workshop on Semantic Web Enabled Software Engineering (SWESE)*, Karlsruhe, Germany, October 2008.
- [Damljjanovic & Bontcheva 09a]
 D. Damljjanovic and K. Bontcheva. Towards enhanced usability of natural language interfaces to knowledge bases. In V. Devedzic and D. Gasevic, editors, *Special issue on Semantic Web and Web 2.0*, volume 6, pages 105–133. Springer-Verlag, Berlin, Germany, 2009.
- [Damljjanovic & Bontcheva 09b]
 D. Damljjanovic and K. Bontcheva. Integrated testbed of case study 1. Technical Report D6.4, TAO Project Deliverable, 2009. <http://www.tao-project.eu/resources/publicdeliverables/d6-4.pdf>.
- [Damljjanovic & Bontcheva 12]
 D. Damljjanovic and K. Bontcheva. Named entity disambiguation using linked data. In *9th Extended Semantic Web Conference (ESWC2012)*, May 2012.
- [Damljjanovic & Devedzic 08]
 D. Damljjanovic and V. Devedzic. Applying semantic web to e-tourism. In Z. Ma and H. Wang, editors, *The Semantic Web for Knowledge and Data Management: Technologies and Practices*. Information Science Reference (IGI Global), 2008.
- [Damljjanovic & Devedzic 09]
 D. Damljjanovic and V. Devedzic. Semantic Web and E-tourism. In M. Khosrow-Pour, editor, *Encyclopedia of Information Science and Technology, Second edition*. IGI Global, 2009.
- [Damljjanovic 10]
 D. Damljjanovic. Towards Portable Controlled Natural Languages for Querying Ontologies. In M. Rosner and N. Fuchs, editors, *Second Workshop on Controlled Natural Languages*, volume 622 of *CEUR Workshop Pre-Proceedings ISSN 1613-0073*. <http://ceur-ws.org>, Marettimo Island, Italy, September 2010.
- [Damljjanovic *et al.* 08a]
 D. Damljjanovic, K. Bontcheva, M. Agatonovic, I. Roberts, and T. Heitz. Integrated testbed of case study 1. Technical Report D6.3, TAO Project Deliverable, 2008. <http://www.tao-project.eu/resources/publicdeliverables/d6-3.pdf>.
- [Damljjanovic *et al.* 08b]
 D. Damljjanovic, K. Bontcheva, V. Tablan, I. Roberts, M. Agatonovic, A. Shafirin, and J. Sun. Case study 1: Domain ontology and semantic augmentation of legacy content. Technical Report D6.2, TAO Project Deliverable, 2008. <http://www.tao-project.eu/resources/publicdeliverables/d6-2.pdf>.
- [Damljjanovic *et al.* 08c]
 D. Damljjanovic, V. Tablan, and K. Bontcheva. A Text-based Query Interface to OWL Ontologies. In *6th Language Resources and Evaluation Conference (LREC)*, Marrakech, Morocco, May 2008. ELRA.

- [Damljanovic *et al.* 09a]
D. Damljanovic, F. Amardeilh, and K. Bontcheva. CA Manager Framework: Creating Customised Workflows for Ontology Population and Semantic Annotation. In *Proceedings of The Fifth International Conference on Knowledge Capture (KCAP'09)*, California, USA, September 2009.
- [Damljanovic *et al.* 09b]
D. Damljanovic, M. Agatonovic, and H. Cunningham. Usability of Natural Language Interfaces for Querying Ontologies (poster). In *Workshop on Controlled Natural Language (CNL 2009)*, 2009.
- [Damljanovic *et al.* 10a]
D. Damljanovic, J. Petrak, and H. Cunningham. Random Indexing for Searching Large RDF Graphs. In *Poster Session at the Proceedings of the 7th Extended Semantic Web Conference (ESWC 2010)*, Lecture Notes in Computer Science, Heraklion, Greece, June 2010. Springer-Verlag.
- [Damljanovic *et al.* 10b]
D. Damljanovic, M. Agatonovic, and H. Cunningham. Identification of the Question Focus: Combining Syntactic Analysis and Ontology-based Lookup through the User Interaction. In *7th Language Resources and Evaluation Conference (LREC)*, La Valletta, Malta, May 2010. ELRA.
- [Damljanovic *et al.* 10c]
D. Damljanovic, M. Agatonovic, and H. Cunningham. Natural Language Interfaces to Ontologies: Combining Syntactic Analysis and Ontology-based Lookup through the User Interaction. In *Proceedings of the 7th Extended Semantic Web Conference (ESWC 2010)*, Lecture Notes in Computer Science, Heraklion, Greece, June 2010. Springer-Verlag.
- [Damljanovic *et al.* 11a]
D. Damljanovic, M. Agatonovic, and H. Cunningham. FREyA: an Interactive Way of Querying Linked Data using Natural Language. In *Proceedings of 1st Workshop on Question Answering over Linked Data (QALD-1), Collocated with the 8th Extended Semantic Web Conference (ESWC 2011)*, Heraklion, Greece, June 2011.
- [Damljanovic *et al.* 11b]
D. Damljanovic, J. Petrak, M. Lupu, H. Cunningham, M. Carlsson, G. Engstrom, and B. Andersson. Random Indexing for Finding Similar Nodes within Large RDF graphs . In *Proceedings of the Fourth International Workshop on Resource Discovery, Collocated with the 8th Extended Semantic Web Conference (ESWC 2011)*, Heraklion, Greece, June 2011.
- [Damljanovic *et al.* 12]
D. Damljanovic, M. Stankovic, and P. Laublet. Linked Data-based Concept Recommendation: Comparison of Different Methods in Open Innovation Scenario. In *Proceedings of the 9th Extended Semantic Web Conference (ESWC)*, 2012.
- [Damljanovic *et al.* 13]
D. Damljanovic, M. Agatonovic, H. Cunningham, and K. Bontcheva. Improving habitability of natural language interfaces for querying ontologies with feedback and clarification dialogues. *Web Semantics: Science, Services and Agents on the World Wide Web*, 19(0), 2013. <http://bit.ly/JWSsd>.
- [Darling *et al.* 12]
W. M. Darling, M. J. Paul, and F. Song. Unsupervised part-of-speech tagging in noisy and esoteric domains with a syntactic-semantic Bayesian HMM. In *Proceedings of the Workshop on Semantic Analysis in Social Media*, pages 1–9. Association for Computational Linguistics, 2012.
- [Darroch & Ratcliff 72]
J. N. Darroch and D. Ratcliff. Generalized iterative scaling for log-linear models. *Annals of Mathematical Statistics*, 43:1470–1480, 1972.

- [Das & Chen 01]
S. Das and M. Chen. Yahoo! for amazon: Extracting market sentiment from stock message boards. In *Proceedings of the Asia Pacific Finance Association Annual Conference (APFA)*, 2001.
- [Daumé & Marcu 06]
H. Daumé, III and D. Marcu. Domain adaptation for statistical classifiers. *J. Artif. Int. Res.*, 26(1):101–126, May 2006.
- [Daumé & Marcu 07]
H. Daumé and D. Marcu. Frustratingly easy domain adaptation. In *Proceedings of the Annual meeting of the Association for Computational Linguistics*, 2007.
- [Dave *et al.* 03]
K. Dave, S. Lawrence, and D. M. Pennock. Mining the peanut gallery: opinion extraction and semantic classification of product reviews. In *WWW '03: Proceedings of the 12th international conference on World Wide Web*, pages 519–528, New York, NY, USA, 2003. ACM.
- [Davidov *et al.* 10]
D. Davidov, O. Tsur, and A. Rappoport. Semi-supervised recognition of sarcastic sentences in Twitter and Amazon. In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning*, pages 107–116. Association for Computational Linguistics, 2010.
- [Davidson 67]
D. Davidson. The logical form of action sentences. In N. Rescher, editor, *The Logic of Decision and Action*. University of Pittsburgh Press, Pittsburgh, 1967.
- [Davidson *et al.* 98]
L. Davidson, J. Kavanagh, K. Mackintosh, I. Meyer, and D. Skuce. Semi-automatic extraction of knowledge-rich contexts from corpora. In *Proc. of 1st Workshop on Computational Terminology (Computerm '98)*, pages 50–56, Montreal, Canada, 1998. COLING-ACL.
- [Davies & Fleiss 82]
M. Davies and J. Fleiss. Measuring Agreement for Multinomial Data. *Biometrics*, 38:1047–1051, 1982.
- [Davies *et al.* 02]
J. Davies, D. Fensel, and F. van Harmelen, editors. *Towards the Semantic Web: Ontology-driven Knowledge Management*. Wiley, 2002.
- [Davies *et al.* 07]
J. Davies, M. Grobelnik, and D. Mladenic, editors. *Semantic Web: Integrating Ontology Management, Knowledge Discovery and Human Language Technologies*. Springer, 2007. Work in progress!
- [Davies *et al.* 13]
J. Davies, J. Hare, S. Samangoeei, J. Preston, N. Jain, D. Dupplaw, and P. H. Lewis. Identifying the geographic location of an image with a multimodal probability density function. In *MediaEval 2013 / Placing: Geo-coordinate Prediction for Social Multimedia*, October 2013.
- [Davis & Dunning 95]
D. Davis and T. Dunning. A trec evaluation of query translation methods for multi-lingual text retrieval. In *Proceedings of the Fourth Text Retrieval Evaluation Conference, NIST*, 1995.
- [Davis & Keller 82]
A. Davis and R. Keller. Data flow program graphs. *IEEE Computer*, 15:26–41, 1982.
- [Davis & Weyuker 83]
M. Davis and E. Weyuker. *Computability, Complexity, and Languages*. Academic Press, New York, 1983.

- [Davis *et al.* 06]
 B. Davis, S. Handschuh, H. Cunningham, and V. Tablan. Further use of Controlled Natural Language for Semantic Annotation of Wikis. In *Proceedings of the 1st Semantic Authoring and Annotation Workshop at ISWC2006*, Athens, Georgia, USA, November 2006.
- [Davis *et al.* 08]
 B. Davis, A. A. Iqbal, A. Funk, V. Tablan, K. Bontcheva, H. Cunningham, and S. Handschuh. Roundtrip ontology authoring. In *Proceedings of the 7th International Semantic Web Conference (ISWC)*, Karlsruhe, Germany, October 2008.
- [Davis *et al.* 09]
 B. Davis, P. Varma, S. Handschuh, L. Dragan, and H. Cunningham. Controlled natural language for semantic annotation. In N. E. Fuchs, editor, *Pre-Proceedings of the Workshop on Controlled Natural Language (CNL 2009)*, volume 448 of *CEUR Workshop Proceedings*. CEUR-WS, April 2009.
- [Davis *et al.* 10a]
 B. Davis, P. Dantuluri, S. Handschuh, and H. Cunningham. Towards controlled natural language for semantic annotation. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 6(4):64–91, 2010.
- [Davis *et al.* 10b]
 B. Davis, P. Dantuluri, L. Dragan, S. Handschuh, and H. Cunningham. On designing controlled natural languages for semantic annotation. In N. E. Fuchs, editor, *Proceedings of the Workshop on Controlled Natural Language (CNL 2009)*, Lecture Notes in Computer Science. Springer, 2010.
- [Davis *et al.* 12]
 A. Davis, A. Veloso, A. Soares, A. Laender, and W. Meira Jr. Named entity disambiguation in streaming data. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, pages 815–824, Jeju Island, Korea, July 2012. Association for Computational Linguistics.
- [Davison & Hinkley 97]
 A. C. Davison and D. V. Hinkley. *Bootstrap Methods And Their Application*. Cambridge University Press, Cambridge, UK, 1997.
- [Dawid & Skene 79]
 A. P. Dawid and A. M. Skene. Maximum likelihood estimation of observer error-rates using the em algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):pp. 20–28, 1979.
- [Day *et al.* 97]
 D. Day, J. Aberdeen, L. Hirschman, R. Kozierok, P. Robinson, and M. Vilain. Mixed-Initiative Development of Language Processing Systems. In *Proceedings of the 5th Conference on Applied Natural Language Processing (ANLP-97)*, 1997.
- [Day *et al.* 98]
 D. Day, P. Robinson, M. Vilain, and A. Yeh. MITRE: Description of the *Alembic* System Used for MUC-7. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)*, 1998.
- [Day *et al.* 04]
 D. Day, C. McHenry, R. Kozierok, and L. Riek. Callisto: A configurable annotation workbench. In *International Conference on Language Resources and Evaluation*, 2004.
- [DBL02]
Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, July 23-26, 2002, Edmonton, Alberta, Canada. ACM, 2002.

- [De Choudhury *et al.* 11a]
M. De Choudhury, S. Counts, and M. Czerwinski. Identifying relevant social media content: Leveraging information diversity and user cognition. In *Proceedings of the 22Nd ACM Conference on Hypertext and Hypermedia*, HT '11, pages 161–170, New York, NY, USA, 2011. ACM.
- [De Choudhury *et al.* 11b]
M. De Choudhury, S. Counts, and M. Czerwinski. Identifying relevant social media content: Leveraging information diversity and user cognition. In *HyperText*, 2011.
- [De Rassenfosse *et al.* 13]
G. De Rassenfosse, H. Dernis, D. Guellec, L. Picci, and B. v. P. de la Potterie. The worldwide count of priority patents: A new indicator of inventive activity. *Research Policy*, 42(3):720–737, 2013.
- [De Smedt 90]
K. De Smedt. IPF: An incremental parallel formulator. In R. Dale, C. Mellish, and M. Zock, editors, *Current Research in Natural Language Generation*, pages 167–192. Academic Press, 1990.
- [De Vries Robbé & Flier 94]
P. De Vries Robbé and F. Flier. Towards a common vocabulary for classification and definition. *Terminology*, 1(1):97–102, 1994.
- [Dean Jones & Visser 98]
T. B.-C. Dean Jones and P. Visser. Methodologies for ontology development. In *Proceedings IT&KNOWS Conference of the 15th IFIP World Computer Congress*, Budapest, 1998. Chapman-Hall.
- [Deane 05]
P. Deane. A nonparametric method for extraction of candidate phrasal terms. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pages 605–613, 2005.
- [deBernières 91]
L. de Bernières. *Señor Vivo and the Coca Lord*. Vintage, 1991.
- [deBruijn *et al.* 04]
J. de Bruijn, F. Martin-Recuerda, D. Manov, and M. Ehrig. State-of-the-art survey on Ontology Merging and Aligning v1. Technical report, SEKT project deliverable D4.2.1, 2004. <http://sw.deri.org/jos/sekt-d4.2.1-mediation-survey-final.pdf>.
- [deCea *et al.* 08]
G. A. de Cea, A. Gómez-Pérez, E. M. Ponsoda, and M.-C. Suárez-Figueroa. Natural language-based approach for helping in the reuse of ontology design patterns. In *Proceedings of the 16th International Conference on Knowledge Engineering and Knowledge Management Knowledge Patterns (EKAW 2008)*, Acitrezza, Italy, September 2008.
- [Decker & Frank 04]
S. Decker and M. Frank. The Social Semantic Desktop. Technical report, DERI Technical Report 2004-05-02, 2004.
- [Decker *et al.* 05]
B. Decker, E. Ras, J. Rech, B. Klein, and C. Hoecht. Self-Organized Reuse of Software Engineering Knowledge Supported by Semantic Wikis. In *Workshop on Semantic Web Enabled Software Engineering (SWESE)*, Galway, Ireland, 2005.
- [Declerck & Crispi 03]
T. Declerck and C. Crispi. Multilingual Linguistic Modules for IE Systems. In *Proceedings of Workshop on Information Extraction for Slavonic and other Central and Eastern European Languages (IESL'03)*, Borovets, Bulgaria, 2003.

- [Declerck & Krieger 06]
T. Declerck and H. Krieger. Translating XBRL into Description Logic: an approach using Protege, Sesame and OWL. In *Proceedings of Business Information Systems (BIS)*, Klagenfurt, Germany, 2006.
- [Declerck & Neumann 00]
T. Declerck and G. Neumann. Using a parameterisable and domain-adaptive information extraction system for annotating large-scale corpora? In *Proceedings of the Workshop Information Extraction meets Corpus Linguistics, LREC-2000*, 2000.
- [Declerck & Reed 03]
R. Declerck and S. Reed. Classifying conditionals: form and function. *English Language and Linguistics*, pages 309–323, 2003.
- [Declerck 01]
T. Declerck. Introduction: Extending nlp tools repositories for the interaction with language data resource repositories. In *Proceedings of the ACL 2001 Workshop on Sharing Tools and Resources*, pages 3–6, 2001.
- [Declerck *et al.* 01]
T. Declerck, P. Wittenburg, and H. Cunningham. The Automatic Generation of Formal Annotations in a Multimedia Indexing and Searching Environment. In *Workshop on Human Language Technology and Knowledge Management*, Toulouse, France, 2001. <http://www.elsnet.org/ac12001-hlt+km.html>.
- [Declerck *et al.* 03a]
T. Declerck, H. Cunningham, H. Saggion, J. Kuper, D. Reidsma, and P. Wittenburg. MUMIS – Advanced Information Extraction for Multimedia Indexing and Searching. In *Proceedings of WIAMIS*, London, 2003.
- [Declerck *et al.* 03b]
T. Declerck, J. Kuper, H. Saggion, A. Samiotou, and P. Wittenburg. Annotation for the semantic web. *Content-based indexing and searching in multimedia documents.*, 2003.
- [Declerck *et al.* 12]
T. Declerck, H.-U. Krieger, and D. Gromann. Acquisition, Representation, and Extension of Multilingual Labels of Financial Ontologies. In R. Costa, M. Silva, and A. L. Soares, editors, *Proceedings of the TKE Workshop “Challenges to Knowledge Representation in Multilingual Contexts”*, pages 17–26. TKE, 6 2012.
- [Deerwester *et al.* 90]
S. Deerwester, S. Dumais, G. Furnas, T. Landauer, and R. Harshman. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41:391–407, 1990.
- [DeJong 82]
G. DeJong. An overview of the FRUMP system. In W. Lehnert and M. Ringle, editors, *Strategies for Natural Language Processing*, pages 149–176. Lawrence Erlbaum, 1982.
- [Dekel *et al.* 04]
O. Dekel, J. Keshet, and Y. Singer. Large Margin Hierarchical Classification. In *Proceedings of the 21st International Conference on Machine Learning (ICML-2004)*, Canada, 2004.
- [deKinderen & Gordijn 07]
S. de Kinderen and J. Gordijn. A consumer needs-driven approach for finding it-service bundles. VU University Amsterdam, 2007.
- [deKinderen & Gordijn 08]
S. de Kinderen and J. Gordijn. e3service — a structured methodology for generating needs-driven it-service bundles in a networked enterprise. In *Symposium on Applied Computing (SAC’08)*, Fortaleza, Brazil, March 2008. ACM.

- [Della Pietra *et al.* 97]
S. Della Pietra, M. Epstein, S. Roukos, and T. Ward. Fertility Models for Statistical Natural Language Understanding. In *Proceedings of the Association for Computational Linguistics*, pages 168–173, July 1997.
- [Della Valle *et al.* 08]
E. Della Valle, D. Cerizza, I. Celino, A. Turati, H. Lausen, N. Steinmetz, M. Erdmann, and A. Funk. Realizing Service-Finder: Web service discovery at web scale. In *European Semantic Technology Conference (ESTC)*, Vienna, September 2008.
- [Dellschaft & Staab ar]
K. Dellschaft and S. Staab. On How to Perform a Gold Standard Based Evaluation of Ontology Learning. In *Proceedings of the 5th International Semantic Web Conference (ISWC'06)*, Athens, GA, USA, to appear.
- [deMarneffe *et al.* 06]
M.-C. de Marneffe, B. MacCartney, and C. Manning. Generating typed dependency parses from phrase structure parses. In *Language Resources and Evaluation Conference*, 2006.
- [Demartini *et al.* 12]
G. Demartini, D. E. Difallah, and P. Cudré-Mauroux. ZenCrowd: Leveraging probabilistic reasoning and crowdsourcing techniques for large-scale entity linking. In *Proceedings of the 21st Conference on World Wide Web*, pages 469–478, 2012.
- [Demidova *et al.* 13]
E. Demidova, D. Maynard, N. Tahmasebi, Y. Stavrakas, V. Plachouras, J. Hare, D. Dupplaw, and A. Funk. Extraction and Enrichment. Deliverable D3.3, ARCOMEM, 2013.
- [Demir *et al.* 10]
E. Demir, M. P. Cary, S. Paley, K. Fukuda, C. Lemer, I. Vastrik, G. Wu, P. D'Eustachio, C. Schaefer, J. Luciano, F. Schacherer, I. Martinez-Flores, Z. Hu, V. Jimenez-Jacinto, G. Joshi-Tope, K. Kandasamy, A. C. Lopez-Fuentes, H. Mi, E. Pichler, I. Rodchenkov, A. Splendiani, S. Tkachev, J. Zucker, G. Gopinath, H. Rajasimha, R. Ramakrishnan, I. Shah, M. Syed, N. Anwar, O. Babur, M. Blinov, E. Brauner, D. Corwin, S. Donaldson, F. Gibbons, R. Goldberg, P. Hornbeck, A. Luna, P. Murray-Rust, E. Neumann, O. Ruebenacker, M. Samwald, M. van Iersel, S. Wimalaratne, K. Allen, B. Braun, M. Whirl-Carrillo, K.-H. Cheung, K. Dahlquist, A. Finney, M. Gillespie, E. Glass, L. Gong, R. Haw, M. Honig, O. Hubaut, D. Kane, S. Krupa, M. Kutmon, J. Leonard, D. Marks, D. Merberg, V. Petri, A. Pico, D. Ravenscroft, L. Ren, N. Shah, M. Sunshine, R. Tang, R. Whaley, S. Letovksy, K. H. Buetow, A. Rzhetsky, V. Schachter, B. S. Sobral, U. Dogrusoz, S. McWeeney, M. Aladjem, E. Birney, J. Collado-Vides, S. Goto, M. Hucka, N. Le Novere, N. Maltsev, A. Pandey, P. Thomas, E. Wingender, P. D. Karp, C. Sander, and G. D. Bader. The BioPAX community standard for pathway data sharing. *Nature Biotechnology*, 28(9):935–942, September 2010.
- [Dempster *et al.* 77]
A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39:1–38, 1977.
- [Derczynski & Bontcheva 14a]
L. Derczynski and K. Bontcheva. PHEME: Veracity in Digital Social Networks. In *Proceedings of the UMAP Project Synergy workshop*, 2014.
- [Derczynski & Bontcheva 14b]
L. Derczynski and K. Bontcheva. Spatio-temporal grounding of claims made on the web, in PHEME. In *Proceedings of the 10th joint ACL-ISO workshop on Interoperable Semantic Annotation*. ACL, 2014.

- [Derczynski & Bontcheva 14c]
L. Derczynski and K. Bontcheva. Passive-aggressive sequence labeling with discriminative post-editing for recognising person entities in tweets. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, volume 2, pages 69–73, 2014.
- [Derczynski & Gaizauskas 10]
L. Derczynski and R. Gaizauskas. Analysing Temporally Annotated Corpora with CAVaT. In *Proceedings of the 7th International Conference on Language Resources and Evaluation*, pages 398–404, 2010.
- [Derczynski & Gaizauskas 11]
L. Derczynski and R. Gaizauskas. A Corpus-based Study of Temporal Signals. In *Proceedings of the 6th Corpus Linguistics Conference*, 2011.
- [Derczynski & Gaizauskas 13a]
L. Derczynski and R. Gaizauskas. Temporal signals help label temporal relations. In *Proceedings of the 51st meeting of the Association for Computational Linguistics (ACL 2013)*, volume 2, pages 645–650, 2013.
- [Derczynski & Gaizauskas 13b]
L. Derczynski and R. Gaizauskas. Empirical Validation of Reichenbach’s Tense Framework. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013)*, pages 71–82, 2013.
- [Derczynski *et al.* 13a]
L. Derczynski, D. Maynard, N. Aswani, and K. Bontcheva. Microblog-Genre Noise and Impact on Semantic Annotation Accuracy. In *Proceedings of the 24th ACM Conference on Hypertext and Social Media*. ACM, 2013.
- [Derczynski *et al.* 13b]
L. Derczynski, A. Ritter, S. Clark, and K. Bontcheva. Twitter Part-of-Speech Tagging for All: Overcoming Sparse and Noisy Data. In *Proceedings of Recent Advances in Natural Language Processing (RANLP)*. Association for Computational Linguistics, 2013.
- [Derczynski *et al.* 13c]
L. Derczynski, B. Yang, and C. Jensen. Towards Context-Aware Search and Analysis on Social Media Data. In *Proceedings of the 16th Conference on Extending Database Technology*. ACM, 2013.
- [Derczynski *et al.* 14]
L. Derczynski, C. Field, and K. Bøgh. DKIE: Open source information extraction for Danish. In S. Wintner, M. Tadia, and B. Babych, editors, *Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pages 61–64. Association for Computational Linguistics, 2014.
- [Derczynski *et al.* 15a]
L. Derczynski, D. Maynard, G. Rizzo, M. van Erp, G. Gorrell, R. Troncy, and K. Bontcheva. Analysis of named entity recognition and linking for tweets. *Information Processing and Management*, 51:32–49, 2015.
- [Derczynski *et al.* 15b]
L. Derczynski, J. Strötgen, R. Campos, and O. Alonso. Time and information retrieval: Introduction to the special issue. *Information Processing & Management*, 51, 2015.
- [Derczynski *et al.* 16]
L. Derczynski, K. Bontcheva, and I. Roberts. Broad twitter corpus: A diverse named entity recognition resource. In *COLING*, pages 1169–1179. ACL, 2016.

- [Derczynski *et al.* 17]
L. Derczynski, K. Bontcheva, M. Liakata, R. Procter, G. W. S. Hoi, and A. Zubiaga. Semeval-2017 task 8: RumourEval: Determining rumour veracity and support for rumours. In *SemEval@ACL*, pages 69–76, 2017.
- [deRoeck *et al.* 98]
A. deRoeck, U. Kruschwitz, P. Neal, P. Scott, S. Steel, R. Turner, and N. Webb. An intelligent directory enquiry assistant. *BT Technology Journal*, 16(3), 1998.
- [Devanbu *et al.* 90]
P. T. Devanbu, R. J. Brachman, P. G. Selfridge, and B. W. Ballard. Lassie—a knowledge-based software information system. In *ICSE '90: Proceedings of the 12th international conference on Software engineering*, pages 249–261, Los Alamitos, CA, USA, 1990. IEEE Computer Society Press.
- [Devaraju & Kuhn 10]
A. Devaraju and W. Kuhn. A process-centric ontological approach for integrating geo-sensor data. In *Proceedings of the Sixth International Conference on Formal Ontology in Information Systems (FOIS)*, pages 199–212, 2010.
- [Devitt & Ahmad 07]
A. Devitt and K. Ahmad. Sentiment polarity identification in financial news: A cohesion-based approach. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 984–991, Prague, Czech Republic, June 2007. Association for Computational Linguistics.
- [Devlin 98]
K. Devlin. *Mathematics: The New Golden Age, Second Edition*. Penguin Books, London, 1998.
- [DFKI 99]
DFKI. The Natural Language Software Registry. <http://www.dfki.de/lt/registry/>, 1999.
- [Diab & Resnik 02]
M. Diab and P. Resnik. An Unsupervised Method for Word Sense Tagging using Parallel Corpora. In *Proceedings of the Association for Computational Linguistics*, pages 255–262, July 2002.
- [Diakopoulos *et al.* 10]
N. Diakopoulos, M. Naaman, and F. Kivran-Swaine. Diamonds in the rough: Social media visual analytics for journalistic inquiry. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, pages 115–122, 2010.
- [Dice 45]
L. R. Dice. Measures of the Amount of Ecologic Association Between Species. *Ecology*, 3(26):297–302, 1945.
- [Diederich *et al.* 88]
J. Diederich, I. Ruhmann, and M. May. KRITON: a knowledge acquisition tool for expert systems. In B. R. Gaines and J. H. Boose, editors, *Knowledge-based Systems: Vol 2: Knowledge Acquisition Tools for Expert Systems*. Academic Press, London, 1988.
- [Diederich *et al.* 07]
J. Diederich, M. Dzbor, and D. Maynard. REASE: The Repository for Learning Units about the Semantic Web. *New Review of Hypermedia and Multimedia: Special Issue on Studying the Users of Digital Education Technologies: Theories, Methods, and Analytical Approaches*, 2007.
- [Dietterich & Michalski 83]
T. Dietterich and R. Michalski. A comparative review of selected methods for learning from examples. In J. Carbonell, R. Michalski, and T. Mitchell, editors, *Machine Learning: An Artificial Intelligence Approach*. Tioga Pub. Co., Palo Alto, 1983.

- [Dietze *et al.* 12a]
S. Dietze, D. Maynard, E. Demidova, T. Risse, W. Peters, K. Doka, and Y. Stavarakas. Preservation of social web content based on entity extraction and consolidation. pages 18–29, 2012.
- [Dietze *et al.* 12b]
S. Dietze, D. Maynard, N. Tahmasebi, Y. Stavarakas, V. Plachouras, E. Demidova, J. Hare, D. Dupplaw, A. Funk, W. Peters, and P. Siehndel. Extraction and Enrichment. Deliverable D3.2, ARCOMEM, 2012.
- [Dietzel & Maynard 15]
A. Dietzel and D. Maynard. Climate change: A chance for political re-engagement? In *Proc. of the Political Studies Association 65th Annual International Conference*, 2015.
- [Difallah *et al.* 13]
D. E. Difallah, G. Demartini, and P. Cudré-Mauroux. Pick-a-crowd: tell me what you like, and I’ll tell you what to do. In *Proceedings of the 22nd international conference on World Wide Web*, pages 367–374. ACM, 2013.
- [Dikaiakos *et al.* 09]
M. D. Dikaiakos, D. Katsaros, P. Mehra, G. Pallis, and A. Vakali. Cloud computing: Distributed internet computing for IT and scientific research. *IEEE Internet Computing*, 13(5):10–13, 2009.
- [Dill *et al.* 03a]
S. Dill, N. Eiron, D. Gibson, D. Gruhl, R. Guha, A. Jhingran, T. Kanungo, S. Rajagopalan, A. Tomkins, J. A. Tomlin, and J. Y. Zien. SemTag and Seeker: Bootstrapping the semantic web via automated semantic annotation. In *Proceedings of WWW’03*, 2003.
- [Dill *et al.* 03b]
S. Dill, J. A. Tomlin, J. Y. Zien, N. Eiron, D. Gibson, D. Gruhl, R. Guha, A. Jhingran, T. Kanungo, S. Rajagopalan, and A. Tomkins. SemTag and Seeker: Bootstrapping the semantic web via automated semantic annotation. In *Proceedings of the 12th International Conference on World Wide Web (WWW2003)*, pages 178–186, Budapest, Hungary, May 2003.
- [Dill *et al.* 03c]
S. Dill, N. Eiron, D. Gibson, D. Gruhl, R. Guha, A. Jhingran, T. Kanungo, K. S. McCurley, S. Rajagopalan, A. Tomkins, et al. A case for automated large-scale semantic annotation. *Web Semantics: Science, Services and Agents on the World Wide Web*, 1(1):115–132, 2003.
- [DiMasi 02]
J. A. DiMasi. The value of improving the productivity of the drug development process: Faster times and better decisions. *Pharmacoeconomics*, 20(3):1–10, 2002.
- [Dimitrov 02a]
M. Dimitrov. *A Light-weight Approach to Coreference Resolution for Named Entities in Text*. MSc Thesis, University of Sofia, Bulgaria, 2002. <http://www.ontotext.com/ie/thesis-m.pdf>.
- [Dimitrov 02b]
M. Dimitrov. *A Light-weight Approach to Coreference Resolution for Named Entities in Text*. MSc Thesis, University of Sofia, Bulgaria, 2002. <http://www.ontotext.com/ie/thesis-m.pdf>.
- [Dimitrov *et al.* 02]
M. Dimitrov, K. Bontcheva, H. Cunningham, and D. Maynard. A Light-weight Approach to Coreference Resolution for Named Entities in Text. In *Proceedings of the Fourth Discourse Anaphora and Anaphor Resolution Colloquium (DAARC)*, Lisbon, 2002.
- [Dimitrov *et al.* 04]
M. Dimitrov, K. Bontcheva, H. Cunningham, and D. Maynard. A Light-weight Approach to Coreference Resolution for Named Entities in Text. In A. Branco, T. McEnery, and R. Mitkov, editors, *Anaphora Processing: Linguistic, Cognitive and Computational Modelling*. John Benjamins, 2004.

- [Dimitrov *et al.* 14]
M. Dimitrov, H. Cunningham, I. Roberts, P. Kostov, A. Simov, P. Rigaux, and H. Lippell. *Anno-Market – Multilingual Text Analytics at Scale on the Cloud*, pages 315–319. Springer International Publishing, Cham, 2014.
- [Dimitrova *et al.* 04]
N. Dimitrova, J. Zimmerman, A. Janevski, L. Agnihotri, N. Haas, D. Li, R. Bolle, S. Velipasalar, T. McGee, and L. Nikolovska. Media Personalisation and Augmentation Through Multimedia Processing and Information Extraction. In L. Ardissono, A. Kobsa, and M. Maybury, editors, *[Ardissono et al. 04]*, chapter 8, pages 201–233. Kluwer, 2004.
- [Ding *et al.* 04]
L. Ding, T. Finin, A. Joshi, R. Pan, R. S. Cost, Y. Peng, P. Reddivari, V. C. Doshi, and J. Sachs. Swoogle: A Search and Metadata Engine for the Semantic Web. In *Proceedings of the Thirteenth ACM Conference on Information and Knowledge Management*, 2004.
- [Ding *et al.* 05]
L. Ding, T. Finin, Y. Peng, P. P. da Silva, and D. L. McGuinness. Tracking RDF Graph Provenance using RDF Molecules. Technical Report TR-CS-05-06, University of Maryland, Baltimore, MD, US, 2005.
- [Ding *et al.* 06]
Y. Ding, D. Embley, and S. Liddle. Automatic Creation and Simplified Querying of Semantic Web Content: An Approach Based on Information-Extraction Ontologies. In *Proceedings of the 1st Asian Semantic Web Conference*, pages 400–414. Springer Berlin/Heidelberg, September 2006.
- [Ding *et al.* 09]
X. Ding, B. Liu, and L. Zhang. Entity discovery and assignment for opinion mining applications. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1125–1134. ACM, 2009.
- [Dingli *et al.* 03]
A. Dingli, F. Ciravegna, and Y. Wilks. Automatic Semantic Annotation using Unsupervised Information Extraction and Integration. In *Workshop on Knowledge Markup and Semantic Annotation*, 2003.
- [Dipper *et al.* 04]
S. Dipper, M. Götze, and M. Stede. Simple Annotation Tools for Complex Annotation Tasks: an Evaluation. In *Proceedings of the LREC Workshop on XML-based Richly Annotated Corpora*, pages 54–62, Lisbon, Portugal, 2004.
- [Dixon 91]
R. Dixon. *A New Approach to English Grammar on Semantic Principles*. Clarendon Press, Oxford, 1991.
- [Do Amaral & Satomura 95]
M. Do Amaral and Y. Satomura. Associating semantic grammars with the SNOMED: Processing medical language and representing clinical facts into a language-independent frame. In *Proc. of 8th World Congress on Medical Informatics (Medinfo)*, pages 18–22, Edmonton, Alberta, 1995.
- [Doan *et al.* 11]
A. Doan, R. Ramakrishnan, and A. Y. Halevy. Crowdsourcing systems on the world-wide web. *Commun. ACM*, 54(4):86–96, April 2011.
- [Doddingtion *et al.* 04]
G. R. Doddingtion, A. Mitchell, M. A. Przybocki, L. A. Ramshaw, S. Strassel, and R. M. Weischedel. The Automatic Content Extraction (ACE) Program - Tasks, Data, and Evaluation. In *Proceedings of LREC*, volume 2, page 1, 2004.

- [Doerr & Crofts 98]
M. Doerr and N. Crofts. Electronic Communication on Diverse Data - The Role of the oo CIDOC Reference Model. In *Proceedings of the CIDOC'98 Conference*, Melbourne, Australia, October 1998.
- [Dolan 94]
W. B. Dolan. Word sense ambiguity: clustering related senses. In *Proceedings of COLING94*, pages 712–716, 1994.
- [Dolbear *et al.* 07]
C. Dolbear, G. Hart, K. Kovacs, J. Goodwin, and S. Zhou. The Rabbit language: description, syntax and conversion to OWL. *Ordnance Survey Research Labs Technical Report*, 2007.
- [Domingue *et al.* 04]
J. Domingue, M. Dzbor, and E. Motta. Magpie: Supporting Browsing and Navigation on the Semantic Web. In N. Nunes and C. Rich, editors, *Proceedings ACM Conference on Intelligent User Interfaces (IUI)*, pages 191–197, 2004.
- [Domingue *et al.* 08]
J. Domingue, L. Cabral, S. Galizia, V. Tanasescu, A. Gugliotta, B. Norton, and C. Pedrinaci. IRS-III: A broker-based approach to semantic web services. *Journal of Web Semantics*, 6(2):109–132, April 2008.
- [Donaldson *et al.* 03]
I. Donaldson, J. Martin, B. de Bruijn, C. Wolting, V. Lay, B. Tuekam, S. Zhang, B. Baskin, G. Bader, K. Michalickova, T. Pawson, and C. Hogue. Prebind and textomy - mining the biomedical literature for protein-protein interactions using a support vector machine. *BMC Bioinformatics*, 4(1):11, 2003.
- [Dong & Dong 00]
Dong and Dong. Hownet. <http://www.keenage.com>, 2000.
- [Dong & Searls 94]
S. Dong and D. B. Searls. Gene structure prediction by linguistic methods. *Genomics* 23, pages 540–551, 1994.
- [Dong *et al.* 04]
X. Dong, A. Halevy, J. Madhavan, E. Nemes, and J. Zhang. Similarity search for web services. In *International Conference on Very Large Databases (VLDB)*, Toronto, 2004.
- [Dong *et al.* 05]
X. Dong, A. Halevy, and J. Madhavan. Reference reconciliation in complex information spaces. In *SIGMOD '05: Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, pages 85–96, New York, NY, USA, 2005. ACM Press.
- [Dörk *et al.* 08]
M. Dörk, M. S. T. Carpendale, C. Collins, and C. Williamson. Visgets: Coordinated visualizations for web-based information exploration and discovery. *IEEE Trans. Vis. Comput. Graph.*, 14(6):1205–1212, 2008.
- [Dork *et al.* 10]
M. Dork, D. Gruen, C. Williamson, and S. Carpendale. A visual backchannel for large-scale events. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1129–1138, November 2010.
- [Dorow *et al.* 05]
B. Dorow, D. Widdows, K. Ling, J.-P. Eckmann, D. Sergi, and E. Moses. Using Curvature and Markov Clustering in Graphs for Lexical Acquisition and Word Sense Discrimination. In *The Second Workshop organized by the MEANING Project (MEANING-2005)*, 2005.

- [Dorr & Jones 96]
 B. Dorr and D. Jones. Role of word sense disambiguation in lexical acquisition: predicting semantics from syntactic cues. In *Proc. of 16th International Conference on Computational Linguistics (COLING)*, Copenhagen, 1996.
- [Dorr et al. 03]
 B. Dorr, D. Zajic, and R. Schwartz. Hedge trimmer: A parse-and-trim approach to headline generation. In D. Radev and S. Teufel, editors, *Proceedings of the HLT-NAACL 03 Text Summarization Workshop*, pages 1–8, 2003.
- [Douglas & Hurst 96]
 S. Douglas and M. Hurst. Controlled language support for Perkins Approved Clear English (PACE). In *CLAW96: Proceedings of the First International Workshop on Controlled Language Applications*, pages 26–27, Leuven, Belgium, 1996.
- [Douglas 07]
 N. Douglas. Twitter blows up at SXSW Conference. *Gawker.com*, 2007. <http://gawker.com/243634/twitter-blows-up-at-sxsw-conference?tag=technextbigthing>.
- [Douglas et al. 95]
 E. Douglas, J. Hobbs, J. Bear, D. Israel, M. Kameyama, A. Kehler, D. Martin, K. Myers, and M. Tyson. SRI International FASTUS System MUC-6 Test results and Analysis. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*. Morgan Kaufmann, California, 1995.
- [Douglis 10]
 F. Douglis. Thanks for the fish - but I'm drowning! *IEEE Internet Computing*, 14:4–6, 2010.
- [Douthat 98]
 A. Douthat. The Message Understanding Conference Scoring Software User's Manual. http://www.itl.nist.gov/iaui/894.02/related_projects/muc_sw/muc_sw_manual.html, 1998.
- [Dowman et al. 05a]
 M. Dowman, V. Tablan, H. Cunningham, and B. Popov. Content augmentation for mixed-mode news broadcasts. In *Proceedings of the 3rd European Conference on Interactive Television: User Centred ITV Systems, Programmes and Applications*, Aalborg University, Denmark, 2005. <http://gate.ac.uk/sale/euro-itv-2005/content-augmentation-for-mixed-mode-news-broadcast-consumption.pdf>.
- [Dowman et al. 05b]
 M. Dowman, V. Tablan, H. Cunningham, and B. Popov. Web-assisted annotation, semantic indexing and search of television and radio news. In *Proceedings of the 14th International World Wide Web Conference*, Chiba, Japan, 2005.
- [Dowman et al. 05c]
 M. Dowman, V. Tablan, H. Cunningham, C. Ursu, and B. Popov. Semantically enhanced television news through web and video integration. In *Second European Semantic Web Conference (ESWC'2005)*, 2005.
- [Downey et al. 04]
 D. Downey, O. Etzioni, S. Soderland, and D. S. Weld. Learning Text Patterns for Web Information Extraction and Assessment. In *Proceedings of AAAI 2004 Workshop on Adaptive Text Extraction and Mining (ATEM'04)*, 2004. <http://www.cs.washington.edu/research/knowitall/papers/DowneyATEM04.pdf>.
- [Downey et al. 07]
 D. Downey, M. Broadhead, and O. Etzioni. Locating complex named entities in web text. In *In Proc. of IJCAI*, 2007.

- [Dowty *et al.* 81]
D. Dowty, R. Wall, and P. Peters. *Introduction to Montague Semantics*. Reidel, Dordrecht, 1981.
- [Doyle 79]
J. Doyle. A truth maintenance system. *Artificial Intelligence* 12, pages 231–272, 1979.
- [Draicchio *et al.* 13]
F. Draicchio, A. Gangemi, V. Presutti, and A. G. Nuzzolese. FRED: from natural language text to RDF and owl in one click. In *Extended Semantic Web Conference*, pages 263–267. Springer, 2013.
- [Draper 86]
S. Draper. Does connectionism constitute a paradigm revolution ? Technical memo, experimental psychology, University of Sussex, 1986.
- [Drouin & Davis 09]
M. Drouin and C. Davis. R u txtng? is the use of text speak hurting your literacy? *Journal of Literacy Research*, 41(1):46–67, 2009.
- [Duan *et al.* 10]
Y. Duan, L. Jiang, T. Qin, M. Zhou, and H.-Y. Shum. An empirical study on learning to rank of tweets. In *COLING*, pages 295–303, 2010.
- [Duboue & McKeown 01a]
P. A. Duboue and K. R. McKeown. Empirically Estimating Order Constraints for Content Planning in Generation. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL-2001)*, 2001.
- [Duboue & McKeown 01b]
P. A. Duboue and K. R. McKeown. Empirically estimating order constraints for content planning in generation. In *Proceedings of ACL-EACL*, Toulouse, July 2001.
- [Dubuc & Lauriston 97]
R. Dubuc and A. Lauriston. Terms and contexts. In S. Wright and G. Budin, editors, *Handbook of Terminology Management*, volume 1: Basic Aspects of Terminology Management, chapter 1.3.3, pages 80–87. John Benjamins, Amsterdam, 1997.
- [DUC 01]
NIST. *Proceedings of the Document Understanding Conference*, September 13 2001.
- [Dumais 91]
S. Dumais. Improving the retrieval of information from external sources. *Behavior Research Methods, Instruments, & Computers*, 23(2):229–236, 1991.
- [Dunning 93a]
T. Dunning. Accurate Methods for the Statistics of Surprise and Coincidence. *Computational Linguistics*, 19(1), March 1993.
- [Dunning 93b]
T. Dunning. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1):61–74, 1993.
- [Dunning 94]
T. Dunning. Statistical Identification of Language. Technical Report Technical Report MCCA-94-273, Computing Research Laboratory, New Mexico State, 1994.
- [Durbin *et al.* 98]
R. Durbin, S. R. Eddy, A. Krogh, and G. Mitchison. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge University Press, 1998.

- [Dybkjær *et al.* 98]
L. Dybkjær, N. Bernsen, H. Dybkjær, D. McKelvie, and A. Mengel. The MATE Markup Framework. MATE Deliverable D1.2. Technical Report D1.2, MATE Project, <http://mate.nis.sdu.dk/>, 1998.
- [Dyer *et al.* 10]
C. Dyer, J. Weese, H. Setiawan, A. Lopez, F. Ture, V. Eidelman, J. Ganitkevitch, P. Blunsom, and P. Resnik. cdec: A decoder, alignment, and learning framework for finite-state and context-free translation models. In *Proceedings of the ACL 2010 System Demonstrations*, pages 7–12. Association for Computational Linguistics, 2010.
- [Dzbor & Motta 06]
M. Dzbor and E. Motta. Study on Integrating Semantic Applications with Magpie. In *15th Intl. Conference on Artificial Intelligence: Methods, Systems & Applications (AIMSA)*, Bulgaria, 2006.
- [Dzbor & Stutt 05]
M. Dzbor and A. Stutt. Prototype of advanced learning platform (ASPL-v1). Technical Report D3.3.3, KnowledgeWeb Deliverable, 2005.
- [Dzbor *et al.* 04]
M. Dzbor, E. Motta, and J. Domingue. Opening up magpie via semantic services. In *Proceedings of ISWC 2004*, Hiroshima, Japan, 2004.
- [EAGLES 99]
EAGLES. EAGLES recommendations. <http://www.ilc.pi.cnr.it/EAGLES96/browse.html>, 1999.
- [EAGLES Computational Lexicons Working Group 96]
EAGLES Computational Lexicons Working Group. Preliminary recommendations on semantic encoding. Technical report, EAGLES, 1996. <http://www.ilc.pi.cnr.it/EAGLES96/rep2/>.
- [Earley 70]
J. Earley. An efficient context-free parsing algorithm. *CACM*, 1970.
- [Easley & Kleinberg 10]
D. Easley and J. Kleinberg. *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press, 2010.
- [Ebeling 06]
J. Ebeling. The electronic text corpus of sumerian literature, 2006.
- [(EC) 10]
E. C. (EC). Europe 2020: A strategy for smart, sustainable and inclusive growth. *Working paper {COM (2010) 2020}*, 2010.
- [Edmondson & Iles 94a]
W. Edmondson and J. Iles. A Non-linear Architecture for Speech and Natural Language Processing. In *Proceedings of International Conference on Spoken Language Processing (ICSLP '94)*, volume 1, pages 29–32, Yokohama, Japan, 1994.
- [Edmondson & Iles 94b]
W. Edmondson and J. Iles. Pantome: An architecture for speech and natural language processing. Paper distributed at Dept. of CS Seminar, Sheffield University., 1994.
- [Edmondson 69]
H. P. Edmondson. New methods in automatic extracting. *J. ACM*, 16(2):264–285, April 1969.
- [Ehrlich & Shami 10]
K. Ehrlich and N. S. Shami. Microblogging inside and outside the workplace. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, pages 42–49. AAAI, 2010.

- [Eickhoff & deVries 12]
C. Eickhoff and A. de Vries. Increasing cheat robustness of crowdsourcing tasks. *Information Retrieval*, pages 1–17, 2012. 10.1007/s10791-011-9181-9.
- [Eisenstein 13]
J. Eisenstein. What to do about bad language on the internet. In *Proceedings of NAACL-HLT*, pages 359–369, 2013.
- [Eisenstein *et al.* 10a]
J. Eisenstein, B. O’Connor, N. Smith, and E. Xing. Discovering sociolinguistic associations with structured sparsity. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 2010.
- [Eisenstein *et al.* 10b]
J. Eisenstein, B. O’Connor, N. Smith, and E. Xing. A latent variable model for geographic lexical variation. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 1277–1287, 2010.
- [Eisenstein *et al.* 12]
J. Eisenstein, D. H. P. Chau, A. Kittur, and E. Xing. Topicviz: Semantic navigation of document collections. In *CHI Work-in-Progress Paper (Supplemental Proceedings)*, 2012.
- [Ejerhed & Dagan 96]
E. Ejerhed and I. Dagan, editors. *Proceedings of the Fourth Workshop on Very Large Corpora*, Copenhagen, 1996. Association for Computational Linguistics.
- [El-Haj *et al.* 10]
M. El-Haj, U. Kruschwitz, and C. Fox. Using Mechanical Turk to create a corpus of Arabic summaries. In *Proceedings of the Seventh conference on International Language Resources and Evaluation*, 2010.
- [ELA99]
ELAN – European Language Activity Network. <http://solaris3.ids-mannheim.de/elan/>, 1999.
- [Elcock 83]
E. Elcock. How complete are knowledge representation systems. *IEEE Computer*, October, 1983.
- [Ellingsworth & Sullivan 03]
M. Ellingsworth and D. Sullivan. Text mining improves business intelligence and predictive modeling in insurance. *DM Review Magazine*, 2003.
- [Elliott & Brzezinski 98]
C. Elliott and J. Brzezinski. Autonomous Agents as Synthetic Characters. *AI Magazine*, 19(2):13–30, 1998.
- [Ellman 97]
J. Ellman. Agent assisted internet resource discovery by generic document profiling. Unpublished report, 1997. <http://osiris.sunderland.ac.uk/cs0jel/research/index.htm>.
- [Ellman *et al.* 97]
J. Ellman, H. Somers, J. Nivre, and A. Multari. Foreign Language Information Extraction: An Application in the Employment Domain. In *Natural Language Processing: Extracting Information for Business Needs*, pages 77–89, London, March 1997. Unicom Seminars Ltd.
- [Elmasri & Navathe 94]
R. Elmasri and S. Navathe. *Fundamentals of Database Systems*. Benjamin/Cummings, Redwood City, California, 1994.

- [Embley *et al.* 98]
 D. Embley, D. Campbell, R. Smith, and S. Liddle. Ontology-based extraction and structuring of information from data-rich unstructured documents. In *Proceedings of the 7th International Conference on Information and Knowledge Management*, pages 52 – 59, Bethesda, MD, 1998.
- [ENC00]
 Article: Soccer. <http://encarta.msn.com>, September 2000.
- [Enguehard & Pantera 94]
 C. Enguehard and L. Pantera. Automatic natural acquisition of a terminology. *Journal of Quantitative Linguistics*, 2(1):27–32, 1994.
- [Ennals *et al.* 10]
 R. Ennals, B. Trushkovsky, and J. Agosta. Highlighting disputed claims on the web. In *WWW'10*, 2010.
- [Epstein *et al.* 96]
 M. Epstein, K. Papineni, S. Roukos, T. Ward, and S. Della Pietra. Statistical Natural Language Understanding Using Hidden Clumpings. In *Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 1, pages 176–179, Atlanta, Georgia, May 1996.
- [Erbach *et al.* 97]
 G. Erbach, G. Neumann, and H. Uszkoreit. Mulinex - multilingual indexing, editing and navigation extensions for the world wide web. In D. Hull and D. Oard, editors, *Cross-Language Text and Speech Retrieval – Papers from the 1997 AAAI Spring Symposium*. AAAI Press, Menlo Park, 1997.
- [Eriksson & Gambäck 97]
 M. Eriksson and B. Gambäck. SVENSK: A Toolbox of Swedish Language Processing Resources. In *Proceedings of the 2nd Conference on Recent Advances in Natural Language Processing (RANLP-2)*, Tzigrav Chark, Bulgaria, 1997.
- [Eriksson 96]
 M. Eriksson. ALEP. <http://www.sics.se/humle/projects/svensk/platforms.html>, 1996.
- [Eriksson 97]
 M. Eriksson. Final Report of Svensk. Technical report, SICS, <http://www.sics.se/humle/projects/svensk/>, 1997.
- [Erk & Pado 06]
 K. Erk and S. Pado. Shalmaneser – a flexible toolbox for semantic role assignment. In *Proceedings of LREC*, volume 6, 2006.
- [Erkan & Radev 04]
 G. Erkan and D. R. Radev. Lexrank: graph-based lexical centrality as salience in text summarization. *Journal Artificial Intelligence Research*, 22(1):457–479, 2004.
- [Erman & Lesser 75]
 L. Erman and V. Lesser. A multi-level organisation for problem solving using many, diverse, cooperating sources of knowledge. In *Proceedings of IJCAI-75*, pages 483–490, 1975.
- [Erman *et al.* 80]
 L. Erman, F. Hayes-Roth, V. Lesser, and D. Reddy. The Hearsay II speech understanding system: integrating knowledge to resolve uncertainty. *Computing Surveys*, 12, 1980.
- [ERO00]
 Soccer dictionary. <http://www.users.erols.com/soccertip/dictionary>, September 2000.

- [Escuderoz *et al.* 04]
G. Escuderoz, L. Marquez, and G. Rigau. TALP System for the English Lexical Sample Task. In *SENSEVAL-3: Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text*, 2004.
- [Estival *et al.* 97]
D. Estival, A. Lavelli, K. Netter, and F. Pianesi, editors. *Computational Environments for Grammar Development and Linguistic Engineering*. Association for Computational Linguistics, July 1997. Madrid, ACL-EACL'97.
- [Esuli & Sebastiani 06a]
A. Esuli and F. Sebastiani. SentiWordNet: A publicly available lexical resource for opinion mining. In *Proceedings of LREC 2006*, 2006.
- [Esuli & Sebastiani 06b]
A. Esuli and F. Sebastiani. SENTIWORDNET: A publicly available lexical resource for opinion mining. In *Proceedings of LREC-06, 5th Conference on Language Resources and Evaluation*, pages 417–422, Genova, IT, 2006.
- [Etzioni *et al.* 04a]
O. Etzioni, M. Cafarella, D. Downey, S. Kok, A. Popescu, T. Shaked, S. Soderland, D. S. Weld, and A. Yates. Web-scale Information Extraction in KnowItAll. In *Proceedings of WWW-2004*, 2004. <http://www.cs.washington.edu/research/knowitall/papers/www-paper.pdf>.
- [Etzioni *et al.* 04b]
O. Etzioni, M. Cafarella, D. Downey, A. Popescu, T. Shaked, S. Soderland, D. S. Weld, and A. Yates. Methods for Domain-Independent Information Extraction from the Web: An Experimental Comparison. In *Proceedings of AAAI-2004*, 2004. <http://www.cs.washington.edu/homes/etzioni/papers/aaai04.pdf>.
- [Etzioni *et al.* 05]
O. Etzioni, M. Cafarella, D. Downey, A. Popescu, T. Shaked, S. Soderland, D. Weld, and A. Yates. Unsupervised named-entity extraction from the web: An experimental study. *Artificial Intelligence*, 165(1):91–134, 2005.
- [Etzioni *et al.* 08]
O. Etzioni, M. Banko, S. Soderland, and D. Weld. Open information extraction from the web. *Communications of the ACM*, 51(12):68–74, 2008.
- [Eugenio & Glass 04]
B. D. Eugenio and M. Glass. The kappa statistic: a second look. *Computational Linguistics*, 1(30), 2004. (squib).
- [Eur00]
Euro 2000. <http://www.euro2000.com>, October 2000.
- [Euzenat 05]
J. Euzenat. Evaluating ontology alignment methods. In Y. Kalfoglou, M. Schorlemmer, A. Sheth, S. Staab, and M. Uschold, editors, *Semantic Interoperability and Integration*, number 04391 in Dagstuhl Seminar Proceedings, 2005.
- [Euzenat *et al.* 06]
J. Euzenat, L. Laera, V. Tamma, and A. Viollet. Negotiation/argumentation techniques among agents complying to different ontologies. Technical Report D2.3.7, KnowledgeWeb Deliverable, 2006.
- [Evans & Gazdar 96]
R. Evans and G. Gazdar. DATR: A Language for Lexical Knowledge Representation. *Computational Linguistics*, 22(1), 1996.

- [Evans & Hartley 90a]
R. Evans and A. F. Hartley. The traffic information collator. *Expert Systems: The International Journal of Knowledge Engineering*, 7(4):209–214, 1990.
- [Evans & Hartley 90b]
R. Evans and A. F. Hartley. The traffic information collator. *Expert Systems: The International Journal of Knowledge*, 7(4):209–214, 1990.
- [Evans & Kilgariff 95]
R. Evans and A. Kilgariff. Mrds, dictionaries, and how to do lexical engineering. In *Proceedings of the 2nd language engineering convention*, pages 125–132. London, 1995.
- [Evans & Zhai 96]
D. A. Evans and C. Zhai. Noun phrase analysis in unrestricted text for information retrieval. In *Proc. of 34th Annual Meeting of the Association for Computational Linguistics (ACL'96)*, pages 17–24, Santa Cruz, USA, 1996.
- [Evans *et al.* 95]
R. Evans, R. Gaizauskas, L. Cahill, J. Walker, J. Richardson, and A. Dixon. Poetic: A system for gathering and disseminating traffic information. *Journal of Natural Language Engineering*, 1(4), 1995.
- [Evans *et al.* 06]
D. M. Evans, J. Marchini, A. P. Morris, and L. R. Cardon. Two-stage two-locus models in genome-wide association. *PLoS Genetics*, 2(9, e157):1424–1432, September 2006.
- [Evans *et al.* 09]
N. Evans, M. Asahar, and Y. Matsumoto. Cocytus: parallel nlp over disparate data. *Traitement Automatique des Langues*, 49:2:271–293, 2009.
- [Eysenck & Keane 90]
M. Eysenck and M. Keane. *Cognitive Psychology: A Student's Handbook*. Lawrence Erlbaum Associates, Hove, U.K., 1990.
- [F. Paradis & Tajarobi 05]
J.-Y. N. F. Paradis and A. Tajarobi. Discovery of business opportunities on the internet with information extraction. In *Workshop on Multi-Agent Information Retrieval and Recommender Systems (IJCAI)*, pages 47–54, Edinburgh, Scotland, 2005.
- [Facebook 11]
Facebook. Statistics. <http://www.facebook.com/press/info.php?statistics>. Accessed on July 21st, 2011., 2011.
- [Falagas *et al.* 08]
M. E. Falagas, E. I. Pitsouni, G. A. Malietzis, and G. Pappas. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *The FASEB Journal*, 22(2):338–342, February 2008.
- [Fallside *et al.* 90]
F. Fallside, H. Lucke, T. P. Marsland, T. P. O'Shea, M. S. J. Owen, R. W. Prager, A. J. Robinson, and N. H. Russell. Continuous speech recognition for the TIMIT database using neural networks. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 445–448, Albuquerque, 1990.
- [Fan & Kambhampati 05]
J. Fan and S. Kambhampati. A snapshot of public web services. *SIGMOD Record*, 34(1):24–32, 2005.

- [Faridani *et al.* 10]
 S. Faridani, E. Bitton, K. Ryokai, and K. Goldberg. Opinion space: A scalable tool for browsing on-line comments. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems (CHI)*, pages 1175–1184, 2010.
- [Farmakiotou *et al.* 02]
 D. Farmakiotou, V. Karkaletsis, I. Koutsias, G. Petasis, and C. D. Spyropoulos. PatEdit: An Information Extraction Pattern Editor for Fast System Customization. In *Proceedings of the LREC 2002 Third International Conference on Language Resources and Evaluation*, pages 1097–1102, Las Palmas, Canary Islands - Spain, 2002.
- [Farquhar *et al.* 96]
 A. Farquhar, R. Fickas, and J. Rice. The Ontolingua Server: a tool for collaborative ontology construction. In *Proceedings of the 10th Banff Knowledge Acquisition Workshop*, Banff, Alberta, Canada, 1996.
- [Farwell & Wilks 89]
 D. Farwell and Y. Wilks. ULTRA - a machine translation system. Memoranda in Computer and Cognitive Science NM1989, Computing Research Lab, New Mexico State University, 1989.
- [Farzindar & Inkpen 12]
 A. Farzindar and D. Inkpen, editors. *Proceedings of the Workshop on Semantic Analysis in Social Media*. Association for Computational Linguistics, Avignon, France, April 2012.
- [Fass 97]
 D. Fass. Processing metonymy and metaphor. In A. Lesgold and V. Patel, editors, *Contemporary Studies in Cognitive Science and Technology*, volume 1. Ablex Publishing Corporation, Greenwich, 1997.
- [Favoron 07]
 R. Favoron. Do you have a business-oriented architecture? BPTrends Web site, May 2007.
- [Fayad *et al.* 00]
 M. Fayad, M. Laitinen, and R. Ward. Software Engineering in the Small. *Communications of the ACM*, 43(3):115–118, 2000.
- [Fayruzov *et al.* 08]
 T. Fayruzov, M. D. Cock, C. Cornelis, and V. Hoste. DEEPER: A Full Parsing Based Approach to Protein Relation Extraction. In E. Marchiori and J. H. Moore, editors, *EvoBIO 2008*, volume 4973 of *Lecture Notes in Computer Science*, pages 36–47. Springer, 2008.
- [Fayyad *et al.* 96]
 U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *AI Magazine*, pages 37–54, 1996.
- [Fazzinga *et al.* 11]
 B. Fazzinga, G. Gianforme, G. Gottlob, and T. Lukasiewicz. Semantic web search based on ontological conjunctive queries. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(4), 2011.
- [Federico *et al.* 00]
 M. Federico, D. Giordani, and P. Coletti. Development and Evaluation of an Italian Broadcast News Corpus. In *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC)*, Athens, Greece, 2000.
- [Feitelson 04]
 D. G. Feitelson. On identifying name equivalences in digital libraries. *Information Research*, 9(4), 2004.

- [Felber 84]
H. Felber. *Terminology Manual*. Unesco and Infoterm, Paris, 1984.
- [Fellbaum 98]
C. Fellbaum, editor. *WordNet - An Electronic Lexical Database*. MIT Press, 1998.
- [Feng & Wang 13a]
W. Feng and J. Wang. Retweet or not?: Personalized tweet re-ranking. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining, WSDM '13*, pages 577–586, New York, NY, USA, 2013. ACM.
- [Feng & Wang 13b]
W. Feng and J. Wang. Retweet or not?: Personalized tweet re-ranking. In *WSDM*, pages 577–586, 2013.
- [Feng *et al.* 09]
D. Feng, S. Besana, and R. Zajac. Acquiring high quality non-expert knowledge from on-demand workforce. In *Proceedings of the 2009 Workshop on The People's Web Meets NLP: Collaboratively Constructed Semantic Resources*, pages 51–56, 2009.
- [Fensel 01]
D. Fensel. *Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce*. Springer-Verlag, 2001.
- [Fensel *et al.* 02a]
D. Fensel, J. Hendler, W. Wahlster, and H. Lieberman, editors. *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. MIT Press, 2002.
- [Fensel *et al.* 02b]
D. Fensel, W. Wahlster, H. Lieberman, and J. Hendler, editors. *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. MIT Press, 2002.
- [Fensel *et al.* 08]
D. Fensel, F. van Harmelen, B. Andersson, P. Brennan, H. Cunningham, E. Della Valle, F. Fischer, Z. Huang, A. Kiryakov, T. K. Lee, L. School, V. Tresp, S. Wesner, M. Witbrock, and N. Zhong. Towards larkc: a platform for web-scale reasoning. In *Proceedings of the IEEE International Conference on Semantic Computing (ICSC 2008)*, Santa Clara, CA, USA, 2008. IEEE Computer Society Press.
- [Fernandes *et al.* 13]
A. C. Fernandes, D. Cloete, M. T. Broadbent, R. D. Hayes, C.-K. Chang, A. Roberts, J. Tsang, M. Soncul, J. Liebscher, R. G. Jackson, R. Stewart, and F. Callard. Development and evaluation of a de-identification procedure for a case register sourced from mental health electronic records. *BMC Medical Informatics and Decision Making*, 2013. Accepted for publication.
- [Fernandez *et al.* 07]
N. Fernandez, J. M. Blazquez, L. Sanchez, and A. Bernardi. IdentityRank: Named Entity Disambiguation in the Context of the NEWS Project. In *Proceedings of the European Semantic Web Conference (ESWC'07), LNCS 4519*, pages 604–654, 2007.
- [Fernandez *et al.* 08]
M. Fernandez, V. Lopez, M. Sabou, V. Uren, D. Vallet, E. Motta, and P. Castells. Semantic search meets the web. In *Semantic Computing, 2008 IEEE International Conference on*, pages 253–260, aug. 2008.
- [Fernández *et al.* 11]
M. Fernández, I. Cantador, V. López, D. Vallet, P. Castells, and E. Motta. Semantically enhanced information retrieval: An ontology-based approach. *Web Semantics*, 9(4):434–452, December 2011.

- [Fernandez *et al.* 14]
M. Fernandez, A. Scharl, K. Bontcheva, and H. Alani. User profile modelling in online communities. In *Proceedings of the Third International Conference on Semantic Web Collaborative Spaces – Volume 1275*, pages 1–15. CEUR-WS. org, 2014.
- [Fernández *et al.* 15]
M. Fernández, G. Burel, H. Alani, L. S. G. Piccolo, C. Meili, and R. Hess. Analysing engagement towards the 2014 earth hour campaign in twitter. In *EnviroInfo & ICT4S 2015: Building the Knowledge Base for Environmental Action and Sustainability*, Copenhagen, Denmark, 2015.
- [Ferragina & Scaiella 10]
P. Ferragina and U. Scaiella. Tagme: On-the-fly annotation of short text fragments (by wikipedia entities). In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, CIKM '10, pages 1625–1628, New York, NY, USA, 2010. ACM.
- [Ferrante & Rackoff 79]
J. Ferrante and C. Rackoff. *The Computational Complexity of Logical Theories*. Springer-Verlag, Berlin, 1979.
- [Ferrucci & Lally 03]
D. Ferrucci and A. Lally. Accelerating Corporate Research in the Development, Application, and Deployment of Human Language Technologies. In *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, 2003.
- [Ferrucci & Lally 04]
D. Ferrucci and A. Lally. UIMA: An Architectural Approach to Unstructured Information Processing in the Corporate Research Environment. *Natural Language Engineering*, 10(3-4):327–348, 2004.
- [Ferrucci *et al.* 09]
D. Ferrucci, A. Lally, K. Verspoor, and E. Nyberg. Unstructured information management architecture (UIMA) version 1.0. OASIS standard, OASIS, <https://www.oasis-open.org/>, March 2009.
- [Fikes & Farquhar 99]
R. Fikes and A. Farquhar. Distributed Repositories of Highly Expressive Reusable Ontologies. *IEEE Intelligent Systems*, 14(2):73–79, 1999.
- [Finch 94]
S. Finch. Exploiting sophisticated representations for document retrieval. In *4th Conference on Applied NLP*, pages 65–71. Stuttgart, October 1994.
- [Finch *et al.* 95]
S. Finch, H. Thompson, and D. McKelvie. Specification of Tool Shell with Discussion of Data and Process Architecture. LRE 62-050 Deliverable 1.1.2, 1995.
- [Finin *et al.* 94]
T. Finin, R. Fritzson, D. McKay, and R. McEntire. KQML – A Language and Protocol for Knowledge and Information Exchange. Technical Report CS-94-02, Computer Science Department, University of Maryland, Baltimore, 1994.
- [Finin *et al.* 10]
T. Finin, W. Murnane, A. Karandikar, N. Keller, J. Martineau, and M. Dredze. Annotating named entities in Twitter data with crowdsourcing. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, pages 80–88, 2010.
- [Fink *et al.* 08]
C. Fink, C. Piatko, J. Mayfield, T. Finin, and J. Martineau. Geolocating blogs from their textual content. In *Working Notes of the AAAI Spring Symposium on Social Semantic Web: Where Web 2.0 Meets Web 3.0*, pages 1–2. AAAI Press, 2008.

- [Finkel *et al.* 05]
 J. Finkel, T. Grenager, and C. Manning. Incorporating non-local information into information extraction systems by Gibbs sampling. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics*, pages 363–370. Association for Computational Linguistics, 2005.
- [Finkler & Neumann 89]
 W. Finkler and G. Neumann. POPEL-HOW: A distributed parallel model for incremental natural language production with feedback. pages 1518–1523, 1989.
- [Finn & Kushmerick 03]
 A. Finn and N. Kushmerick. Active learning selection strategies for information extraction. In *ECML-03 Workshop on Adaptive Text Extraction and Mining*, 2003.
- [Firmin & Chrzanowski 99]
 T. Firmin and M. Chrzanowski. An Evaluation of Automatic Text Summarization Systems. In I. Mani and M. Maybury, editors, *Advances in Automatic Text Summarization*, pages 325–336. The MIT Press, 1999.
- [Fischer *et al.* 96]
 D. Fischer, W. Mohr, and L. Rostek. A Modular, Object-Oriented and Generic Approach for Building Terminology Maintenance Systems. In *TKE '96: Terminology and Knowledge Engineering*, pages 245–258, Frankfurt, 1996.
- [Fiscus *et al.* 98]
 J. G. Fiscus, G. Doddington, J. S. Garofolo, and A. Martin. Nist’s 1998 topic detection and tracking evaluation (tdt2). In *Proc. of the DARPA Broadcast News Workshop*, Virginia, US, 1998.
- [Fleischman & Hovy 02]
 M. Fleischman and E. Hovy. Fine grained classification of named entities. In *Proceedings of the 19th International Conference on Computational Linguistics (COLING'02)*, Taipei, Taiwan, 2002.
- [Fleiss 75]
 J. L. Fleiss. Measuring agreement between two judges on the presence or absence of a trait. *Biometrics*, 31:651–659, 1975.
- [Fleiss 81]
 J. L. Fleiss. *Statistical Methods for Rates and Proportions*. John Wiley & Sons, New York, 1981.
- [Florian *et al.* 03]
 R. Florian, A. Ittycheriah, H. Jing, and T. Zhang. Named Entity Recognition through Classifier Combination. In *Proceedings of CoNLL-2003*, pages 168–171. Edmonton, Canada, 2003.
- [Fokkens *et al.* 13a]
 A. Fokkens, M. Van Erp, M. Postma, T. Pedersen, P. Vossen, and N. Freire. Offspring from reproduction problems: What replication failure teaches us. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, volume 1, pages 1691–1701, 2013.
- [Fokkens *et al.* 13b]
 A. Fokkens, M. Van Erp, P. Vossen, S. Tonelli, W. R. Van Hage, B. SynerScope, L. Serafini, R. Sprugnoli, and J. Hoeksema. GAF: A grounded annotation framework for events. In *NAACL HLT*, volume 2013, page 11. Citeseer, 2013.
- [Fokoue *et al.* 06]
 A. Fokoue, A. Kershenbaum, L. Ma, E. Schonberg, and K. Srinivas. The summary abox: Cutting ontologies down to size. In *Proc. of the Int. Semantic Web Conf. (ISWC2006)*, pages 136–145, 2006.

- [Fong *et al.* 01]
 S. Fong, A. Sun, and K. K. Wong. Price Watcher Agent for E-Commerce. In *Proceedings of the second Asia-Pacific Conference on Intelligent Agent Technology (IAT-2001)*, pages 294–299, Maebashi Terra, Maebashi City, Japan, 2001.
- [Foo00]
 Football unlimited web page. <http://www.footballunlimited.com>, September 2000.
- [Fornaciari *et al.* 13]
 T. Fornaciari, F. Celli, and M. Poesio. The effect of personality type on deceptive communication style. In *Intelligence and Security Informatics Conference (EISIC), 2013 European*, pages 1–6, Aug 2013.
- [Fornasari *et al.* 05]
 F. Fornasari, A. Tommasi, C. Zavattari, R. Gagliardi, T. Declerck, and M. Nannipieri. Xbrl web-based business intelligence services. In P. Cunningham and M. Cunningham, editors, *Innovation and the Knowledge Economy: Issues, Applications, Case Studies. Proceedings of eChallenge 2005*. IOS Press, 2005.
- [forNational Statistics 10]
 O. for National Statistics. National Population Projections, 2010-based projections: Principal projection - GB population in age groups. <http://www.ons.gov.uk/ons/rel/npp/national-population-projections/2010-based-projections/rft-table-a2-1-principal-projection---uk-population-in-age-groups.xls>, 2010.
- [Forney 73]
 G. D. Forney, Jr. The Viterbi algorithm. *Proceedings of the IEEE*, 61:268–278, 1973.
- [Forsyth & Martell 07]
 E. Forsyth and C. Martell. Lexical and discourse analysis of online chat dialog. In *International Conference on Semantic Computing*, pages 19–26. IEEE, 2007.
- [Fort & Sagot 10]
 K. Fort and B. Sagot. Influence of pre-annotation on pos-tagged corpus development. In *Proceedings of the Fourth Linguistic Annotation Workshop*, pages 56–63, 2010.
- [Fort *et al.* 11]
 K. Fort, G. Adda, and K. B. Cohen. Amazon mechanical turk: Gold mine or coal mine? *Computational Linguistics*, 37(2):413–420, 2011.
- [Foster & White 04]
 M. E. Foster and M. White. Techniques for Text Planning with XSLT. In *RDF/RDFS and OWL in Language Technology: 4th Workshop on NLP and XML (NLPXML-2004)*, Barcelona, Spain, 2004.
- [Foster 11a]
 I. Foster. Globus online: Accelerating and democratizing science through cloud-based services. *IEEE Internet Computing*, 15:70–73, 2011.
- [Foster 11b]
 I. Foster. Globus online: Accelerating and democratizing science through cloud-based services. *IEEE Internet Computing*, 15(3):70–73, 2011.
- [Foster *et al.* 01]
 I. Foster, C. Kesselmann, and S. Tuecke. Anatomy of the grid: Enabling scalable virtual organizations. *International Journal of Supercomputing Applications*, 2001.
- [Foster *et al.* 02]
 I. Foster, C. Kesselmann, J. Nick, and S. Tuecke. Grid services for distributed system integration. *Computer*, pages 37–46, June 2002.

- [Foster *et al.* 11]
 J. Foster, Ö. Çetinoglu, J. Wagner, J. Le Roux, S. Hogan, J. Nivre, D. Hogan, J. Van Genabith, et al. #hardtoparse: POS Tagging and Parsing the Twittersverse. In *Proceedings of the AAAI Workshop On Analyzing Microtext*, pages 20–25, 2011.
- [Fowler & Scott 97]
 M. Fowler and K. Scott. *UML Distilled*. Addison-Welsey, Reading, MA, 1997.
- [Fowler & Scott 00]
 M. Fowler and K. Scott. *UML Distilled, Second Edition*. Addison-Welsey, Reading, MA, 2000.
- [Fowler 97]
 M. Fowler. *Analysis Patterns: Reusable Object Models*. Addison-Welsey, Reading, MA, 1997.
- [Fowler *et al.* 99]
 H. Fowler, F. Fowler, and J. Pearsall, editors. *The Concise Oxford English Dictionar*. Oxford University Press, 1999.
- [Fox 04]
 B. Fox. Let the PC entertain you. *New Scientist*, page 22, July 2004.
- [Frakes & Baeza-Yates 92]
 W. Frakes and R. Baeza-Yates, editors. *Information retrieval, data structures and algorithms*. Prentice Hall, New York, Englewood Cliffs, N.J., 1992.
- [Francis & Kucera 79]
 W. N. Francis and H. Kucera. Brown corpus manual. *Brown University*, 1979.
- [Francopoulo *et al.* 09]
 G. Francopoulo, N. Bel, M. George, N. Calzolari, M. Monachini, M. Pet, and C. Soria. Multilingual resources for nlp in the lexical markup framework (lmf). *Language Resources and Evaluation*, 43:57–70, 2009. 10.1007/s10579-008-9077-5.
- [Francopoulo *et al.* 16]
 G. Francopoulo, J. Mariani, P. Paroubek, and F. Vernier. Providing and Analyzing NLP Terms for our Community. *Computerm 2016*, page 94, 2016.
- [Frank *et al.* 05]
 E. Frank, M. Hall, G. Holmes, R. Kirkby, B. Pfahringer, and I. Witten. Weka: A machine learning workbench for data mining. *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*, pages 1305–1314, 2005.
- [Frantzi & Ananiadou 95]
 K. Frantzi and S. Ananiadou. Statistical measures for terminological extraction. In *Proc. of 3rd International Conference on Statistical Analysis of Textual Data (JADT '95)*, pages 297–308, 1995.
- [Frantzi & Ananiadou 96a]
 K. Frantzi and S. Ananiadou. Extracting nested collocations. In *Proc. of the 16th International Conference on Computational Linguistics (COLING)*, pages 41–46, Copenhagen, Denmark, 1996.
- [Frantzi & Ananiadou 96b]
 K. Frantzi and S. Ananiadou. A hybrid approach to term recognition. In *Proceedings of NLP+IA 96*, volume 1, pages 93–98, Moncton, Canada, June 1996.
- [Frantzi & Ananiadou 97]
 K. Frantzi and S. Ananiadou. Automatic term recognition using contextual cues. In *Proceedings of 3rd DELOS Workshop*, Zurich, Switzerland, 1997.

- [Frantzi & Ananiadou 99]
K. Frantzi and S. Ananiadou. The C-Value/NC-Value domain independent method for multi-word term extraction. *Journal of Natural Language Processing*, 6(3):145–179, 1999.
- [Frantzi 97]
K. Frantzi. Incorporating context information for the extraction of terms. In *Proceedings of ACL/EACL '97*, pages 501–503, Madrid, Spain, July 1997.
- [Frantzi 98]
K. Frantzi. *Automatic Recognition of Multi-Word Terms*. Unpublished PhD thesis, Manchester Metropolitan University, England, 1998.
- [Frederick P. Brooks 78]
J. Frederick P. Brooks. *The Mythical Man-Month: Essays on Software Engineering*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1978.
- [Frege 92a]
G. Frege. Über sinn und bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 1892.
- [Frege 92b]
G. Frege. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und Philosophische Kritik*, 100:25–50, 1892. (Translated by M. Black under the title “On Sense and Reference”, in P. Geach and M. Black, *Translations from the Philosophical Writings of Gottlob Frege*, Oxford, 1952.)
- [Freitag & McCallum 99]
D. Freitag and A. K. McCallum. Information Extraction with HMMs and Shrinkage. In *Proceedings of Workshop on Machine Learning for Information Extraction*, pages 31–36, 1999.
- [Freire *et al.* 12]
N. Freire, J. Borbinha, and P. Calado. An approach for named entity recognition in poorly structured data. *The Semantic Web: Research and Applications*, pages 718–732, 2012.
- [Freitag & Kushmerick 00]
D. Freitag and N. Kushmerick. Boosted Wrapper Induction. In *Seventeenth National Conference on Artificial Intelligence (AAAI-2000): Twelfth Innovative Applications of Artificial Intelligence Conference (IAAI-2000)*, pages 577–583, 2000.
- [Freitag 97]
D. Freitag. Using grammatical inference to improve precision in information extraction. *ICML-97 Workshop on Automata Induction, Grammatical Inference, and Language Acquisition*, 1997.
- [Freitag 98a]
D. Freitag. Information extraction from html: Application of a general learning approach. *Proceedings of the Fifteenth Conference on Artificial Intelligence AAAI-98*, pages 517–523, 1998.
- [Freitag 98b]
D. Freitag. Information Extraction from HTML: Application of a general Machine Learning approach. In *Proceedings of Fifteenth National Conference on Artificial Intelligence (AAAI-98)*, 1998.
- [Freitag 98c]
D. Freitag. *Machine Learning for Information Extraction in Informal Domains*. Unpublished PhD thesis, Carnegie Mellon University, 1998.
- [Freitag 98d]
D. Freitag. Multistrategy Learning for Information Extraction. In *Proceedings of the Fifteenth International Conference on Machine Learning*, pages 161–169, 1998.

- [Freitag 00]
D. Freitag. Machine Learning for Information Extraction in Informal Domains. *Machine Learning*, 39(2/3):169–202, 2000.
- [Freitas *et al.* 10]
J. Freitas, A. Calado, D. Braga, P. Silva, and M. S. Dias. Crowd-sourcing platform for large-scale speech data collection. In *VI Jornadas en Tecnología del Habla and II Iberian SLTech Workshop*, 2010.
- [Friburger & Maurel 02]
N. Friburger and D. Maurel. Textual similarity based on proper names. In *Proceedings of the Workshop on Mathematical/Formal Methods in Information Retrieval (SIGIR 2002 MF/IR 2002)*, Tampere, Finland, Aug 2002.
- [Friedman *et al.* 78a]
J. Friedman, D. Moran, and D. Warren. Dynamic interpretations: Semantics in a process version of montague grammar. Paper presented to the sloan foundation workshop on formal semantics, Stanford, 1978.
- [Friedman *et al.* 78b]
J. Friedman, D. Moran, and D. Warren. Evaluating english sentences in a logical model: A process version of montague grammar. Paper presented to the 7th international conference on computational linguistics, Bergen, 1978.
- [Frietsch *et al.* 16]
R. Frietsch, P. Neuhäusler, O. Rothengatter, and K. Jonkers. Societal grand challenges from a technological perspective: Methods and identification of classes of the international patent classification ipc. Technical report, Fraunhofer ISI Discussion Papers Innovation Systems and Policy Analysis, 2016.
- [Frijters *et al.* 10]
R. Frijters, M. van Vugt, R. Smeets, R. van Schaik, J. de Vlieg, and W. Alkema. Literature mining for the discovery of hidden connections between drugs, genes and diseases. *PLoS Computational Biology*, 6(9):e1000943, 09 2010.
- [Frisch 86]
A. Frisch. Knowledge retrieval as specialized inference. Phd thesis, department of computer science, University of Rochester, Rochester, NY, 1986.
- [Fröhlich & Werner 95]
M. Fröhlich and M. Werner. Demonstration of the Graph Visualization System daVinci. In *Proceedings of DIMACS Workshop on Graph Drawing '94, LNCS 894*. Springer-Verlag, 1995.
- [Fromreide *et al.* 14]
H. Fromreide, D. Hovy, and A. Søgaard. Crowdsourcing and annotating NER for Twitter #drift. In *Proceedings of LREC*, pages 2544–2547. European Language Resources Association, 2014.
- [Fu & Booth 75a]
K. Fu and T. Booth. Grammatical inference: Introduction and survey – part i. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-5(1), 1975.
- [Fu & Booth 75b]
K. Fu and T. Booth. Grammatical inference: Introduction and survey – part ii. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-5(4), 1975.
- [Fu 74]
K. Fu. *Syntactic Methods in Pattern Recognition*. Academic Press, New York, 1974.

- [Fuchs & Schwitter 96]
N. Fuchs and R. Schwitter. Attempto Controlled English (ACE). In *CLAW96: Proceedings of the First International Workshop on Controlled Language Applications*, Leuven, Belgium, 1996.
- [Fuchs *et al.* 99]
N. E. Fuchs, U. Schwertel, and R. Schwitter. Attempto controlled english - not just another logic specification language. In P. Flener, editor, *Logic-Based Program Synthesis and Transformation*, number 1559 in Lecture Notes in Computer Science, Manchester, UK, June 1999. Eighth International Workshop LOPSTR'98, Springer.
- [Fuchs *et al.* 05a]
N. E. Fuchs, S. Hoefler, K. Kaljurand, G. Schneider, and U. Schwertel. Extended Discourse Representation Structures in Attempto Controlled English. Technical Report ifi-2005.08, Department of Informatics, University of Zurich, Zurich, Switzerland, 2005.
- [Fuchs *et al.* 05b]
N. E. Fuchs, K. Kaljurand, T. Kuhn, and G. Schneider. REWERSE review meeting 2005: Demonstrations of the Attempto Parsing Engine APE. Technical report, Department of Informatics and Institute of Computational Linguistics, University of Zurich, May 2005.
- [Fuchs *et al.* 06]
N. E. Fuchs, K. Kaljurand, T. Kuhn, G. Schneider, L. Royer, and M. Schröder. Attempto Controlled English and the semantic web. Deliverable I2D7, REWERSE Project, April 2006.
- [Fuchs *et al.* 08]
N. E. Fuchs, K. Kaljurand, and T. Kuhn. Attempto Controlled English for Knowledge Representation. In C. Baroglio, P. A. Bonatti, J. Małuszyński, M. Marchiori, A. Polleres, and S. Schaffert, editors, *Reasoning Web, Fourth International Summer School 2008*, number 5224 in Lecture Notes in Computer Science, pages 104–124. Springer, 2008.
- [Fuhr 99]
N. Fuhr. Resource Discovery in Distributed Digital Libraries. In *Proceedings of Digital Libraries '99: Advanced Methods and Technologies*, St. Petersburg State University, 1999.
- [Fuhr 00]
N. Fuhr. Information Retrieval in Digital Libraries: Dealing with Structure (invited talk). In *Proceedings of the DELOS Workshop: Information Seeking, Searching and Querying in Digital Libraries*, 2000.
- [Fujino & Isozaki 07]
A. Fujino and H. Isozaki. Multi-label Patent Classification at NTT Communication Science Laboratories. In *Proceedings of the Sixth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access*, pages 381–384, Tokyo, Japan, 2007.
- [Fuller *et al.* 04]
S. Fuller, D. Revere, P. Bugni, and G. Martin. A knowledgebase system to enhance scientific discovery: Telemakus. *Biomedical Digital Libraries*, 1(1):2, 2004.
- [Fung & Lo 98]
P. Fung and Y. Lo. An IR approach for translating new words from unparallel, comparable texts. In *Proc. of 36th Annual ACL Conference*, pages 414–420, Montreal, Canada, 1998.
- [Funk & Bontcheva 10a]
A. Funk and K. Bontcheva. Ontology-based categorization of web services with machine learning. In *Proceedings of the seventh international conference on Language Resources and Evaluation (LREC) [LRE10]*.

- [Funk & Bontcheva 10b]
A. Funk and K. Bontcheva. Effective development with gate and reusable code for semantically analysing heterogeneous documents. In *New Challenges for NLP Frameworks workshop at LREC*, Valletta, Malta, May 2010.
- [Funk & Bontcheva 10c]
A. Funk and K. Bontcheva. Ontology-based categorization of web services with machine learning. In *7th Language Resources and Evaluation Conference (LREC)*, Valletta, Malta, May 2010.
- [Funk *et al.* 06]
A. Funk, B. Davis, V. Tablan, K. Bontcheva, and H. Cunningham. Controlled language IE components version 2. Deliverable D2.2.2, SEKT, 2006.
- [Funk *et al.* 07a]
A. Funk, D. Maynard, H. Saggion, and K. Bontcheva. Ontological integration of information extracted from multiple sources. In *Multi-source Multilingual Information Extraction and Summarization (MMIES) workshop at Recent Advances in Natural Language Processing (RANLP07)*, pages 9–15, Borovets, Bulgaria, September 2007.
- [Funk *et al.* 07b]
A. Funk, V. Tablan, K. Bontcheva, H. Cunningham, B. Davis, and S. Handschuh. CLOnE: Controlled Language for Ontology Editing. In *Proceedings of the 6th International Semantic Web Conference (ISWC 2007)*, Busan, Korea, November 2007.
- [Funk *et al.* 08]
A. Funk, Y. Li, H. Saggion, K. Bontcheva, and C. Leibold. Opinion analysis for business intelligence applications. In *First Int. Workshop on Ontology-Supported Business Intelligence*, Karlsruhe, October 2008. ACM.
- [Funk *et al.* 10]
A. Funk, I. Roberts, and W. Peters. Implementing a variety of linguistic annotations through a common web-service interface. In *Language Resource and Language Technology Standards workshop at LREC*, Valletta, Malta, May 2010.
- [Furnas *et al.* 87]
G. W. Furnas, T. K. Landauer, L. M. Gomez, and S. T. Dumais. The vocabulary problem in human-system communication. *Commun. ACM*, 30:964–971, November 1987.
- [Furuta & Urbina 02]
R. Furuta and E. Urbina. On the characteristics of scholarly annotations. In *Proceedings of the 13th ACM Conference on Hypertext and Hypermedia*, pages 78–79, College Park, Maryland, USA, 2002. ACM, New York, NY, USA.
- [Fusaro *et al.* 11]
V. A. Fusaro, P. Patil, E. Gafni, D. P. Wall, and P. J. Tonellato. Biomedical cloud computing with amazon web services. *PLoS Comput Biol*, 7(8):e1002147, 08 2011.
- [G.A. *et al.* 93]
M. G.A., R. Beckwith, C. Fellbaum, D. Gross, and K. Miller. Introduction to WordNet: On-line. Distributed with the WordNet Software., 1993.
- [Gaggi & Celentano 02]
O. Gaggi and A. Celentano. Modelling Synchronized Hypermedia Presentations. Technical report, Department of Computer Science, Ca' Foscari University, Venezia, Italy, 2002.
- [Gahleitner *et al.* 05]
E. Gahleitner, W. Behrendt, J. Palkoska, and E. Weippl. On cooperatively creating dynamic ontologies. In *HYPertext '05: Proceedings of the sixteenth ACM conference on Hypertext and hypermedia*, pages 208–210, New York, NY, USA, 2005. ACM Press.

- [Gaizauskas & Humphreys 96a]
 R. Gaizauskas and K. Humphreys. Quantitive Evaluation of Coreference Algorithms in an Information Extraction System. In *DAARC96 - Discourse Anaphora and Anaphor Resolution Colloquium*. Lancaster University, 1996.
- [Gaizauskas & Humphreys 96b]
 R. Gaizauskas and K. Humphreys. Using verb semantic role information to extend partial parses via a co-reference mechanism. In J. Carroll, editor, *Proceedings of the Workshop on Robust Parsing*, pages 103–113, Prague, Czech Republic, August 1996. European Summer School in Language, Logic and Information.
- [Gaizauskas & Humphreys 96c]
 R. Gaizauskas and K. Humphreys. XI: A simple prolog-based language for cross-classification and inheritance. In *Proceedings of the 7th International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA96)*, pages 86–95, Sozopol, Bulgaria, 1996.
- [Gaizauskas & Humphreys 96d]
 R. Gaizauskas and K. Humphreys. XI: A Simple Prolog-based Language for Cross-Classification and Inhetotance. In *Proceedings of the 7th International Conference in Artificial Intelligence: Methodology, Systems, Applications*, pages 86–95, Sozopol, Bulgaria, 1996.
- [Gaizauskas & Humphreys 97a]
 R. Gaizauskas and K. Humphreys. Quantitative Evaluation of Coreference Algorithms in an Information Extraction System. Technical report CS-97-19, Department of Computer Science, University of Sheffield, 1997.
- [Gaizauskas & Humphreys 97b]
 R. Gaizauskas and K. Humphreys. Quantitive Evaluation of Coreference Algorithms in an Information Extraction System. In S. Botley and T. McEnery, editors, *Discourse Anaphora and Anaphor Resolution*. University College London Press, 1997.
- [Gaizauskas & Humphreys 97c]
 R. Gaizauskas and K. Humphreys. Using a semantic network for information extraction. Technical Report CS-97-03, Department of Computer Science, University of Sheffield, 1997.
- [Gaizauskas & Robertson 97]
 R. Gaizauskas and A. Robertson. Coupling information retrieval and information extraction: A new text technology for gathering information from the web. In *Proceedings of RIAO'97*, Montreal, 1997.
- [Gaizauskas & Rodgers 97]
 R. Gaizauskas and P. Rodgers. NL Module Evaluation in GATE. In *Proceedings of the SALT club workshop on Evaluation in Speech and Language Technology*, 1997.
- [Gaizauskas & Wilks 98a]
 R. Gaizauskas and Y. Wilks. Information Extraction: Beyond Document Retrieval. *Journal of Documentation*, 54(1):70–105, 1998.
- [Gaizauskas & Wilks 98b]
 R. Gaizauskas and Y. Wilks. Information extraction: Beyond document retrieval. *Journal of Documentation*, 54(1), 1998.
- [Gaizauskas 92a]
 R. Gaizauskas. *Deriving Answers to Logical Queries by Answer Composition*. Unpublished PhD thesis, University of Sussex, 1992.
- [Gaizauskas 92b]
 R. Gaizauskas. Deriving answers to logical queries via answer composition. In G. Wiggins, C. Mellish, and T. Duncan, editors, *3rd U.K. Annual Conference on Logic Programming, Edinburgh 1991*, pages 112–134, London, 1992. Springer-Verlag.

- [Gaizauskas 95a]
R. Gaizauskas. Investigations into the grammar underlying the Penn Treebank II. Research Memorandum CS-95-25, Department of Computer Science, University of Sheffield, 1995.
- [Gaizauskas 95b]
R. Gaizauskas. XI: A Knowledge Representation Language Based on Cross-Classification and Inheritance. Technical Report CS-95-24, Department of Computer Science, University of Sheffield, 1995.
- [Gaizauskas 97]
R. Gaizauskas. A Review of *Evaluating Natural Language Processing Systems*, Sparck-Jones and Galliers, 1996. *Journal of Natural Language Engineering*, 1997.
- [Gaizauskas *et al.* 92a]
R. Gaizauskas, R. Evans, L. J. Cahill, J. Richardson, and J. Walker. POETIC: A system for gathering and disseminating traffic information. In *Proceedings of the International Conference on Artificial Intelligence Applications in Transportation Engineering*, pages 79–98. Institute of Transportation Studies, University of California, Irvine, 1992.
- [Gaizauskas *et al.* 92b]
R. Gaizauskas, R. Evans, L. J. Cahill, J. Richardson, and J. Walker. Poetic: A system for gathering and disseminating traffic information. In S. G. Ritchie and C. T. Hendrickson, editors, *Conference Preprints of the International Conference on Artificial Intelligence Applications in Transportation Engineering*, pages 79–98. San Buenaventura, California, 1992.
- [Gaizauskas *et al.* 93]
R. Gaizauskas, L. Cahill, and R. Evans. Description of the Sussex system used for MUC-5. In *Proceedings of the Fifth Message Understanding Conference (MUC-5)*. ARPA, Morgan Kaufmann, California, 1993.
- [Gaizauskas *et al.* 95]
R. Gaizauskas, T. Wakao, K. Humphreys, H. Cunningham, and Y. Wilks. Description of the LaSIE system as used for MUC-6. In *Proceedings of the Sixth Message Understanding Conference (MUC-6), 6–8 November 1995*, pages 207–220. Morgan Kaufmann, California, 1995.
- [Gaizauskas *et al.* 96a]
R. Gaizauskas, P. Rodgers, H. Cunningham, and K. Humphreys. GATE User Guide. <http://www.dcs.shef.ac.uk/nlp/gate>, 1996.
- [Gaizauskas *et al.* 96b]
R. Gaizauskas, T. Wakao, K. Humphreys, H. Cunningham, and Y. Wilks. Description of the LaSIE system as used for MUC-6. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, pages 207 – 220, San Francisco, California, 1996. Morgan Kaufmann, California.
- [Gaizauskas *et al.* 96c]
R. Gaizauskas, H. Cunningham, Y. Wilks, P. Rodgers, and K. Humphreys. GATE – an Environment to Support Research and Development in Natural Language Engineering. In *Proceedings of the 8th IEEE International Conference on Tools with Artificial Intelligence (ICTAI-96)*, Toulouse, France, October 1996. <ftp://ftp.dcs.shef.ac.uk/home/robertg/ictai96.ps>.
- [Gaizauskas *et al.* 97a]
R. Gaizauskas, K. Humphreys, S. Azzam, and Y. Wilks. Concepticons *vs.* lexicons: An architecture for multilingual information extraction. In M. Pazienza, editor, *Proceedings of the Summer School on Information Extraction (SCIE-97)*, LNCS/LNAI, pages 28–43. Springer-Verlag, 1997.
- [Gaizauskas *et al.* 97b]
R. Gaizauskas, K. Humphreys, S. Azzam, and Y. Wilks. Concepticons *vs.* Lexicons: An Architecture

for Multilingual Information Extraction. In M. Pazienza, editor, *Information Extraction. A multidisciplinary Approach to an Emerging Information Technology. Lectures Notes in Artificial Intelligence*, volume 1299, pages 28–43. Springer, 1997.

[Gaizauskas *et al.* 98]

R. Gaizauskas, M. Hepple, and C. Huyck. A scheme for comparative evaluation of diverse parsing systems. In *Proceedings of the 1st International Conference on Language Resources and Evaluation (LREC'98)*, pages 143–149, Granada, Spain, 1998.

[Gaizauskas *et al.* 03a]

R. Gaizauskas, L. Burnard, P. Clough, and S. Piao. Using the XARA XML-Aware corpus query tool to investigate the METER Corpus. In *In Proceedings of the Corpus Linguistics 2003 Conference*, pages 227–236, Lancaster, UK, 2003.

[Gaizauskas *et al.* 03b]

R. Gaizauskas, M. A. Greenwood, M. Hepple, I. Roberts, H. Saggion, and M. Sargaison. The University of Sheffield's TREC 2003 Q&A Experiments. In *In Proceedings of the 12th Text REtrieval Conference*, 2003.

[Gaizauskas *et al.* 04]

R. Gaizauskas, M. A. Greenwood, M. Hepple, I. Roberts, H. Saggion, and M. Sargaison. The University of Sheffield's TREC 2004 Q&A Experiments. In *In Proceedings of the 13th Text REtrieval Conference*, 2004.

[Gaizauskas *et al.* 05a]

R. Gaizauskas, M. A. Greenwood, M. Hepple, H. Harkema, H. Saggion, and A. Sanka. The University of Sheffield's TREC 2005 Q&A Experiments. In *In Proceedings of the 11th Text REtrieval Conference*, 2005.

[Gaizauskas *et al.* 05b]

R. Gaizauskas, M. Hepple, H. Saggion, M. A. Greenwood, and K. Humphreys. SUPPLE: A practical parser for natural language engineering applications. In *Proceedings of the Ninth International Workshop on Parsing Technology*, pages 200–201. Association for Computational Linguistics, 2005.

[Gaizauskas *et al.* 12]

R. Gaizauskas, E. Barker, C. Chang, L. Derczynski, M. Phiri, and C. Peng. Applying ISO-Space to Healthcare Facility Design Evaluation Reports. In *Proceedings of the Joint ISA-7, SRSL-3 and I2MRT Workshop on Semantic Annotation and the Integration and Interoperability of Multimodal Resources and Tools*, pages 31–38, 2012.

[Gale *et al.* 91]

W. Gale, K. Church, P. Hanks, and D. Hindle. Using statistics in lexical analysis. In U. Zernik, editor, *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*, pages 115–164. Lawrence Erlbaum Associates, Hillsdale, N.J., 1991.

[Gale *et al.* 92a]

W. Gale, K. Church, and D. Yarowsky. One sense per discourse. In *Proceedings of the DARPA Speech and Natural Language Workshop*, pages 233–237, Harriman, New York, February 1992.

[Gale *et al.* 92b]

W. A. Gale, C. K. W., and Y. D. One sense per discourse. In *Proceedings of the DARPA Speech and Natural Language Workshop*, Harriman, New York, February 1992.

[Gale *et al.* 92c]

W. Gale, K. Church, and D. Yarowsky. Work on statistical methods for word sense disambiguation. In *Probabilistic Approaches to Natural Language: Papers from the 1992 Fall Symposium*, pages 54–60, Cambridge, Massachusetts, 1992. AAAI Press.

- [Gale *et al.* 92d]
W. A. Gale, K. W. Church, and D. Yarowsky. A method for disambiguating word senses in a large corpus. *Computers and the Humanities*, 26:415–439, 1992.
- [Gales 92]
M. J. F. Gales. An improved approach to the hidden Markov model decomposition of speech and noise. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 1, pages 233–236, San Francisco, 1992.
- [Gallagher 78]
R. Gallagher. Variations on a theme by huffman. *IEEE Transactions on Information Theory*, IT-24(6):668–674, 1978.
- [Gallaire & Minker 77]
H. Gallaire and J. Minker. *Logic and Databases*. Plenum Press, New York, 1977.
- [Gallaire *et al.* 84]
H. Gallaire, J. Minker, and J.-M. Nicolas. Logic and databases: A deductive approach. *Computing Surveys*, 16(2):153–185, 1984.
- [Gallant 91]
S. Gallant. A practical approach for representing contexts and for performing word-sense disambiguation using neural networks. 3:293–309, 1991.
- [Gallier 86]
J. Gallier. *Logic for Computer Science*. Harper and Row, New York, 1986.
- [Gallin 75]
D. Gallin. Intensional and higher order modal logic. 1975.
- [Galperin 05]
M. Y. Galperin. The molecular biology database collection: 2005 update. *Nucleic Acids Research*, 33:D5–D24, 2005.
- [Gamallo *et al.* 06]
P. Gamallo, M. Gonzalez, A. Agustini, G. Lopes, and V. de Lima. Mapping syntactic dependencies onto semantic relations. In *Proc. of the ECAI Workshop on Machine Learning and Natural Language Processing for Ontology Engineering*, 2006.
- [Gambäck & Olsson 00]
B. Gambäck and F. Olsson. Experiences of Language Engineering Algorithm Reuse. In *Second International Conference on Language Resources and Evaluation (LREC)*, pages 155–160, Athens, Greece, 2000.
- [Gamma *et al.* 95]
E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns – Elements of Reusable Object-Oriented Software*. Addison-Wesley, Reading, MA, 1995.
- [Ganesan *et al.* 10]
K. Ganesan, C. Zhai, and J. Han. Opinosis: A graph based approach to abstractive summarization of highly redundant opinions. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING '10)*, 2010.
- [Ganesan *et al.* 12a]
K. Ganesan, C. Zhai, and E. Viegas. Micropinion generation: an unsupervised approach to generating ultra-concise summaries of opinions. In *Proceedings of the 21st World Wide Web Conference*, pages 869–878, 2012.

- [Ganesan *et al.* 12b]
K. Ganesan, C. Zhai, and E. Viegas. Micropinion generation: an unsupervised approach to generating ultra-concise summaries of opinions. In *Proceedings of the 21st International Conference on World Wide Web*, WWW '12, pages 869–878, 2012.
- [Gangemi 05]
A. Gangemi. Ontology design patterns for semantic web content. In *The Semantic Web–ISWC 2005*, pages 262–276. Springer, 2005.
- [Garcia 05]
M. A. M. A. Garcia, F.S. Working with jbpm: From workflow management to business intelligence - white paper. Technical report, BT Intelligent Systems Research Centre, 2005.
- [Garcia-Castro *et al.* 05]
R. Garcia-Castro, D. Maynard, H. Wache, D. Foxvog, and R. G. Cabero. Specification of a methodology, general criteria and benchmark suites for benchmarking ontology tools. Technical Report D2.1.4, KnowledgeWeb Deliverable, 2005.
- [García-Castro *et al.* 07]
R. García-Castro, S. David, and J. Prieto-González. D1.2.2.1.2 benchmarking the interoperability of ontology development tools using owl as interchange language. Technical report, Knowledge Web, September 2007.
- [Gardiner & Java 90]
J. Gardiner and R. Java. Recollective experience in word and nonword recognition. *Memory and Cognition*, 18:23–30, 1990.
- [Gardner 05]
S. P. Gardner. Ontologies and semantic data integration. *Drug and Discovery Today*, 10(14):1001–1007, July 2005.
- [Garey & Johnson 79]
M. Garey and D. Johnson. *Computers and Intractability*. W.H. Freeman and Co., San Francisco, 1979.
- [Garlan & Perry 95]
D. Garlan and D. Perry. Introduction to the Special Issue on Software Architecture. *IEEE Transactions on Software Engineering*, 21(4), April 1995.
- [Garlan *et al.* 95]
D. Garlan, R. Allen, and J. Ockerbloom. Architectural Mismatch or Why it's hard to build systems out of existing parts. In *Proceedings of the Seventeenth International Conference on Software Engineering*, Seattle WA, April 1995.
- [Garside *et al.* 87]
R. Garside, G. Leech, and G. Sampson, editors. *The Computational Analysis of English*. Longman, London, 1987.
- [Gattani *et al.* 13]
A. Gattani, D. S. Lamba, N. Garera, M. Tiwari, X. Chai, S. Das, S. Subramaniam, A. Rajaraman, V. Harinarayan, and A. Doan. Entity extraction, linking, classification, and tagging for social media: a wikipedia-based approach. *Proceedings of the VLDB Endowment*, 6(11):1126–1137, 2013.
- [Gauvain *et al.* 94]
J. L. Gauvain, L. F. Lamel, G. Adda, and M. Adda-Decker. The LIMSI continuous speech dictation system: Evaluation on the ARPA Wall Street Journal task. volume 1, pages 557–560, 1994.

- [Gawron *et al.* 82]
 J. Gawron, J. King, J. Lamping, E. Loebner, E. Paulson, G. Pullum, I. Sag, and T. Wasow. Processing english with a generalized phrase structure grammar. Computer science laboratory technical note csl-82-5, Hewlett-Packard Company, Palo Alto CA, 1982.
- [Gazdar & Mellish 89a]
 G. Gazdar and C. Mellish. *Natural Language Processing in Prolog*. Addison-Wesley, Reading, MA, 1989.
- [Gazdar & Mellish 89b]
 G. Gazdar and C. Mellish. *Natural Language Processing in Prolog*. Addison-Wesley, Reading, MA, 1989.
- [Gazdar & Mellish 89c]
 G. Gazdar and C. Mellish. *Natural Language Processing in Prolog: An Introduction to Computational Linguistics*. Addison-Wesley Publishing Company, 1989.
- [Gazdar 96]
 G. Gazdar. Paradigm merger in natural language processing. In R. Milner and I. Wand, editors, *Computing Tomorrow: Future Research Directions in Computer Science*, pages 88–109. Cambridge University Press, 1996.
- [Gazdar *et al.* 82]
 G. Gazdar, G. Pullum, and I. Sag. Auxiliaries and related phenomena in a restrictive theory of grammar. *Language*, 58(3), 1982.
- [Gazdar *et al.* 85]
 G. Gazdar, E. Klein, G. Pullum, and I. Sag. *Generalized Phrase Structure Grammar*. Basil Blackwell, Oxford, 1985.
- [Geelan 09]
 J. Geelan. Twenty-one experts define cloud computing. *Cloud Computing Journal*, pages 1–5, 2009.
- [Geldof & Van deVelde 97a]
 S. Geldof and W. Van de Velde. An architecture for template based (hyper)text generation. In *Proceedings of the 6th European Workshop on Natural Language Generation (EWNLG'97)*, pages 28 – 37, 1997.
- [Geldof & Van deVelde 97b]
 S. Geldof and W. Van de Velde. Context-sensitive hypertext generation. In *Working Notes of Natural Language Processing for the World Wide Web*, pages 54 – 61, March 1997. AAAI-97 Spring Symposium Series.
- [Gentile *et al.* 11]
 A. L. Gentile, V. Lanfranchi, S. Mazumdar, and F. Ciravegna. Extracting semantic user networks from informal communication exchanges. In *Proceedings of the 10th International Conference on the Semantic Web, ISWC'11*, pages 209–224. Springer-Verlag, 2011.
- [Georgiev *et al.* 11]
 G. Georgiev, K. Pentchev, A. Avramov, T. Primov, and V. Momtchev. Scalable interlinking of biomedical entities and scientific literature in linked life data. In *Proceedings of the second workshop on Collaborative Annotation of a Large Biomedical Corpus (CALBC II), 16-18 March 2011*, pages 59–61, Cambridge, UK, 2011.
- [Gerber *et al.* 11]
 A. Gerber, L. Gao, and J. Hunter. A Scoping Study of (Who, What, When, Where) Semantic Tagging Services. Technical report, eResearch Lab, The University of Queensland, Queensland, Australia, 2011.

- [Geser & (eds.) 04]
G. Geser and J. P. (eds.). An expedition report. In *DigiCULT - The Future Digital heritage Space*, volume Thematic Issue 7. DigiCULT Consortium, 2004.
- [Geser 04]
G. Geser. Introduction and overview. In *DigiCULT - Virtual Communities and Collaboration in the Heritage Sector*, volume Thematic Issue 5. DigiCULT Consortium, 2004.
- [Gevrey & Rüger 02]
J. Gevrey and S. Rüger. Link-based approaches for text retrieval. In *Proceedings of TREC-10*, pages 279–285, Gaithersburg, MD, 2002.
- [Ghose *et al.* 07]
A. Ghose, P. G. Ipeirotis, and A. Sundararajan. Opinion mining using econometrics: A case study on reputation systems. In *Proceedings of the Association for Computational Linguistics*. The Association for Computational Linguistics, 2007.
- [Gibbins *et al.* 08]
N. Gibbins, T. R. Payne, and H. H. Wang. Tao suite best practice documentation. Technical Report D5.4, TAO Project Deliverable, 2008. <http://www.tao-project.eu/resources/publicdeliverables/d5-4-final.pdf>.
- [Gillard 96]
P. Gillard. Personal communication, 1996. Longmans Group.
- [Gillick & Liu 10]
D. Gillick and Y. Liu. Non-expert evaluation of summarization systems is risky. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, pages 148–151, 2010.
- [Gimenez & Marquez 03]
J. Gimenez and L. Marquez. Fast and Accurate Part-of-Speech Tagging: The SVM Approach Revisited. In *Proceedings of the International Conference RANLP-2003 (Recent Advances in Natural Language Processing)*, pages 158–165. John Benjamins Publishers, 2003.
- [Gimpel *et al.* 11]
K. Gimpel, N. Schneider, B. O’Connor, D. Das, D. Mills, J. Eisenstein, M. Heilman, D. Yogatama, J. Flanigan, and N. Smith. Part-of-speech tagging for twitter: annotation, features, and experiments. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 42–47. ACL, 2011.
- [Gindl *et al.* 10a]
S. Gindl, A. Weichselbraun, and A. Scharl. Cross-domain contextualisation of sentiment lexicons. In *Proceedings of 19th European Conference on Artificial Intelligence (ECAI-2010)*, pages 771–776, 2010.
- [Gindl *et al.* 10b]
S. Gindl, A. Scharl, and A. Weichselbraun. Generic high-throughput methods for multilingual sentiment detection. In *Digital Ecosystems and Technologies (DEST), 2010 4th IEEE International Conference on*, pages 239–244. IEEE, 2010.

- [Girju *et al.* 03]
 R. Girju, A. Badulescu, and D. Moldovan. Learning Semantic Constraints for the Automatic Discovery of Part-Whole Relations. In M. Hearst and M. Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 80–87, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [Gish 90]
 H. Gish. A probabilistic approach to the understanding and training of neural network classifiers. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1361–1364, Albuquerque, 1990.
- [Giunchiglia *et al.* 09a]
 F. Giunchiglia, U. Kharkevich, and I. Zaihrayeu. Concept search. In *The Semantic Web: Research and Applications, 6th European Semantic Web Conference*, volume 5554 of *Lecture Notes in Computer Science*, pages 429–444. Springer, 2009.
- [Giunchiglia *et al.* 09b]
 F. Giunchiglia, U. Kharkevich, and I. Zaihrayeu. Concept search. In L. Aroyo, P. Traverso, F. Ciravegna, P. Cimiano, T. Heath, E. Hyvönen, R. Mizoguchi, E. Oren, M. Sabou, and E. Simperl, editors, *The Semantic Web: Research and Applications*, volume 5554 of *Lecture Notes in Computer Science*, pages 429–444. Springer Berlin, Heidelberg, 2009.
- [Glaser & Schütze 12]
 A. Glaser and H. Schütze. Automatic generation of short informative sentiment summaries. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 276–285, Avignon, France, April 2012.
- [Glass *et al.* 04]
 J. Glass, T. Hazen, L. Hetherington, and C. Wang. Analysis and processing of lecture audio data: Preliminary investigations. In *Proc., HLT-NAACL 2004 Workshop on Interdisciplinary Approaches to Speech Indexing and Retrieval*, Boston, 2004.
- [Gledhill 85]
 C. Gledhill. Genre. In P. Cook, editor, *The Cinema Book*. British Film Institute, London, 1985.
- [Gligorov *et al.* 07]
 R. Gligorov, W. ten Kate, Z. Aleksovski, and F. van Harmelen. Using google distance to weight approximate ontology matches. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 767–776, New York, NY, USA, 2007. ACM Press.
- [Gliozzo *et al.* 05]
 A. M. Gliozzo, C. Giuliano, and R. Rinaldi. Instance Filtering for Entity Recognition. *SIGKDD Explorations*, 2005. Special Issue: Text Mining and Natural Language Processing.
- [Glover *et al.* 02]
 E. Glover, K. Tsioutsoulouklis, S. Lawrence, D. Pennock, and G. Flake. Using Web Structure for Classifying and Describing Web Pages. In *Proceedings of the International World Wide Web Conference*, Honolulu, Hawaii, May 7–11 2002.
- [Go *et al.* 09a]
 A. Go, R. Bhayani, , and L. Huang. Twitter sentiment classification using distant supervision. Technical Report CS224N Project Report, Stanford University, 2009.
- [Go *et al.* 09b]
 A. Go, L. Huang, and R. Bhayani. Twitter Sentiment Analysis. Technical report, Stanford University, 2009.

- [Goble & Roure 09]
C. Goble and D. D. Roure. The impact of workflow tools on data-centric research. *Data Intensive Computing: The Fourth Paradigm of Scientific Discovery*, May 2009.
- [Goetz & Suhre 04]
T. Goetz and O. Suhre. Design and implementation of the uima common analysis system. *IBM SYSTEMS JOURNAL*, 43(3):476–489, 2004.
- [Goh *et al.* 01]
K.-S. Goh, E. Chang, and K.-T. Cheng. Support Vector Machine Pairwise Classifiers With Error Reduction for Image Classification. In *Proceedings of the 2001 ACM workshops on Multimedia: multimedia information retrieval*, pages 32–37, Ottawa, Canada, 2001.
- [Gök *et al.* 15]
A. Gök, A. Waterworth, and P. Shapira. Use of web mining in studying innovation. *Scientometrics*, 102(1):653–671, 2015.
- [Gold 67]
E. Gold. Language identification in the limit. *Information and Control*, 10, 1967.
- [Goldfarb & Prescod 98a]
C. Goldfarb and P. Prescod. *The XML Handbook*. Prentice Hall, New York, 1998.
- [Goldfarb & Prescod 98b]
C. Goldfarb and P. Prescod. *The XML Handbook*. Prentice Hall, 1998.
- [Goldfarb 81]
W. D. Goldfarb. The undecidability of the second-order unification problem. *Theoretical Computer Science*, 13:225–230, 1981.
- [Goldfarb 90]
C. F. Goldfarb. *The SGML Handbook*. Oxford University Press, 1990.
- [Goldstein 94]
N. Goldstein. *The Associated Press Stylebook and Libel Manual. Fully Revised and Updated*. ERIC, 1994.
- [Goldstein *et al.* 00]
J. Goldstein, V. O. Mittal, J. G. Carbonell, and M. Kantrowitz. Multi-document summarization by sentence extraction. In *Proceedings of ANLP/NAACL workshop on Automatic Summarization*, Seattle, WA, April 2000.
- [Gomez *et al.* 94]
F. Gomez, R. Hull, and C. Segami. Acquiring knowledge from encyclopaedic texts. In *4th Conference on Applied NLP*, pages 84–90. Stuttgart, October 1994.
- [Gomez-Rodriguez *et al.* 14]
M. Gomez-Rodriguez, K. Gummadi, and B. Schölkopf. Quantifying information overload in social media and its impact on social contagions. In *ICWSM '14: Proceedings of the 8th International AAAI Conference on Weblogs and Social Media*, 2014.
- [Gong & Liu 01]
Y. Gong and X. Liu. Generic Text Summarization Using Relevance Measure and Latent Semantic Analysis. In *Proceedings of SIGIR 2001*, pages 19–25, 2001.
- [Goni *et al.* 97]
J. Goni, J. Gonzalez, and A. Moreno. ARIES: A lexical platform for engineering Spanish processing tools. *Journal of Natural Language Engineering*, 3(4):317–347, 1997.

- [Gooch 12]
P. Gooch. Badrex: In situ expansion and coreference of biomedical abbreviations using dynamic regular expressions. Technical report, City University London, London, 2012.
- [Goodwin *et al.* 08]
J. Goodwin, C. Dolbear, and G. Hart. Geographical linked data: The administrative geography of Great Britain on the Semantic Web. *Transactions in GIS*, 12:19–30, 2008.
- [Gopalakrishnan *et al.* 94]
P. S. Gopalakrishnan, D. Nahamoo, M. Padmanabhan, and M. A. Picheny. A channel-bank-based phone detection strategy. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 2, pages 161–164, Adelaide, 1994.
- [Gopalakrishnan *et al.* 95]
P. S. Gopalakrishnan, L. R. Bahl, and R. L. Mercer. A tree-search strategy for large vocabulary continuous speech recognition. volume 1, pages 572–575, Detroit, 1995.
- [Gordon 79]
M. Gordon. *The Denotational Description of Programming Languages: an Introduction*. Springer, New York, 1979.
- [Gorin *et al.* 91]
A. Gorin, S. Levinson, A. Gertner, and E. Goldman. Adaptive acquisition of language. *Computer Speech and Language*, 5:101–132, 1991.
- [Gorlick & Quilici 94]
M. Gorlick and A. Quilici. Visual Programming-in-the-Large versus Visual Programming-in-the-Small. In *Proceedings VL'94 Tenth Annual IEEE Conference on Visual Languages, St.Louis*. IEEE Computer Society Press, 1994.
- [Gorrell & Bontcheva 14]
G. Gorrell and K. Bontcheva. Classifying twitter favourites: Like, bookmark or thanks? *Journal of the Association of Information Science and Technology*, tba:tba, 2014.
- [Gorrell *et al.* 14]
G. Gorrell, J. Petrak, K. Bontcheva, G. Emerson, and T. Declerck. Multilingual resources and evaluation of knowledge modelling - v2. Technical Report D2.3.2, Trendminer Project Deliverable, 2014.
- [Gorrell *et al.* 15]
G. Gorrell, J. Petrak, and K. Bontcheva. Using @Twitter conventions to improve #lod-based named entity disambiguation. In *The Semantic Web. Latest Advances and New Domains*, pages 171–186. Springer, 2015.
- [Görz *et al.* 96]
G. Görz, M. Kessler, J. Spilker, and H. Weber. Research on Architectures for Integrated Speech/Language Systems in Verbmobil. In *Proceedings of COLING-96*, Copenhagen, 1996.
- [Gotoh & Renals 00]
Y. Gotoh and S. Renals. Information extraction from broadcast news. *Philosophical Transactions of the Royal Society of London, Series A*, 358(1769):1295–1309, 2000. <http://www.dcs.shef.ac.uk/~sjr/pubs/2000/rs00-preprint.html>.
- [Gotoh *et al.* 98]
Y. Gotoh, S. Renals, R. Gaizauskas, G. Williams, and H. Cunningham. Named Entity Tagged Language Models for LVCSR. Technical Report CS-98-05, Department of Computer Science, University of Sheffield, 1998.

- [Göttler 89]
H. Göttler. Graph Grammars: A New Paradigm for Implementing Visual Languages. In *ESEC'89. 2nd European Software Engineering Conference. University of Warwick, Coventry, U.K. LNCS 387*, pages 336–350. Springer-Verlag, 1989.
- [Gouws *et al.* 11]
S. Gouws, D. Metzler, C. Cai, and E. Hovy. Contextual bearing on linguistic variation in social media. In *Proceedings of the Workshop on Languages in Social Media, LSM '11*, pages 20–29, 2011.
- [Gowda & Diday 92]
K. C. Gowda and E. Diday. Symbolic clustering using a new dissimilarity measure. *IEEE Transactions on Systems Management and Cybernetics*, 22:368–378, 1992.
- [Graf *et al.* 10]
E. Graf, I. Frommholz, M. Lalmas, and K. van Rijsbergen. Knowledge modeling in prior art search. In H. Cunningham, A. Hanbury, and S. Rüger, editors, *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference). LNCS volume number: 6107*, Lecture Notes in Computer Science, Vienna, Austria, May 2010. Springer.
- [Gray 84]
P. Gray. *Logic, Algebra, and Databases*. Ellis Horwood, Chichester, 1984.
- [Gray, Jr. & Markel 76]
A. H. Gray, Jr. and J. D. Markel. Distance measures for speech processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 24:380–391, 1976.
- [Grcar *et al.* 07]
M. Grcar, D. Mladenic, M. Grobelnik, B. Fortuna, and J. Brank. Ontology learning implementation. Technical Report D2.2, TAO Project Deliverable, 2007. <http://www.tao-project.eu/resources/publicdeliverables/d2-2.pdf>.
- [Green & Petre 96]
T. G. Green and M. Petre. Usability analysis of visual programming environments: a ‘cognitive dimensions’ framework. *Journal of Visual Languages and Visual Computing*, 7(2):134–174, 1996.
- [Green *et al.* 95]
P. D. Green, M. P. Cooke, and M. D. Crawford. Auditory Scene Analysis and HMM-Recognition of Speech in Noise. pages 401–404, 1995.
- [Greenfield *et al.* 16]
K. Greenfield, R. Caceres, M. Coury, K. Geyer, Y. Gwon, J. Matterer, A. Mensch, C. Sahin, and O. Simek. A reverse approach to named entity extraction and linking in microposts. In *6th workshop on 'Making Sense of Microposts'*, pages 67–69, 2016.
- [Greenhalgh *et al.* 09]
T. Greenhalgh, H. W. Potts, G. Wong, P. Bark, and D. Swinglehurst. Tensions and paradoxes in electronic patient record research: a systematic literature review using the meta-narrative method. *Milbank Quarterly*, 87(4):729–788, Dec 2009.
- [Greenspun 99]
P. Greenspun. *Philip and Alex's Guide to Web Publishing*. Morgan Kaufman, California, 1999.
- [Greenwood & Bontcheva 12]
M. Greenwood and K. Bontcheva. Multi-lingual summarisation of stream media software. Technical Report D4.1.1, TrendMiner Project Deliverable, 2012.
- [Greenwood *et al.* 02]
M. A. Greenwood, I. Roberts, and R. Gaizauskas. The University of Sheffield's TREC 2002 Q&A Experiments. In *In Proceedings of the 11th Text REtrieval Conference*, 2002.

- [Greenwood *et al.* 10]
M. A. Greenwood, H. Cunningham, N. Aswani, I. Roberts, and V. Tablan. GATE Mimir: Philosophy, Development, Deployment and Evaluation. Research Memorandum CS-10-05, Department of Computer Science, University of Sheffield, 2010.
- [Greenwood *et al.* 11]
M. A. Greenwood, V. Tablan, and D. Maynard. GATE Mimir: Answering Questions Google Can't. In *Proceedings of the 10th International Semantic Web Conference (ISWC2011)*, October 2011.
- [Greenwood *et al.* 12]
M. A. Greenwood, N. Aswani, and K. Bontcheva. Reputation Profiling with GATE. In *CLEF 2012 Labs and Workshop Notebook Papers*, 2012.
- [Greenwood *et al.* 16]
M. A. Greenwood, V. Solachidis, O. Papadopoulou, K. Apostoliditis, D. Galanopoulos, D. Tatzoglou, V. Mezaris, B. Eldesouky, N. K. Tran, C. Hube, C. Niederée, J. Petrak, and G. Gorrell. ForgetIT Deliverable D6.4: Contextualisation Framework and Evaluation, February 2016.
- [Grefenstette 94]
G. Grefenstette. *Explorations in Automatic Thesaurus Discovery*. Kluwer Academic Publishers, 1994.
- [Gridinoc *et al.* 08]
L. Gridinoc, M. Sabou, M. D'Aquin, M. Dzbor, and E. Motta. Semantic browsing with PowerMagpie. In *ESWC'08: Proceedings of the 5th European semantic web conference on The semantic web*, pages 802–806. Springer-Verlag, 2008.
- [Grillo *et al.* 03]
G. Grillo, F. Licciulli, S. Liuni, E. Sbisà, and G. Pesole. Patsearch: A program for the detection of patterns and structural motifs in nucleotide sequences. *Nucleic Acids Res*, 31(13):3608–3612, July 2003.
- [Grishman & Kittredge 86]
R. Grishman and R. Kittredge, editors. *Analyzing Language in Restricted Domains: Sublanguage Description and Processing*. Lawrence Erlbaum, Hillsdale, New Jersey, 1986.
- [Grishman & Sterling 93]
R. Grishman and J. Sterling. Description of the Proteus system as used for MUC-5. In *Proceedings of the Fifth Message Understanding Conference (MUC-5)*, pages 181–194. Morgan Kaufmann, California, 1993.
- [Grishman & Sundheim 96a]
R. Grishman and B. Sundheim. Message Understanding Conference - 6: A Brief History. In *Proceedings of the 16th International Conference on Computational Linguistics*, Copenhagen, June 1996. Association for Computational Linguistics, Morristown, NJ, USA.
- [Grishman & Sundheim 96b]
R. Grishman and B. Sundheim. Message understanding conference - 6: A brief history. In *Proceedings of the 16th International Conference on Computational Linguistics*, Copenhagen, Denmark, Jun 1996.
- [Grishman & Sundheim 96c]
R. Grishman and B. Sundheim. Message Understanding Conference-6: A Brief History. In *Proceedings of the 16th International Conference on Computational Linguistics*, Copenhagen, June 1996.
- [Grishman 95a]
R. Grishman. The NYU system for MUC-6 or where's the syntax. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, pages 167–176. Morgan Kaufmann, California, 1995.

- [Grishman 95b]
R. Grishman. TIPSTER Architecture Design Document Version 1.52 (Tinman Architecture). Technical report, Department of Computer Science, New York University, 1995. <http://www.cs.nyu.edu/~tipster>.
- [Grishman 95c]
R. Grishman. TIPSTER Architecture Design Document Version 2.0 (Tinman Architecture). Technical report, Department of Computer Science, New York University, 1995. <http://www.cs.nyu.edu/~tipster>.
- [Grishman 97a]
R. Grishman. Information Extraction: Techniques and Challenges. In *Information Extraction: a Multidisciplinary Approach to an Emerging Information Technology*, Frascati, Italy, 1997. Springer.
- [Grishman 97b]
R. Grishman. Information extraction: Techniques and challenges. In M. T. Paziienza, editor, *Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology, International Summer School (SCIE-97)*, volume 1299 of *Lecture Notes in Computer Science*, pages 10–27, Frascati, Italy, Jul 1997. Springer Verlag.
- [Grishman 97c]
R. Grishman. TIPSTER Architecture Design Document Version 2.3. Technical report, DARPA, 1997. http://www.itl.nist.gov/div894/894.02/related_projects/tipster/.
- [Grishman 01]
R. Grishman. Adaptive Information Extraction and Sublanguage Analysis. In *Proceedings of Workshop on Adaptive Text Extraction and Mining at Seventeenth International Joint Conference on Artificial Intelligence*, Seattle, USA, 2001.
- [Gréar *et al.* 05]
M. Gréar, D. Mladeníć, and M. Grobelnik. User Profile Inference Module. Technical report, SEKT project deliverable D5.5.2, 2005.
- [Grobelnik & Mladenic 06]
M. Grobelnik and D. Mladenic. Knowledge Discovery for Ontology Construction. In J. Davies, R. Studer, and P. Warren, editors, *Semantic Web Technologies*. John Wiley and Sons, 2006.
- [Grosz *et al.* 86]
B. Grosz, K. Sparck-Jones, and B. Webber, editors. *Readings in Natural Language Processing*. Morgan Kaufmann, California, Los Altos, CA, 1986.
- [Grover *et al.* 93]
C. Grover, J. Carroll, and T. Briscoe. The alvey natural language tools, 1993.
- [Grover *et al.* 00]
C. Grover, A. Holt, E. Klein, and M. Moens. Designing a controlled language for interactive model checking. In *CLAW 2000: Third International Workshop on Controlled Language Applications*, Seattle, 2000.
- [Grover *et al.* 02]
C. Grover, S. McDonald, D. N. Gearailt, V. Karkaletsis, D. Farmakiotou, G. Samaritakis, G. Petasis, M. T. Paziienza, M. Vindigni, F. Vichot, and F. Wolinski. Multilingual XML-Based Named Entity Recognition for E-Retail Domains. In *Proceedings of the LREC 2002 Third International Conference on Language Resources and Evaluation*, pages 1060–1067, Las Palmas, Canary Islands - Spain, 2002.
- [Gruber 76]
J. Gruber. *Lexical structures in syntax and semantics*. North-Holland, 1976.

- [Gruber 93]
T. R. Gruber. A Translation Approach to Portable Ontologies. *Knowledge Acquisition*, 5(2):199–220, 1993.
- [Gruenstein *et al.* 09]
E. Gruenstein, I. McGraw, and A. Sutherl. A self-transcribing speech corpus: collecting continuous speech with an online educational game. In *the Speech and Language Technology in Education (SLaTE) Workshop*, 2009.
- [Gruhl *et al.* 04]
D. Gruhl, L. Chavet, D. Gibson, J. Meyer, P. Pattanayak, A. Tomkins, and J. Zien. How to build a WebFountain: An architecture for very large-scale text analytics. *IBM Systems Journal*, 43(1):64–77, 2004.
- [Gruhl *et al.* 09]
D. Gruhl, M. Nagarajan, J. Pieper, C. Robson, and A. Sheth. Context and Domain Knowledge Enhanced Entity Spotting in Informal Text. In *Proceedings of the 8th International Semantic Web Conference (ISWC'2009)*, 2009.
- [Gualtieri & Ruffolo 05]
A. Gualtieri and M. Ruffolo. An ontology-based framework for representing organizational knowledge. In *Proceedings of I-KNOW '05*, Graz, Austria, 2005.
- [Guarino & Welty 00]
N. Guarino and C. Welty. Ontological analysis of taxonomic relationships. In *Proceedings of ER-2000: The 19th International Conference on Conceptual Modeling*, 2000.
- [Guarino & Welty 04]
N. Guarino and C. A. Welty. An overview of ontoclean. In Staab and Studer [Staab & Studer 04], pages 151–172.
- [Guarino *et al.* 99]
N. Guarino, C. Masolo, and G. Vetere. OntoSeek: Content-Based Access to the Web. *IEEE Intelligent Systems*, 14(3):70–80, 1999.
- [Guellec & van Pottelsberghe de laPotterie 07]
D. Guellec and B. van Pottelsberghe de la Potterie. *The Economics of the European Patent System—IP Policy for Innovation and Competition*. Oxford University Press, Oxford, UK, February 2007.
- [Guerin & Miyazaki 06]
B. Guerin and Y. Miyazaki. Analyzing rumours, gossip, and urban legends through their conversational properties. *The Psychological Record*, 56:23–34, 2006.
- [Guerin & Miyazaki 10]
B. Guerin and Y. Miyazaki. Analyzing rumors, gossip, and urban legends through their conversational properties. 56, 01 2010.
- [Guha & Garg 04]
R. V. Guha and A. Garg. Disambiguating People in Search. In *Proceedings of the 13th World Wide Web Conference (WWW 2004)*, ACM Press, 2004.
- [Guha & Lenat 90]
R. Guha and D. Lenat. Cyc: A midterm report. *AI Magazine*, Fall 1990.
- [Guha *et al.* 03]
R. Guha, R. McCool, and E. Miller. Semantic search. In *Proceedings of the 12th international conference on World Wide Web, WWW '03*, pages 700–709, New York, NY, USA, 2003. ACM.

- [Guo 89]
C.-M. Guo. Constructing a Machine Tractable Dictionary from Longman Dictionary of Contemporary English. Technical Report MCCS-89-156, Computing Research Laboratory, New Mexico State University, 1989.
- [Guo *et al.* 03]
L. Guo, F. Shao, C. Botev, and J. Shanmugasundaram. Xrank: ranked keyword search over xml documents. In *Proceedings of the 2003 ACM SIGMOD international conference on Management of data*, SIGMOD '03, pages 16–27, New York, NY, USA, 2003. ACM.
- [Gupta *et al.* 88]
V. Gupta, M. Lennig, and P. Mermelstein. Fast search strategy in a large vocabulary speech recognizer. *Journal of the Acoustical Society of America*, 84:2007–2017, 1988.
- [Gupta *et al.* 10]
N. Gupta, G. Di Fabbrizio, and P. Haffner. Capturing the stars: Predicting ratings for service and product reviews. In *Proceedings of the NAACL HLT 2010 Workshop on Semantic Search*, pages 36–43, Los Angeles, California, June 2010.
- [Gurevych *et al.* 12a]
I. Gurevych, J. ECKLE-KOHLER, S. Hartmann, M. Matuschek, C. M. Meyer, and C. Wirth. Uby - a large-scale unified lexical-semantic resource based on lmf. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 580–590, Avignon, France, 2012.
- [Gurevych *et al.* 12b]
I. Gurevych, J. ECKLE-KOHLER, S. Hartmann, M. Matuschek, C. M. Meyer, and C. Wirth. UBY: A large-scale unified lexical-semantic resource based on LMF. In *13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 580–590, 2012.
- [Guthrie *et al.* 90]
L. Guthrie, B. Slator, Y. Wilks, and R. Bruce. Is there content in empty heads? In *Proceedings of the 13th International Conference on Computational Linguistics (COLING-90)*, volume 3, pages 138–143, Helsinki, Finland, 1990.
- [Guthrie *et al.* 91]
J. Guthrie, L. Guthrie, Y. Wilks, and H. Aidinejad. Subject Dependent Co-occurrence and Word Sense Disambiguation. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pages 146–152, 1991.
- [Guthrie *et al.* 04]
L. Guthrie, R. Basili, F. Zanzotto, K. Bontcheva, H. Cunningham, D. Guthrie, J. Cui, M. Cammisa, J. Liu, C. Martin, K. Haralambiev, M. Holub, K. Machery, and F. Jelinek. Large Scale Experiments for Semantic Labelling of Noun Phrases in Raw Text. In *Proceedings of the Language Resources and Evaluation Conference (LREC 2004)*, 2004.
- [Guy *et al.* 15]
I. Guy, R. Levin, T. Daniel, and E. Bolshinsky. Islands in the stream: A study of item recommendation within an enterprise social stream. In R. A. Baeza-Yates, M. Lalmas, A. Moffat, and B. A. Ribeiro-Neto, editors, *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, Santiago, Chile, August 9-13, 2015*, pages 665–674. ACM, 2015.
- [Haas *et al.* 11]
K. Haas, P. Mika, P. Tarjan, and R. Blanco. Enhanced results for web search. In *Proceeding of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pages 725–734, 2011.

- [Haase *et al.* 04a]
P. Haase, Y. Sure, and D. Vrandečić. D3.1.1 Ontology Management and Evolution – Survey, Methods and Prototypes. Technical report, SEKT EU Project Deliverable, 2004.
- [Haase *et al.* 04b]
P. Haase, M. Ehrig, A. Hotho, and B. Schnizler. Personalized Information Access in a Bibliographic Peer-to-Peer System. In *Proceedings of the AAAI Workshop on Semantic Web Personalization*, pages 1–12. AAAI Press, July 2004.
- [Hachey *et al.* 05]
B. Hachey, M. Becker, C. Grover, and E. Klein. Selective Sampling for Information Extraction with a Committee of Classifiers. In *PASCAL Challenges Workshop*, Southampton, 2005.
- [Hacioglu *et al.* 04]
K. Hacioglu, S. Pradhan, W. Ward, J. H. Martin, and D. Jurafsky. Semantic Role Labeling by Tagging Syntactic Chunks. In *Proceedings of CoNLL-2004*, pages 110–113, 2004.
- [Haddad *et al.* 05]
H. Haddad, L. M. Liebrock, A. Omicini, and R. L. Wainwright, editors. *Proceedings of the 2005 ACM Symposium on Applied Computing (SAC), Santa Fe, New Mexico, USA, March 13-17, 2005*. ACM, 2005.
- [Haeb-Umbach & Ney 94]
R. Haeb-Umbach and H. Ney. Improvements in beam search for 10 000-word continuous-speech recognition. *IEEE Transactions on Signal Processing*, 2:353–356, 1994.
- [Haeberli 88]
P. Haeberli. ConMan: A Visual Programming Language for Interactive Graphics. *ACM Computer Graphics*, 22(4):103–111, 1988. ACM SIGGRAPH '88.
- [Hahn & Schnattinger 98]
U. Hahn and K. Schnattinger. Towards text knowledge engineering. In *Proc. of 15th National Conference on Artificial Intelligence (AAAI-98)*, pages 524–531, Menlo Park, CA, 1998. MIT Press.
- [Hakimov *et al.* 12]
S. Hakimov, S. A. Oto, and E. Dogdu. Named Entity Recognition and Disambiguation using Linked Data and Graph-based Centrality Scoring. In *Proceedings of the 4th International Workshop on Semantic Web Information Management*, 2012.
- [Halaschek-Wiener *et al.* 06]
C. Halaschek-Wiener, J. Golbeck, A. Schain, M. Grove, B. Parsia, and J. A. Hendler. Annotation and Provenance Tracking in Semantic Web Photo Libraries. In *IPAW*, volume 4145 of *Lecture Notes in Computer Science*. Springer, 2006.
- [Halevy *et al.* 09]
A. Halevy, P. Norvig, and F. Pereira. The unreasonable effectiveness of data. *IEEE Intelligent Systems*, 24(2):8–12, 2009.
- [Hallett 06]
C. Hallett. Generic Querying of Relational Databases Using Natural Language Generation Techniques. In *Proceedings of the Fourth International Natural Language Generation Conference*, pages 95–102. Association for Computational Linguistics, July 2006.
- [Halliday 85]
M. Halliday. *An Introduction to Functional Grammar*. Edward Arnold, London, 1985.

- [Halpin & Lavrenko 11]
H. Halpin and V. Lavrenko. Relevance feedback between hypertext and semantic web search: Frameworks and evaluation. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(4), 2011.
- [Hamidian & Diab 16]
S. Hamidian and M. T. Diab. Rumor identification and belief investigation on twitter. In *Proceedings of NAACL-HLT*, pages 3–8, 2016.
- [Hammond & Pasin 15]
T. Hammond and M. Pasin. The nature.com ontologies portal. In *5th Workshop on Linked Science*, 2015.
- [Hammond *et al.* 05]
T. Hammond, T. Hannay, and et al. Social bookmarking tools (i): A general review. *D-Lib Magazine*, 11(4), April 2005.
- [Hammond *et al.* 10]
M. Hammond, R. Hawtin, L. Gillam, and C. Oppenheim. Cloud computing for research – Final Report. Technical report, JISC, 2010. <http://www.jisc.ac.uk/whatwedo/programmes/researchinfrastructure/usingcloudcomp.aspx>.
- [Hampapur 99a]
A. Hampapur. Semantic video indexing: approach and issues. *SIGMOD-Record*, 28(1):32–9, March 1999.
- [Hampapur 99b]
A. Hampapur. Semantic video indexing: Approach and issues. *ACM Sigmod Record*, 28(1), March 1999.
- [Hampton *et al.* 11]
K. N. Hampton, L. S. Goulet, L. Rainie, and K. Purcell. Social networking sites and our lives: How people’s trust, personal relationships, and civic and political involvement are connected to their use of social networking sites and other technologies. Technical Report June 16, Pew Research Center’s Internet and American Life Project, <http://pewinternet.org/Reports/2011/Technology-and-social-networks.aspx>, 2011.
- [Hamza *et al.* 02]
O. Hamza, D. M. V. Tablan, C. Ursu, H. Cunningham, and Y. Wilks. Named Entity Recognition in Romanian. Technical report, Department of Computer Science, University of Sheffield, 2002.
- [Hamza *et al.* 03]
O. Hamza, K. Bontcheva, D. Maynard, V. Tablan, and H. Cunningham. Named Entity Recognition in Romanian. In *Proceedings of Workshop on Information Extraction for Slavonic and other Central and Eastern European Languages (IESL’03)*, Borovets, Bulgaria, 2003.
- [Han & Baldwin 11]
B. Han and T. Baldwin. Lexical normalisation of short text messages: makin sens a #twitter. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, HLT ’11, pages 368–378, 2011.
- [Han *et al.* 03]
H. Han, C. L. Giles, and H. Zha. A model-based k-means algorithm for name disambiguation. In *Proceedings of the 2nd International Semantic Web Technologies for Searching and Retrieving Scientific Data*, Florida, USA, 2003.

- [Han *et al.* 04]
H. Han, L. Giles, H. Zha, C. Li, and K. Tsioutsoulouklis. Two supervised learning approaches for name disambiguation in author citations. In *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL'04)*, 2004.
- [Han *et al.* 12]
B. Han, P. Cook, and T. Baldwin. Automatically constructing a normalisation dictionary for microblogs. In *Proceedings of the conference on Empirical Methods in Natural Language Processing*, pages 421–432. ACL, 2012.
- [Han *et al.* 14]
B. Han, A. J. Yepes, A. MacKinlay, and Q. Chen. Identifying twitter location mentions. In *Proceedings of the 2014 Australasian Language Technology Association Workshop*, pages 157–162, 2014.
- [Handschuh & Staab 02]
S. Handschuh and S. Staab. Authoring and Annotation of Web Pages in CREAM. In *Proceedings of WWW'02*, Honolulu, Hawaii, USA, 2002.
- [Handschuh 05]
S. Handschuh. *Creating Ontology-based Metadata by Annotation for the Semantic Web*. Unpublished PhD thesis, 2005.
- [Handschuh *et al.* 01]
S. Handschuh, S. Staab, and A. Maedche. CREAM – creating relational metadata with a component-based, ontology-driven framework. In *Proceedings of K-CAP 2001*, Victoria, BC, Canada, 2001.
- [Handschuh *et al.* 02]
S. Handschuh, S. Staab, and F. Ciravegna. S-CREAM — Semi-automatic CREAtion of Metadata. In *13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02)*, pages 358–372, Sigüenza, Spain, 2002.
- [Hanks 79]
P. Hanks. *Collins English Dictionary*. Collins, London & Glasgow, 1979.
- [Hansen *et al.* 99]
K. Hansen, C. Yndigeñ, and K. Grnbk. Dynamic use of digital library material - supporting users with typed links in open hypermedia. In *Proceedings of the European Conference on Digital Libraries*, Paris, France, September 1999. Springer.
- [Harabagiu & Hickl 11]
S. Harabagiu and A. Hickl. Relevance Modeling for Microblog Summarization. In *Proceedings of the Fifth International Conference on Weblogs and Social Media (ICWSM)*, 2011.
- [Harabagiu & Maiorano]
S. Harabagiu and S. J. Maiorano. Acquisition of linguistic patterns for knowledge-based information extraction.
- [Hardie *et al.* 06]
A. Hardie, P. Baker, T. McEnery, and B. Jayaram. Corpus-building for South Asian languages. *Trends in Linguistics. Studies and Monographs*, 175:211–242, 2006.
- [Hare *et al.* 06]
J. S. Hare, P. H. Lewis, P. G. B. Enser, and C. J. Sandom. A linear-algebraic technique with an application in semantic image retrieval. In H. Sundaram, M. R. Naphade, J. R. Smith, and Y. Rui, editors, *CIVR*, volume 4071 of *Lecture Notes in Computer Science*, pages 31–40. Springer, 2006.

- [Hare *et al.* 11]
 J. S. Hare, S. Samangooei, and D. P. Dupplaw. OpenIMAJ and ImageTerrier: Java libraries and tools for scalable multimedia analysis and indexing of images. In *Proceedings of the 19th ACM International Conference on Multimedia*, pages 691–694, New York, NY, USA, 2011. ACM.
- [Hare *et al.* 13]
 J. Hare, S. Samangooei, D. Dupplaw, and P. H. Lewis. Twitter’s visual pulse. In *3rd ACM International conference on multimedia retrieval*, pages 297–298, April 2013.
- [Hargittai *et al.* 12a]
 E. Hargittai, W. R. Neuman, and O. Curry. Taming the information tide: Perceptions of information overload in the american home. *Inf. Soc.*, 28(3):161–173, 2012.
- [Hargittai *et al.* 12b]
 E. Hargittai, W. R. Neuman, and O. Curry. Taming the information tide: Perceptions of information overload in the american home. *Inf. Soc.*, 28(3):161–173, 2012.
- [Harkema H 04]
 H. M. Harkema H, Gaizauskas R. A large scale terminology resource for biomedical text processing. In *Proceedings of the HLT-NAACL Workshop: BioLINK 2004, Linking Biological Literature, Ontologies and Databases*, 2004.
- [Harley & Glennon 97]
 A. Harley and D. Glennon. Sense tagging in action: Combining different tests with additive weights. In *Proceedings of the SIGLEX Workshop “Tagging Text with Lexical Semantics”*, pages 74–78. Association for Computational Linguistics, Washington, D.C., 1997.
- [Harris 54]
 Z. Harris. Distributional structure. *Word*, 10(23):146–162, 1954.
- [Harris 68]
 Z. Harris. *Mathematical Structures of Language*. Wiley (Interscience), New York, 1968.
- [Harrison 91]
 P. Harrison. Evaluating Syntax Performance of Parsers/Grammars of English. In *Proceedings of the Workshop on Evaluating Natural Language Processing Systems, ACL*, 1991.
- [Hartmann *et al.* 05]
 J. Hartmann, Y. Sure, P. Haase, and M. del Carmen Suarez-Figueroa. OMV – Ontology Metadata Vocabulary. In *ISWC 2005*, 2005.
- [Hasbullah & Maynard 16]
 S. S. Hasbullah and D. Maynard. Automated Content Analysis: A Sentiment Analysis on Malaysian Government Social Media. In *ACM IMCOM (International Conference on Ubiquitous Information Management and Communication)*, Danang, Vietnam, 2016.
- [Hashimoto & Yukawa 07]
 K. Hashimoto and T. Yukawa. Term weighting classification system using the chi-square statistic for the classification subtask at ntcir-6 patent retrieval task. In *Proceedings of the Sixth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access*, pages 385–389, 2007.
- [Hassan *et al.* 09]
 A. Hassan, D. R. Radev, J. Cho, and A. Joshi. Content based recommendation and summarization in the blogosphere. In *Proceedings of the Third International Conference on Weblogs and Social Media (ICWSM)*, 2009.

- [Hatcher & Gospodnetic 04]
E. Hatcher and O. Gospodnetic. *Lucene in Action (In Action series)*. Manning Publications Co., Greenwich, CT, USA, 2004.
- [Hatzivassiloglou & McKeown 97]
V. Hatzivassiloglou and K. McKeown. Predicting the semantic orientation of adjectives. In *ACL*, pages 174–181, 1997.
- [Hatzivassiloglou 96]
V. Hatzivassiloglou. Do we need linguistics when we have statistics? a comparative analysis of the contributions of linguistic cues to a statistical word grouping system. In J. Klavans and P. Resnick, editors, *The Balancing Act: combining symbolic and statistical approaches to language*, pages 67–94. MIT Press, 1996.
- [Hauff *et al.* 13]
C. Hauff, B. Thomee, and M. Trevisiol. Working Notes for the Placing Task at MediaEval 2013, 2013.
- [Haugeland 81]
J. Haugeland, editor. *Mind Design*. MIT Press, Cambridge, MA, 1981.
- [Havel 93]
I. Havel. Artificial thought and emergent mind. In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, volume 1, pages 758–766, 1993.
- [Haveliwala *et al.* 03]
T. Haveliwala, S. Kamvar, and G. Jeh. An analytical comparison of approaches to personalizing pagerank. Technical report, Stanford University, 2003.
- [Hayes-Roth 94]
F. Hayes-Roth. Architecture-Based Acquisition and Development of Software: Guidelines and Recommendations from the ARPA Domain-Specific Software Architecture (DSSA) Program. Technical report, Techknowledge Federal Systems, 1994. <http://www.oswego.com/dssa/>, visited 29th March 1999.
- [Hearst 92]
M. A. Hearst. Automatic acquisition of hyponyms from large text corpora. In *Conference on Computational Linguistics (COLING'92)*, Nantes, France, 1992. Association for Computational Linguistics.
- [Hecht *et al.* 11]
B. Hecht, L. Hong, B. Suh, and E. H. Chi. Tweets from Justin Bieber’s heart: The dynamics of the location field in user profiles. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, New York, NY, USA, 2011.
- [Heckman *et al.* 07]
D. Heckman, E. Schwarzkopf, D. Mori, J. Dengler, and A. Krner. The user model and context ontology GUMO revisited for future Web 2.0 extensions. In *Proceedings of the Int. Workshop on Contexts and Ontologies: Representation and Reasoning. CEUR Workshop Proceedings*, volume 298, 2007.
- [Heckmann *et al.* 05]
D. Heckmann, T. Schwartz, B. Brandherm, M. Schmitz, and M. von Wilamowitz-Moellendorff. GUMO - the general user model ontology. In *Proceedings of the 10th International Conference on User Modeling*, pages 428–432, 2005.
- [Heilman & Smith 10]
M. Heilman and N. A. Smith. Rating computer-generated questions with mechanical turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, pages 35–40, 2010.

- [Hendler & Stoffel 99]
 J. Hendler and K. Stoffel. Back-End Technology for High-Performance Knowledge Representation Systems. *IEEE Intelligent Systems*, 14(3):63–69, 1999.
- [Hendrix *et al.* 77]
 G. G. Hendrix, E. D. Sacerdoti, D. Sagalowicz, and J. Slocum. Developing a natural language interface to complex data. In *VLDB '1977: Proceedings of the third international conference on Very large data bases*, pages 292–292. VLDB Endowment, 1977.
- [Hendry & Harper 96]
 D. Hendry and D. Harper. An Architecture for Implementing Extensible Information-Seeking Environments. In *Proceedings of SIGIR-96.*, pages 94–100, 1996.
- [Henrich 96]
 A. Henrich. Document Retrieval Facilities for Repository-Based System Development Environments. In *Proceedings of SIGIR '96*, pages 101–109, Zurich, 1996.
- [Hepp 10]
 M. Hepp. HyperTwitter: Collaborative knowledge engineering via Twitter messages. In *Knowledge Engineering and Management by the Masses - 17th International Conference EKAW 2010*, pages 451–461, 2010.
- [Hepp *et al.* 05]
 M. Hepp, D. Bachlechner, and K. Siorpaes. Ontowiki: Community-driven ontology engineering and ontology usage based on wikis, 2005.
- [Hepp *et al.* 06]
 M. Hepp, D. Bachlechner, and K. Siorpaes. Harvesting Wiki Consensus - Using Wikipedia Entries as Ontology Elements. In *Proceedings of the 1st Workshop "SemWiki2006 - From Wiki to Semantics" at ESWC 2006*, 2006.
- [Hepple 00a]
 M. Hepple. Independence and Commitment: Assumptions for Rapid Training and Execution of Rule-based Part-of-Speech Taggers. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics*, Hong Kong, October 2000.
- [Hepple 00b]
 M. Hepple. Independence and commitment: Assumptions for rapid training and execution of rule-based POS taggers. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics (ACL-2000)*, Hong Kong, October 2000.
- [Herman J.M. Steeneken 95]
 D. A. V. L. Herman J.M. Steeneken. Multi-lingual assessment of speaker independent large vocabulary speech-recognition systems: the SQALE project (speech recognition quality assessment for language engineering). In *Proceedings European Conference on Speech Communication and Technology*, 1995.
- [Hermansky 90]
 H. Hermansky. Perceptual linear predictive (PLP) analysis of speech. *Journal of the Acoustical Society of America*, 87:1738–1752, 1990.
- [Hermansky *et al.* 91]
 H. Hermansky, N. Morgan, A. Bayya, and P. Kohn. RASTA-PLP speech analysis. Technical Report TR-91-069, International Computer Science Institute, Berkeley CA, 1991.
- [Herr *et al.* 04]
 P. Herr, F. Kardes, and J. Kim. Effects of word-of-mouth and product attribute information on persuasion: an accessibility-diagnostics perspective. *Journal of Consumer Research*, 17:454–462, 2004.

- [Hersh & Leone 95]
W. Hersh and T. Leone. The sapphire server. In *Proceedings of the 19th Annual Symposium on Computer Applications in Medical Care*, pages 858–862, 1995.
- [Hersh 05]
W. Hersh. Evaluation of biomedical text-mining systems: Lessons learned from information retrieval. *Briefings in Bioinformatics*, 6(4):344–356, 2005.
- [Herviou-Picard 96]
M.-L. Herviou-Picard. Les outils d’indexation AlethIP issus du projet GRAAL: principes et utilisation. Technical Report HN-46/96/022, Direction des Etudes et Recherches, EDF, Clamart, France, 1996.
- [Herzog *et al.* 04]
G. Herzog, A. Ndiaye, S. Merten, H. Kirchmann, T. Becker, and P. Poller. Large-Scale Software Integration for Spoken Language and Multimodal Dialog Systems. *Natural Language Engineering*, 2004.
- [Hess & Kushmerick 03]
A. Hess and N. Kushmerick. Learning to attach semantic metadata to web services. In *Proceedings of the Second International Semantic Web Conference (ISWC 2003)*, pages 258–273. Springer, 2003.
- [Heß & Kushmerick 04]
A. Heß and N. Kushmerick. Machine learning for annotating semantic web services. In *AAAI Spring Symposium on Semantic Web Services*, March 2004.
- [Heß *et al.* 04]
A. Heß, E. Johnston, and N. Kushmerick. ASSAM: A tool for semi-automatically annotating web services with semantic metadata. In *International Semantic Web Conference (ISWC)*, Hiroshima, November 2004.
- [Hettne *et al.* 10]
K. Hettne, E. van Mulligan, M. Schuemie, B. Schijvennaars, and J. Kors. Rewriting and suppressing UMLS terms for improved biomedical term identification. *Journal of Biomedical Semantics*, 1(5):1–14, 2010.
- [Hewitt *et al.* 80]
C. Hewitt, G. Attardi, and M. Simi. Knowledge embedding in the description system omega. *Proceedings of the First National Annual Conference on Artificial Intelligence*, 1980.
- [Heyer & Witschel 05]
G. Heyer and H. Witschel. Terminology and metadata—on how to efficiently build an ontology. *TermNet News – Newsletter of International Cooperation in Terminology*, 87, 2005.
- [Hiemstra *et al.* 01]
D. Hiemstra, W. Kraaij, R. Pohlmann, and T. Westerveld. Twenty-one at clef2000: Translation resources, merging strategies and relevance feedback. In C. Peters, editor, *Cross-language Information Retrieval and Evaluation, Proceedings of CLEF 2000 workshop*, pages 102–116. Springer-Verlag, 2001.
- [Hilbert & López 11]
M. Hilbert and P. López. The World’s Technological Capacity to Store, Communicate, and Compute Information. *Science*, 332(6025):60–65, 2011.
- [Hildebrand *et al.* 77]
D. K. Hildebrand, J. D. Laing, and H. Rosenthal. *Analysis of Ordinal Data*. Quantitative Applications in the Social Sciences. Sage, 1977.

- [Hildebrand *et al.* 06]
M. Hildebrand, J. van Ossenbruggen, and J. Hardman. /facet: A Browser for Heterogeneous Semantic Web Repositories. In *Proceedings of the 5th International Semantic Web Conference*, 2006.
- [Hils 92]
D. Hils. Visual Languages and Computing Survey: Data Flow Visual Programming Languages. *Journal of Visual Languages and Computing*, pages 69–101, 1992.
- [Hiltz & Turoff 85]
S. R. Hiltz and M. Turoff. Structuring computer-mediated communication systems to avoid information overload. *Commun. ACM*, 28:680–689, July 1985.
- [Hindle & Rooth 90]
D. Hindle and M. Rooth. Structural ambiguity and lexical relations. In *Proceedings of the DARPA Speech and Natural Language Workshop*. 1990.
- [Hindle 90]
D. Hindle. Noun classification from predicate-argument structures. In *Proc. of 28th Annual Meeting of the Association for Computational Linguistics*, pages 268–275, Pittsburgh, Pennsylvania, 1990.
- [Hirschhorn & Daly 05]
J. N. Hirschhorn and M. J. Daly. Genome-wide association studies for common diseases and complex traits. *NATURE REVIEWS-GENETICS*, 6:95–108, February 2005.
- [Hirschman & Sager 82]
L. Hirschman and N. Sager. Automatic information formatting of a medical sublanguage. In Kittredge and Lehrberger, editors, *Sublanguage: Studies of Language in Restricted Semantic Domains*, pages 27–69. Walter de Gruyter, 1982.
- [Hirschman *et al.* 75]
L. Hirschman, R. Grishman, and N. Sager. Grammatically based automatic word class formation. *Information Processing and Retrieval*, 11:39–57, 1975.
- [Hirst 81]
G. Hirst. *Anaphora in Natural Language Understanding*. Springer-Verlag, Berlin, 1981.
- [Hirst 87]
G. Hirst. *Semantic Interpretation and the Resolution of Ambiguity*. Cambridge University Press, Cambridge, 1987.
- [Hoadley *et al.* 10]
C. M. Hoadley, H. Xu, J. J. Lee, and M. B. Rosson. Privacy as information access and illusory control: The case of the facebook news feed privacy outcry. *Electronic Commerce Research and Applications*, 9(1):50 – 60, 2010. Special Issue: Social Networks and Web 2.0.
- [Hobbs & Rosenschein 77]
J. Hobbs and S. Rosenschein. Making computational sense of montague’s intensional logic. *Artificial Intelligence*, 9, 1977.
- [Hobbs 85]
J. Hobbs. Resolving pronoun references. In B. G. K. S.-J. B. Webber, editor, *Readings in Natural Language*, volume 44, pages 311–338. Morgan Kaufmann, California, 1985.
- [Hobbs 86]
J. Hobbs. The TACITUS project. *Computational Linguistics*, 12(3):220–222, 1986.
- [Hobbs 91]
J. Hobbs. Description of the TACITUS system as used for MUC-3. In *Proceedings of the Third Message Understanding Conference MUC-3*, pages 200–206. Morgan Kaufmann, California, 1991.

- [Hobbs 93]
J. Hobbs. The Generic Information Extraction System. In *Proceedings of the Fifth Message Understanding Conference (MUC-5)*. Morgan Kaufmann, California, 1993. http://www.itl.nist.gov/div894/894.02/related_projects/tipster/gen_ie.htm.
- [Hobbs et al. 92]
J. Hobbs, D. Appelt, M. Tyson, J. Bear, and D. Israel. Description of the FASTUS system as used for MUC-4. In *Proceedings of the Fourth Message Understanding Conference MUC-4*, pages 268–275. Morgan Kaufmann, California, 1992.
- [Hobbs et al. 93]
J. Hobbs, M. Stickel, D. Appelt, and P. Martin. Interpretation as abduction. *Artificial Intelligence*, 63:69–142, 1993.
- [Hobbs et al. 96]
J. Hobbs, D. Appelt, J. Bear, D. Israel, M. Kameyama, M. Stickel, and M. Tyson. FASTUS: A cascaded finite-state transducer for extracting information from natural language text. In *Finite State Devices for Natural Language Processing*, pages 383–406. MIT Press, 1996.
- [Hochberg et al. 94a]
M. M. Hochberg, S. J. Renals, and A. J. Robinson. ABBOT: The CUED hybrid connectionist-hmm large-vocabulary recognition system. In *Proceedings of Spoken Language Systems Technology Workshop*, 1994.
- [Hochberg et al. 94b]
M. M. Hochberg, S. J. Renals, A. J. Robinson, and D. J. Kershaw. Large vocabulary continuous speech recognition using a hybrid connectionist-HMM system. pages 1499–1502, Yokohama, Japan, 1994.
- [Hochberg et al. 95]
M. M. Hochberg, S. J. Renals, A. J. Robinson, and G. D. Cook. Recent improvements to the ABBOT large vocabulary CSR system. pages 69–72, Detroit, 1995.
- [Hodges 85a]
W. Hodges. The logical basis of prolog. Paper presented at a meeting on “Prolog and its Relevance to Philosophy” at King’s College on October 25, 1985.
- [Hodges 85b]
W. Hodges. Reducing first-order logic to horn logic. Unpublished Paper, 1985.
- [Hoefler 04]
S. Hoefler. The syntax of Attempto Controlled English: An abstract grammar for ACE 4.0. Technical Report ifi-2004.03, Department of Informatics, University of Zurich, 2004.
- [Hoffart et al. 11]
J. Hoffart, M. A. Yosef, I. Bordino, H. Furstenuau, M. Pinkal, M. Spaniol, B. Taneva, S. Thater, and G. Weikum. Robust disambiguation of named entities in text. In *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 782–792, 2011.
- [Hoffart et al. 14]
J. Hoffart, Y. Altun, and G. Weikum. Discovering emerging entities with ambiguous names. In *Proceedings of the 23rd International Conference on World Wide Web, WWW ’14*, pages 385–396, 2014.
- [Hoffmann & Valencia]
R. Hoffmann and A. Valencia. Implementing the ihop concept for navigation of biomedical literature. *Bioinformatics*, 21(suppl 2):ii252–ii258.

- [Hoffmann 09]
L. Hoffmann. Crowd control. *Communications of the ACM*, 52(3):16–17, 2009.
- [Hogan *et al.* 06]
A. Hogan, A. Harth, and S. Decker. Reconrank: A scalable ranking method for semantic web data with context. In *Second International Workshop on Scalable Semantic Web Knowledge Base Systems (SSWS 2006)*, Athens, GA, USA, 2006.
- [Hogger 84]
C. Hogger. *Introduction to Logic Programming*. Academic Press, London, 1984.
- [Holmberg & Hellsten 14]
K. Holmberg and I. Hellsten. Analyzing the climate change debate on twitter – content and differences between genders. In *Proceedings of the ACM WebScience conference*, pages 287–288, Bloomington, IN, USA, 2014.
- [Honeycutt & Herring 09]
C. Honeycutt and S. Herring. Beyond microblogging: Conversation and collaboration via Twitter. In *Proceedings of the 42nd Hawaii International Conference on System Sciences*, pages 1–10, 2009.
- [Hong *et al.* 11]
L. Hong, G. Convertino, and E. H. Chi. Language matters in twitter: A large scale study. In L. A. Adamic, R. A. Baeza-Yates, and S. Counts, editors, *Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, July 17-21, 2011*. The AAAI Press, 2011.
- [Hong *et al.* 12]
L. Hong, R. Bekkerman, J. Adler, and B. D. Davison. Learning to rank social update streams. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '12*, pages 651–660, New York, NY, USA, 2012. ACM.
- [Hong *et al.* 13a]
L. Hong, A. S. Doumith, and B. D. Davison. Co-factorization machines: Modeling user interests and predicting individual decisions in twitter. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining, WSDM '13*, pages 557–566, New York, NY, USA, 2013. ACM.
- [Hong *et al.* 13b]
L. Hong, A. S. Doumith, and B. D. Davison. Co-factorization machines: Modeling user interests and predicting individual decisions in twitter. In *WSDM*, 2013.
- [Horacek 97]
H. Horacek. Designing and Exploiting an Ontology for Multi-Lingual Generation. In *IJCAI-97 Workshop on Ontologies and Multilingual NLP*, 1997.
- [Hornby 80]
A. Hornby, editor. *Oxford Advanced Learner's Dictionary of Current English*. Oxford University Press, 1980.
- [Horowitz & Hill 89]
P. Horowitz and W. Hill. *The Art of Electronics*. Cambridge University Press, 1989.
- [Horridge *et al.* 06]
M. Horridge, N. Drummond, J. Goodwin, A. L. Rector, R. Stevens, and H. Wang. The manchester owl syntax. In *OWLed*, volume 216, 2006.
- [Horrocks & Patel-Schneider 03]
I. Horrocks and P. Patel-Schneider. Three theses of representation in the semantic web, 2003.

- [Horrocks & vanHarmelen 01]
I. Horrocks and F. van Harmelen. Reference Description of the DAML+OIL (March 2001) Ontology Markup Language. Technical report, 2001. <http://www.daml.org/2001/03/reference.html>.
- [Hosseini *et al.* 10]
M. Hosseini, I. J. Cox, N. Millic-Frayling, and V. Vinay. Measuring performance volatility of retrieval systems. In H. Cunningham, A. Hanbury, and S. Rüger, editors, *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference)*. LNCS volume number: 6107, Lecture Notes in Computer Science, Vienna, Austria, May 2010. Springer.
- [Hotho *et al.* 03]
A. Hotho, S. Staab, and G. Stumme. WordNet improves text document clustering. In *Proc. of the SIGIR 2003 Semantic Web Workshop*, 2003.
- [Hovy & Arens 91]
E. Hovy and Y. Arens. Automatic generation of formatted text. In *Proceedings of the 8th AAAI Conference*, pages 92 – 97, Anaheim, CA, 1991.
- [Hovy 88a]
E. Hovy. *Generating Natural Language Under Pragmatic Constraints*. Lawrence Erlbaum, Hillsdale, NJ, 1988.
- [Hovy 88b]
E. Hovy. Generating natural language with a phrasal lexicon. In D. McDonald and L. Bolc, editors, *Natural Language Generation Systems*, pages 353 – 384. Springer-Verlag, Berlin, 1988.
- [Hovy 91]
E. H. Hovy. Approaches to the planning of coherent text. In C. L. Paris, W. R. Swartout, and W. C. Mann, editors, *Natural language generation in artificial intelligence and computational linguistics*. Kluwer Academic Publishers, July 1991. Presented at the Fourth International Workshop on Natural Language Generation. Santa Catalina Island, CA, July, 1988.
- [Hovy 10]
E. Hovy. Annotation. In *Tutorial Abstracts of ACL*, 2010.
- [Hovy *et al.* 06]
E. Hovy, M. P. Marcus, M. Palmer, L. A. Ramshaw, and R. M. Weischedel. Ontonotes: The 90% solution. In *Proceedings of HLT-NAACL*, 2006.
- [Hovy *et al.* 13a]
D. Hovy, T. Berg-Kirkpatrick, A. Vaswani, and E. Hovy. Learning whom to trust with MACE. In *Proc. of NAACL-HLT*, pages 1120–1130, 2013.
- [Hovy *et al.* 13b]
E. Hovy, T. Mitamura, and M. Palmer, editors. *Workshop on Events: Definition, Detection, Coreference, and Representation*. Association for Computational Linguistics, Atlanta, Georgia, June 2013.
- [Hovy *et al.* 14]
D. Hovy, B. Plank, and A. Søgaard. When POS data sets don't add up: Combatting sample bias. In *Proc. LREC*, pages 4472–4475. LREC, 2014.
- [Howe 98]
D. Howe. The free on-line dictionary of computing. <http://wombat.doc.ic.ac.uk/foldoc/>, 1998.
- [Hripcsak & Heitjan 02]
G. Hripcsak and D. Heitjan. Measuring agreement in medical informatics reliability studies. *Journal of Biomedical Informatics*, 35:99–110, 2002.

- [Hripcsak & Rothschild 05]
G. Hripcsak and A. S. Rothschild. Agreement, the F-measure, and Reliability in Information Retrieval. *Journal of the American Medical Informatics Association*, 12(3):296–298, 2005.
- [Hristovski *et al.* 05]
D. Hristovski, B. Peterlin, J. A. Mitchell, and S. M. Humphrey. Using literature-based discovery to identify disease candidate genes. *International Journal of Medical Informatics*, 74(2–4):289 – 298, 2005. <ce:title>MIE 2003</ce:title>.
- [Hsu & Lin 02]
C.-W. Hsu and C.-J. Lin. A comparison of methods for multi-class support vector machines. *IEEE Transactions on Neural Networks*, 13:415–425, 2002.
- [Hsueh *et al.* 09]
P.-Y. Hsueh, P. Melville, and V. Sindhwani. Data quality from crowdsourcing: a study of annotation selection criteria. In *Proceedings of the NAACL HLT 2009 Workshop on Active Learning for Natural Language Processing*, pages 27–35, 2009.
- [Hu & Liu 04a]
M. Hu and B. Liu. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177. ACM, 2004.
- [Hu & Liu 04b]
M. Hu and B. Liu. Mining and summarizing customer reviews. In *KDD '04: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177, New York, NY, USA, 2004. ACM.
- [Hu *et al.* 03]
B. Hu, S. Dasmahapatra, and N. Shadbolt. From Lexicon To Mammographic Ontology: Experiences and Lessons. In D. Calvanese, G. De Giacomo, and E. Franconi, editors, *Proceedings of the International Workshop on Description Logics (DL'2003)*, pages 229–233, 2003.
- [Hu *et al.* 07]
M. Hu, A. Sun, and E.-P. Lim. Comments-oriented blog summarization by sentence extraction. In *Proceedings of the Sixteenth ACM conference on Conference on Information and Knowledge Management (CIKM)*, pages 901–904, 2007.
- [Hu *et al.* 13]
Y. Hu, K. Talamadupula, S. Kambhampati, et al. Dude, srsly?: The surprisingly formal nature of Twitter’s language. *Proceedings of ICWSM*, 2013.
- [Huang & Qiu 08]
Z. Huang and Y. Qiu. Construction and aggregation of citation semantic link network. In *SKG '08: Proceedings of the 2008 Fourth International Conference on Semantics, Knowledge and Grid*, pages 247–254, Washington, DC, USA, 2008. IEEE Computer Society.
- [Huang *et al.* 90]
D. Huang, Y. Ariki, and M. Jack. *Hidden Markov Models for Speech Recognition*. Edinburgh University Press, 1990.
- [Huang *et al.* 11a]
M. Huang, Y. Yang, and X. Zhu. Quality-biased Ranking of Short Texts in Microblogging Services. In *Proceedings of the 5th International Joint Conference on Natural Language Processing*, pages 373–382, 2011.
- [Huang *et al.* 11b]
M. Huang, Y. Yang, and X. Zhu. Quality-biased Ranking of Short Texts in Microblogging Services. In *IJCNLP*, pages 373–382, 2011.

- [Huang *et al.* 12a]
H. Huang, A. Zubiaga, H. Ji, H. Deng, D. Wang, H. K. Le, T. F. Abdelzaher, J. Han, A. Leung, J. P. Hancock, and C. R. Voss. Tweet ranking based on heterogeneous networks. In *COLING 2012, 24th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers, 8-15 December 2012, Mumbai, India*, pages 1239–1256, 2012.
- [Huang *et al.* 12b]
H. Huang, A. Zubiaga, H. Ji, H. Deng, D. Wang, H. K. Le, T. F. Abdelzaher, J. Han, A. Leung, J. P. Hancock, and C. R. Voss. Tweet ranking based on heterogeneous networks. In *COLING*, pages 1239–1256, 2012.
- [Huang *et al.* 14]
H. Huang, Y. Cao, X. Huang, H. Ji, and C.-Y. Lin. Collective tweet wikification based on semi-supervised graph regularization. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pages 380–390, 2014.
- [Hubmann-Haidvogel *et al.* 12]
A. Hubmann-Haidvogel, A. M. P. Brasoveanu, A. Scharl, M. Sabou, and S. Gindl. Visualizing contextual and dynamic features of micropost streams. In *Proceedings of the #MSM2012 Workshop, CEUR*, volume 838, 2012.
- [Huertas-Rosero *et al.* 08]
A. F. Huertas-Rosero, L. A. Azzopardi, and K. van Rijsbergen. Characterising through erasing. In *Proceedings of the Second Quantum Interaction Symposium (QI-2008)*. College Publications, 2008.
- [Huffman 95]
S. Huffman. Learning information extraction patterns from examples. *Workshop on new approaches to learning for natural language processing (IJCAI-95)*, pages 127–142, 1995.
- [Huggins & Nickerson 85]
A. W. F. Huggins and R. S. Nickerson. Speech quality evaluation using ‘phoneme-specific’ sentences. *Journal of the Acoustical Society of America*, 77:1896–1906, 1985.
- [Hughes & Bird 03]
B. Hughes and S. Bird. Grid-enabling natural language engineering by stealth. In *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 31–38, 2003.
- [Hull 93]
D. A. Hull. Using statistical testing in the evaluation of retrieval experiments. In *Research and Development in Information Retrieval*, pages 329–338, 1993.
- [Hull *et al.* 01]
D. Hull, S. Ait-Mokhatar, M. Chuat, A. Eisele, E. Gaussier, G. Grefenstette, P. Isabelle, C. Samuelsson, and F. Segond. Language technologies and patent search and classification. *World Patent Information*, 23:265–268, 2001.
- [Humphreys *et al.* 96]
K. Humphreys, R. Gaizauskas, H. Cunningham, and S. Azzam. CREOLE Module Specifications. <http://www.dcs.shef.ac.uk/nlp/gate/>, 1996.
- [Humphreys *et al.* 98]
K. Humphreys, R. Gaizauskas, S. Azzam, C. Huyck, B. Mitchell, H. Cunningham, and Y. Wilks. Description of the LaSIE system as used for MUC-7. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)*. http://www.itl.nist.gov/iaui/894.02/related_projects/muc/index.html, 1998.

- [Humphreys *et al.* 99]
K. Humphreys, R. Gaizauskas, M. Hepple, and M. Sanderson. The University of Sheffield TREC-8 Q&A System. In *In Proceedings of the 8th Text REtrieval Conference*, 1999.
- [Humphreys *et al.* 00a]
K. Humphreys, G. Demetriou, and R. Gaizauskas. Two applications of information extraction to biological science journal articles: Enzyme interactions and protein structures. In *Proc. of Pacific Symposium on Biocomputing (PSB-2000)*, Honolulu, Hawaii, 2000.
- [Humphreys *et al.* 00b]
K. Humphreys, R. Gaizauskas, and H. Cunningham. LaSIE Technical Specifications. Technical report, Department of Computer Science. University of Sheffield, 2000.
- [Hung *et al.* 08]
R. J. Hung, J. D. McKay, V. Gaborieau, P. Boffetta, and et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature*, 452:633–637, April 2008.
- [Hurlock & Wilson 11]
J. Hurlock and M. L. Wilson. Searching Twitter: Separating the Tweet from the Chaff. In *Proceedings of the Fifth AAAI Conference on Weblogs and Social Media (ICWSM)*, 2011.
- [Hutchins & Somers 92a]
W. J. Hutchins and H. L. Somers. *An Introduction to Machine Translation*. Academic Press, London, 1992.
- [Hutchins & Somers 92b]
W. Hutchins and H. Somers. *An Introduction to Machine Translation*. Academic Press, 1992.
- [Hutchins 86]
W. Hutchins. *Machine Translation: past, present, future*. Chichester : Ellis Horwood, 1986.
- [Hutto & Gilbert 14]
C. Hutto and E. Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth International AAAI Conference on Weblogs and Social Media*, 2014.
- [Huyck 94]
C. Huyck. Plink: An intelligent natural language parser (cse-tr-218-94). Technical report, Computer Science and Engineering Division, The University of Michigan, 1994.
- [Hwa 04]
R. Hwa. Sample Selection for Statistical Parsing. *Computational Linguistics*, 30(3):253–276, 2004.
- [Hwa *et al.* 02]
R. Hwa, P. Resnik, A. Weinberg, and O. Kolak. Evaluating translational correspondence using annotation projection. Technical report, University of Maryland, 2002.
- [Hwa *et al.* 05]
R. Hwa, P. Resnik, A. Weinberg, C. Cabezas, and O. Kolak. Bootstrapping parsers via syntactic projection across parallel texts. *Journal of Natural Language Engineering*, 11(03):311–325, 2005.
- [Hyvönen 83]
E. Hyvönen. Graph Grammar Approach to Natural Language Parsing and Understanding. In *Proceedings of the Eighth International Joint Conference on Artificial Intelligence (IJCAI'83)*, volume 2, pages 671–674. Morgan Kaufmann, California, 1983.
- [IBM 99a]
IBM. Building Object-Oriented Frameworks. <http://www.ibm.com/java/education/-oobuilding/index.html>, March 1999.

- [IBM 99b]
IBM. Leveraging Object-Oriented Frameworks. <http://www.ibm.com/java/education/-ooleveraging/index.html>, March 1999.
- [Ibrahim & Cummins 89]
M. Ibrahim and F. Cummins. TARO: An Interactive, Object-Oriented Tool for Building Natural Language Systems. In *IEEE International Workshop on Tools for Artificial Intelligence*, pages 108–113, Los Angeles, 1989.
- [Ide & Priest-Dorman 99]
N. Ide and G. Priest-Dorman. Corpus Encoding Standard. <http://www.cs.vassar.edu/CES/>, 1999.
- [Ide & Romary 02]
N. Ide and L. Romary. Standards for language resources. In *Proceedings of 3rd Language Resources and Evaluation Conference (LREC'2002)*, Gran Canaria, Spain, 2002.
- [Ide & Romary 04]
N. Ide and L. Romary. Standards for language resources. *Natural Language Engineering*, 10:211–225, 2004.
- [Ide & Suderman 05]
N. Ide and K. Suderman. Integrating linguistic resources: The american national corpus model. In *Proceedings of Human Language Technology Conference/Conference on Empirical Methods in Natural Language Processing HLT/EMNLP 2005*, Vancouver, B.C., Canada, 2005.
- [Ide & Suderman 12]
N. Ide and K. Suderman. Bridging the gaps: interoperability for language engineering architectures using GrAF. *Language Resources and Evaluation*, 46:1–15, mar 2012.
- [Ide & Veronis 94]
N. Ide and J. Veronis. Have we wasted our time? In *Proceedings of the International Workshop on the Future of the Dictionary*, Grenoble, 1994.
- [Ide & Veronis 98]
N. Ide and J. Veronis. Introduction to the special issue on word sense disambiguation: The state of the art. *Computational Linguistics*, 24(1):1–40, 1998.
- [Ide 98a]
N. Ide. Corpus Encoding Standard: SGML Guidelines for Encoding Linguistic Corpora. In *Proceedings of the First International Language Resources and Evaluation Conference*, pages 463–470, Granada, Spain, 1998.
- [Ide 98b]
N. Ide. Encoding Linguistic Corpora. In *Proceedings of the Sixth Workshop on Very Large Corpora*, pages 9–17, Montreal, 1998.
- [Ide et al. 00]
N. Ide, P. Bonhomme, and L. Romary. XCES: An XML-based Standard for Linguistic Corpora. In *Proceedings of the second International Conference on Language Resources and Evaluation (LREC 2000)*, 30 May – 2 Jun 2000, pages 825–830, Athens, Greece, 2000.
- [Ignat et al. 03]
C. Ignat, B. Pouliquen, A. Ribeiro, and R. Steinberger. Extending and Information Extraction Tool Set to Eastern-European Languages. In *Proceedings of Workshop on Information Extraction for Slavonic and other Central and Eastern European Languages (IESL'03)*, Borovets, Bulgaria, 2003.
- [Indurkha & Damerou 10]
N. Indurkha and F. J. Damerou. *Handbook of Natural Language Processing*, volume 2. CRC Press, 2010.

- [Ingwersen 96]
P. Ingwersen. Cognitive perspectives of information retrieval interaction: Elements of a cognitive ir theory. *Journal of Documentation*, 52(1):3–50, 1996.
- [INLG’96 96]
Proceedings of the Eighth International Workshop on Natural Language Generation, Herstmonceux, Sussex, UK, 1996.
- [Inouye & Kalita 11]
D. Inouye and J. K. Kalita. Comparing Twitter summarization algorithms for multiple post summaries. In *SocialCom/PASSAT*, pages 298–306, 2011.
- [Ireson & Ciravegna 05]
N. Ireson and F. Ciravegna. Pascal Challenge The Evaluation of Machine Learning for Information Extraction. In *Proceedings of Dagstuhl Seminar Machine Learning for the Semantic Web* (<http://www.smi.ucd.ie/Dagstuhl-MLSW/proceedings/>), 2005.
- [Ireson & Ciravegna 10]
N. Ireson and F. Ciravegna. Toponym resolution in social media. In *Proceedings of the 9th International Semantic Web Conference (ISWC)*, pages 370–385, 2010.
- [Irvine & Klementiev 10]
A. Irvine and A. Klementiev. Using Mechanical Turk to annotate lexicons for less commonly used languages. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, pages 108–113, 2010.
- [Isard *et al.* 98]
A. Isard, D. McKelvie, and H. Thompson. Towards a Minimal Standard for Dialogue Transcripts: A New Sgml Architecture for the HCRC Map Task Corpus. In *Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP ’98)*, Sydney, 1998.
- [Isbell *et al.* 02]
C. Isbell, B. Amento, G. Bell, and J. Helfman. Ishmail: Managing Massive Amounts of Mail. In *CSCW 2002 Workshop: Redesigning Email for the 21st Century*, 2002.
- [ISO 1087 90]
ISO 1087. International organisation for standardisation (ISO) - terminology - vocabulary, 1990.
- [ISO 08]
ISO. Language resource management—morpho-syntactic annotation framework (MAF). Standard ISO/DIS 24611, ISO TC37/SC4, December 2008.
- [ISO 10]
ISO. Language resource management—syntactic annotation framework (SynAF). Standard ISO/DIS 24615, ISO TC37/SC4, 2010.
- [ISO-704 87]
ISO-704. International organisation for standardisation (ISO) - principles and methods of terminology, 1987.
- [Isozaki & Kazawa 02]
H. Isozaki and H. Kazawa. Efficient Support Vector Classifiers for Named Entity Recognition. In *Proceedings of the 19th International Conference on Computational Linguistics (COLING’02)*, pages 390–396, Taipei, Taiwan, 2002.
- [Issar & Ward 93]
S. Issar and W. Ward. CMU’s Robust Spoken Language Understanding System. In *European Conference on Speech Communication and Technology (EUROSPEECH)*, volume 3, pages 2147–2149, Berlin, Germany, September 1993.

- [IST 02]
IST. *Guide for Proposers FP6/2002/IST/1 Integrated Projects*. Information Society Technologies, European Commission, 2002. B_PGIP_200201_EN.DOC.
- [ISW06]
5th International Semantic Web Conference (ISWC2006), 2006.
- [Ittycheriah *et al.* 03]
A. Ittycheriah, L. Lita, N. Khambhatla, N. Nikolov, S. Roukos, and M. Stys. Identifying and Tracking Entity Mentions in a Maximum Entropy Framework. In *Proceedings of HLT-NAACL 2003*, pages 40–42, Edmonton, Canada, 2003.
- [Iwayama *et al.* 05]
M. Iwayama, A. Fujii, and N. Kando. Overview of Classification Subtask at NTCIR-5 Patent Retrieval Task. In *Proceedings of the Fifth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-lingual Information Access, 6–9 December 2005*, pages 278–286, 2005.
- [Iwayama *et al.* 07]
M. Iwayama, A. Fujii, and N. Kando. Overview of Classification Subtask at NTCIR-6 Patent Retrieval Task. In *Proceedings of the Sixth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access*, 2007.
- [Iyer & Ostendorf 96]
R. Iyer and M. Ostendorf. Modelling long distance dependences in language: Topic mixtures vs. dynamic cache models. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 1, Philadelphia PA, 1996.
- [Jackson 75]
M. Jackson. *Principles of Program Design*. Academic Press, London, 1975.
- [Jacobs & Rau 90a]
P. S. Jacobs and L. F. Rau. Scisor: Extracting information from on-line news. *Communications of the ACM*, 33(11):88–97, 1990.
- [Jacobs & Rau 90b]
P. Jacobs and L. Rau. Scisor: Extracting information from on-line news. *Communications of the ACM*, 33(11):88–97, 1990.
- [Jacobs 92a]
P. S. Jacobs, editor. *Text-Based Intelligent Systems: Current Research and Practice in Information Extraction and Retrieval*. Lawrence Erlbaum, Hillsdale, NJ, 1992.
- [Jacobs 92b]
P. Jacobs, editor. *Text-Based Intelligent Systems: Current Research and Practice in Information Extraction and Retrieval*. Lawrence Erlbaum, Hillsdale, NJ, 1992.
- [Jacobs *et al.* 95]
P. Jacobs, G. Krupka, and L. Rau. Description of the SHOGUN System as used for MUC-5. In *Proceedings of the Fifth Message Understanding Conference (MUC-5)*, pages 109–120. Morgan Kaufmann, California, 1995.
- [Jacobson *et al.* 99]
I. Jacobson, G. Booch, and J. Rumbaugh. *The Unified Software Development Process*. Addison-Wesley, Reading, MA, 1999.

- [Jacoby *et al.* 74]
 J. Jacoby, D. E. Speller, and C. A. Kohn. Brand choice behavior as a function of information load. *Journal of Marketing Research*, 11(1):pp. 63–69, 1974.
- [Jacquemin & Royaute 94]
 C. Jacquemin and J. Royaute. Retrieving terms and their variants in a lexicalised unification-based framework. In *Proc. of 13th Annual International ACM-SIGIR Conference in Research and Development in Information Retrieval*, pages 132–141, Dublin, 1994.
- [Jacquemin 96]
 C. Jacquemin. What is the tree that we see through the window: A linguistic approach to windowing and term variation. *Information Processing and Management*, 32(4):445–458, 1996.
- [Jacquemin 97]
 C. Jacquemin. Recognition and acquisition: Two inter-related activities in corpus-based term extraction. *Terminology*, 4(2):245 ff., 1997.
- [Jagadesh & Qian 92]
 H. Jagadesh and X. Qian. Integrity maintenance in an object-oriented database. In *Proceedings of 18th International Conference on Very Large Databases*, pages 469–481, Vancouver, Canada, 1992.
- [Jain *et al.* 99]
 A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: A review. *ACM Computing Surveys*, 31(3):264–323, Sep 1999.
- [Janowicz *et al.* 13]
 K. Janowicz, A. Bröring, C. Stasch, S. Schade, T. Everding, and A. Llaves. A restful proxy and data model for linked sensor data. *International Journal of Digital Earth*, 6(3):233–254, 2013.
- [Janowicz *et al.* ss]
 K. Janowicz, S. Scheider, T. Pehle, and G. Hart. Geospatial semantics and linked spatiotemporal data – past, present, and future. *Semantic Web – Interoperability, Usability, Applicability*, In Press.
- [Jansen *et al.* 00]
 B. J. Jansen, A. Spink, and T. Saracevic. Real life, real users, and real needs: A study and analysis of user queries on the web. *Information Processing and Management*, 36(2):207–227, 2000.
- [Jansen *et al.* 09]
 B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury. Twitter power: Tweets as electronic word of mouth. *JASIST*, 60(11):2169–2188, 2009.
- [Janssen 90]
 S. Janssen. Automatic sense disambiguation with LDOCE: enriching syntactically analyzed corpora with semantic data. In J. Aarts and W. Meijs, editors, *Theory and Practice in Corpus Linguistics, Language and Computers: Studies in Practical Linguistics*, No.4, pages 105–135. Rodopi, Amsterdam, 1990.
- [Järvelin & Kekäläinen 02a]
 K. Järvelin and J. Kekäläinen. Cumulated gain-based evaluation of IR techniques. *ACM Trans. Inf. Syst.*, 20(4):422–446, October 2002.
- [Järvelin & Kekäläinen 02b]
 K. Järvelin and J. Kekäläinen. Cumulated gain-based evaluation of IR techniques. *ACM Trans. Inf. Syst.*, 20(4):422–446, October 2002.
- [Java *et al.* 07]
 A. Java, X. Song, T. Finin, and B. Tseng. Why we twitter: understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web*

mining and social network analysis, WebKDD/SNA-KDD '07, pages 56–65, New York, NY, USA, 2007. ACM.

[jbp]

JBPM User Guide. <http://docs.jboss.com/jbpm/v3/userguide>.

[Jelinek & Lafferty 91]

F. Jelinek and J. D. Lafferty. Computation of the probability of initial substring generation by stochastic context-free grammars. *Computational Linguistics*, 17:315–323, 1991.

[Jelinek 69]

F. Jelinek. Fast sequential decoding algorithm using a stack. *IBM Journal of Research and Development*, 13:675–685, 1969.

[Jelinek 76]

F. Jelinek. Continuous speech recognition by statistical methods. *Proceedings of the IEEE*, 64:532–556, 1976.

[Jelinek 85]

F. Jelinek. Self-organising language modelling for speech recognition. IBM Technical Disclosure Bulletin 28, IBM, 1985.

[Jelinek 90]

F. Jelinek. Self-organised language modelling for speech recognition. In A. Waibel and K. F. Lee, editors, *Readings in Speech Recognition*, pages 450–506. Morgan Kaufmann, California, 1990.

[Jelinek 91]

F. Jelinek. Up from trigrams! The struggle for improved language models. In *Proceedings European Conference on Speech Communication and Technology*, pages 1037–1039, 1991.

[Jelinek 97]

F. Jelinek. *Statistical Methods for Speech Recognition*. MIT Press, Cambridge, MA, 1997.

[Jelinek *et al.* 75]

F. Jelinek, L. R. Bahl, and R. L. Mercer. Design of a linguistic statistical decoder for the recognition of continuous speech. *IEEE Transactions on Information Theory*, 21:250–256, 1975.

[Jelinek *et al.* 90]

F. Jelinek, R. Mercer, and S. Roukos. Classifying words for improved statistical language models. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 621–624, Albuquerque, 1990.

[Jelinek *et al.* 91]

F. Jelinek, B. Merialdo, S. Roukos, and M. Strauss. A dynamic language model for speech recognition. In *Proceedings DARPA Workshop on Speech and Natural Language*, pages 293–295, February 1991.

[Jelinek *et al.* 92]

F. Jelinek, J. Lafferty, and R. Mercer. Basic Methods of Probabilistic Context Free Grammars. *Speech Recognition and Understanding. Recent Advances*, 75:345–360, 1992.

[Ji & Grishman 11]

H. Ji and R. Grishman. Knowledge base population: Successful approaches and challenges. In *Proc. of ACL'2011*, pages 1148–1158, 2011.

[Ji *et al.* 10a]

H. Ji, R. Grishman, H. T. Dang, K. Griffitt, and J. Ellis. Overview of the tac 2010 knowledge base population track. In *Proceedings of the Third Text Analysis Conference*, 2010.

- [Ji *et al.* 10b]
H. Ji, R. Grishman, H. T. Dang, K. Griffitt, and J. Ellis. Overview of the tac 2010 knowledge base population track. In *Proc. of the 3rd Text Analysis Conf.*, 2010.
- [Jiang & Stroulia 04]
Y. Jiang and E. Stroulia. Towards reengineering web sites to web-services providers. In *Proceedings of the Eighth European Conference on Software Maintenance and Reengineering (CSMR 2004)*, pages 296–305, 2004.
- [Jimenez *et al.* 09]
S. Jimenez, C. Becerra, A. Gelbukh, and F. Gonzalez. Generalized mongue-elkan method for approximate text string comparison. In *Proc. of CICLing*, 2009.
- [Jin *et al.* 06]
Y. Jin, R. T. McDonald, K. Lerman, M. A. Mandel, S. Carroll, M. Y. Liberman, F. C. Pereira, R. S. Winters, , and P. S. White. Automated recognition of malignancy mentions in biomedical literature. *BMC Bioinformatics*, 7:492–499, 2006.
- [Jing & McKeown 98a]
H. Jing and K. McKeown. Combining Multiple, Large-Scale Resources in a Reusable Lexicon for Natural Language Generation. In *Proceedings of the 36th ACL and the 17th COLING (ACL-COLING '98)*, pages 607–613, Montreal, 1998.
- [Jing & McKeown 98b]
H. Jing and K. McKeown. Combining multiple, large-scale resources in a reusable lexicon for natural language generation. In *COLING-ACL '98, Proceedings of the Conference*, pages 607 – 613, Montreal, Canada, August 1998.
- [Jing & McKeown 99]
H. Jing and K. McKeown. The decomposition of human-written summary sentences. In *Research and Development in Information Retrieval*, pages 129–136, 1999.
- [Jing & McKeown 00]
H. Jing and K. McKeown. Cut and Paste Based Text Summarization. In *Proceedings of the 1st Meeting of the North American Chapter of the Association for Computational Linguistics*, pages 178–185, Seattle, Washington, USA, April 29 – May 4, April 2000.
- [Jing 00]
H. Jing. Sentence Reduction for Automatic Text Summarization. In *Proceedings of the 6th Applied Natural Language Processing Conference*, pages 310–315, Seattle, Washington, USA, April 29 – May 4, April 2000.
- [Joachims 98]
T. Joachims. Text categorization with support vector machines: Learning with many relevant features. In C. Nédellec and C. Rouveirol, editors, *Proceedings of ECML-98, 10th European Conference on Machine Learning*, number 1398 in Lecture Notes in Computer Science, pages 137–142, Chemnitz, Germany, 1998. Springer Verlag, Heidelberg.
- [Joachims 99a]
T. Joachims. Making large-scale SVM learning practical. In B. Schölkopf, C. J. C. Burges, and A. J. Smola, editors, *Advances in Kernel Methods – Support Vector Learning*, pages 169–184, 1999.
- [Joachims 99b]
T. Joachims. Transductive inference for text classification using support vector machines. In *Proceedings of the 16th International Conference on Machine Learning (ICML-99)*, 1999.
- [Joachims 02]
T. Joachims. Optimizing search engines using clickthrough data. In *KDD [DBL02]*, pages 133–142.

- [Joachims 06]
T. Joachims. Training linear svms in linear time. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '06, pages 217–226, New York, NY, USA, 2006. ACM.
- [Joachims *et al.* 97]
T. Joachims, D. Freitag, and T. Mitchell. WebWatcher: A tour guide for the web. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'97)*, 1997.
- [Johansson 86]
S. Johansson. The tagged {LOB} corpus: User's manual, 1986.
- [Johansson *et al.* 09a]
M. Johansson, Y. Li, J. Wakefield, M. Greenwood, T. Heitz, I. Roberts, H. Cunningham, P. Brennan, A. Roberts, and J. Mckay. Using prior information attained from the literature to improve ranking in genome-wide association studies. In *The American Society of Human Genetics 59th Annual Meeting, October 20–24, 2009*, Honolulu, Hawaii, USA, October 2009.
- [Johansson *et al.* 09b]
M. Johansson, Y. Li, J. Wakefield, M. Greenwood, T. Heitz, I. Roberts, H. Cunningham, P. Brennan, A. Roberts, and J. Mckay. Using prior information attained from the literature to improve ranking in genome-wide association studies. In *International Genetic Epidemiology Society (IGES) 18th Annual Conference*, Kahuku, Hawaii, USA, October 2009.
- [Johansson *et al.* 12a]
M. Johansson, A. Roberts, D. Chen, Y. Li, M. Delahaye, N. Aswani, M. A. Greenwood, S. Benhamou, P. Lagiou, I. Holcátová, F. Merletti, L. Richiardi, K. Kjaerheim, A. Agudo, X. Castellsagué, G. J. Macfarlane, T. V. Macfarlane, L. Barzan, C. Canova, N. S. Thakker, D. I. Conway, A. Znaor, C. M. Healy, W. Ahrens, D. Zaridze, N. Szeszenia-Dabrowska, J. Lissowska, E. Fabianova, D. Mates, V. Bencko, L. Foretova, V. Janout, M. P. Curado, S. Koifman, A. Menezes, V. Wünsch-Filho, J. E. Neto, P. Boffetta, S. Franceschi, R. Herrero, L. F. Garrote, R. Talamini, S. Boccia, D. Zelenika, M. Lathrop, H. Cunningham, P. Brennan, J. Wakefield, and J. D. Mckay. Incorporation of Prior Information From the Medical Literature in GWAS of Oral Cancer Identifies Novel Susceptibility Variant on Chromosome 4 – The AdAPT Method. *PLoS One*, 2012.
- [Johansson *et al.* 12b]
M. Johansson, A. Roberts, D. Chen, Y. Li, M. Delahaye-Sourdeix, N. Aswani, M. A. Greenwood, S. Benhamou, P. Lagiou, I. Holcátová, L. Richiardi, K. Kjaerheim, A. Agudo, X. Castellsagué, T. V. M. andLuigi Barzan, C. Canova, N. S. Thakker, D. I. Conway, A. Znaor, C. M. Healy, W. Ahrens, D. Zaridze, N. Szeszenia-Dabrowska, J. Lissowska, E. Fabianova, I. N. Mates, V. Bencko, L. Foretova, V. J. andMaria Paula Curado, S. Koifman, A. Menezes, V. Wünsch-Filho, J. E. Neto, P. Boffetta, S. Franceschi, R. Herrero, L. F. Garrote, R. Talamini, S. Boccia, P. Galan, L. Vatten, P. Thomson, D. Zelenika, M. Lathrop, G. Byrnes, H. Cunningham, P. Brennan, J. Wakefield, and J. D. Mckay. Using prior information from the medical literature in gwas of oral cancer identifies novel susceptibility variant on chromosome 4 – the adapt method. *PLoS ONE*, 7(5):e36888, 05 2012.
- [Johnson & Lindenstrauss 84]
W. B. Johnson and J. Lindenstrauss. Extensions to lipschitz mapping into hilbert space. *Contemporary Mathematics*, 26, 1984.
- [Johnson & Rosenberg 95]
J. Johnson and R. Rosenberg. A Data Management Strategy for Transportable Natural Language Interfaces. *International Journal of Intelligent Systems*, 10(9):771–808, 1995.
- [Johnson 97]
R. Johnson. Frameworks Home Page. <http://st-www.cs.uiuc.edu/users/johnson/frameworks.html>, 1997.

- [Jonassen 82]
D. Jonassen, editor. *The Technology of Text*. Educational Technology Publications, 1982.
- [Jones & Mewhort 07]
M. N. Jones and D. J. K. Mewhort. Representing word meaning and order information in a composite holographic lexicon. *Psychological Review*, 114:1–37, 2007.
- [Jones & Paice 92]
P. A. Jones and C. D. Paice. A 'select and generate' approach to automatic abstracting. In A. McEnry and C. Paice, editors, *Proceedings of the 14th British Computer Society Information Retrieval Colloquium*, pages 151–154. Springer Verlag, 1992.
- [Jones 72]
K. S. Jones. A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 28(1):11–21, 1972.
- [Jones 90]
C. Jones. *Systematic Software development using VDM*. Prentice Hall, New York, 1990.
- [Jones 96]
D. Jones. *Analogical Natural Language Processing*. Studies in Computational Linguistics. UCL Press, 1996.
- [Jones 05]
R. Jones. *Learning to Extract Entities from Labeled and Unlabeled Text*. Unpublished PhD thesis, School of Computer Science, Carnegie Mellon University, 2005.
- [Jones et al. 95]
G. Jones, H. Lloyd-Thomas, and J. Wright. Lattice parsing and application of integrated language models for speech recognition. In *Proceedings European Conference on Speech Communication and Technology*, pages 1789–1792, Madrid, 1995.
- [Jonquet et al. 09]
C. Jonquet, N. H. Shah, and M. A. Musen. The open biomedical annotator. *Summit on Translational Bioinformatics*, pages 56–60, 2009.
- [Jönsson et al. 04]
A. Jönsson, F. Andén, L. Degerstedt, A. Flycht-Eriksson, M. Merkel, and S. Norberg. Experiences from combining dialogue system development with information extraction techniques. In M. T. Maybury, editor, *New directions in Question Answering*, pages 153–164. AAAI/MIT Press, 2004.
- [Jordan et al. 96]
P. Jordan, B. Thomas, B. Weerdmeester, and A. McClelland, editors. *Usability Evaluation in Industry*. Taylor and Francis, 1996.
- [Joshi & Srinivas 94]
A. Joshi and B. Srinivas. Disambiguation of super parts of speech (or supertags): Almost parsing. In *Proc. of 15th International Conference on Computational Linguistics (COLING)*, Kyoto, Japan, 1994.
- [Joshi 87]
A. K. Joshi. The relevance of tree adjoining grammar to generation. In G. Kempen, editor, *Natural Language Generation: Recent Advances in Artificial Intelligence, Psychology, and Linguistics*. Kluwer Academic Publishers, Boston/Dordrecht, 1987. Paper presented at the Third International Workshop on Natural Language Generation, August 1986, Nijmegen, The Netherlands.
- [Joubert et al. 94]
M. Joubert, M. Fieschi, and J.-J. Robert. A conceptual model for retrieval with UMLS. In *Proc. of JAMIA '94*, 1994.

- [Jouis & Elhadi 97]
C. Jouis and W. Elhadi. Aupelf project: Term and semantic relation extraction tools evaluation paradigm. In *Proc. of SALT Workshop on Evaluation in Speech and Language Technology*, Sheffield, UK, 1997.
- [Jouis 94]
C. Jouis. Contextual approach: SEEK, a linguistic and computational tool for use in knowledge acquisition. In *Proc. of 1st European Conference in Cognitive Science in Industry*, 1994.
- [Juang & Rabiner 85]
B. H. Juang and L. R. Rabiner. Mixture autoregressive hidden Markov models for speech signals. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33:1404–1412, 1985.
- [Judd & Kalita 13]
J. Judd and J. Kalita. Better twitter summaries? In L. Vanderwende, H. D. III, and K. Kirchhoff, editors, *Human Language Technologies: Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, June 9-14, 2013, Westin Peachtree Plaza Hotel, Atlanta, Georgia, USA*, pages 445–449. The Association for Computational Linguistics, 2013.
- [Jurafsky & Martin 00]
D. Jurafsky and J. H. Martin. *Speech and Language Processing*. Prentice Hall, Upper Saddle River, NJ 07458, 2000.
- [Jurafsky *et al.* 94]
D. Jurafsky, C. Wooters, G. Tajchman, J. Segal, A. Stolcke, and N. Morgan. Integrating experimental models of syntax, phonology, and accent/dialect in a speech recognizer. In *Proceedings of the AAAI Workshop on Integration of Natural Language and Speech Processing*, pages 107–115, Seattle, 1994.
- [Jurafsky *et al.* 95]
D. Jurafsky, C. Wooters, J. Segal, A. Stolcke, E. Fosler, and G. T. and N. Morgan. Using a context-free grammar as a language model for speech recognition. In *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing*, pages 189–192, Detroit MI, 1995.
- [Justeson & Katz 95]
J. Justeson and S. Katz. Technical terminology: some linguistic properties and an algorithm for identification in text. *Natural Language Engineering*, 1:9–27, 1995.
- [Kagdi *et al.* 07]
H. H. Kagdi, M. L. Collard, and J. I. Maletic. A survey and taxonomy of approaches for mining software repositories in the context of software evolution. *Journal of Software Maintenance*, 19(2):77–131, 2007.
- [Kageura & Umino 96]
K. Kageura and B. Umino. Methods of automatic term recognition. *Terminology*, 3(2):259–289, 1996.
- [Kahane *et al.* 15]
B. Kahane, A. Mogoutov, J. P. Cointet, L. Villard, and P. Laredo. A dynamic query to delineate emergent science and technology: the case of nano science and technology. *Content and technical structure of the Nano S&T Dynamics Infrastructure*, pages 47–70, 2015.
- [Kaljurand & Fuchs 06]
K. Kaljurand and N. E. Fuchs. Bidirectional mapping between OWL DL and Attempto Controlled English. In *Fourth Workshop on Principles and Practice of Semantic Web Reasoning* [PPS06].
- [Kaljurand & Fuchs 07]
K. Kaljurand and N. E. Fuchs. Verbalizing OWL in Attempto Controlled English. In *OWLED*, volume 258, 2007.

- [Kaljurand 06a]
K. Kaljurand. From ACE to OWL and from OWL to ACE. In Bry and Schwertel [Bry & Schwertel 06].
- [Kaljurand 06b]
K. Kaljurand. Writing OWL ontologies in ACE. Technical report, University of Zurich, August 2006.
- [Kalyanpur *et al.* 05]
A. Kalyanpur, B. Parsia, E. Sirin, B. Cuenca-Grau, and J. Hendler. Swoop: A 'Web' Ontology Editing Browser. *Journal of Web Semantics*, 4(2), 2005.
- [Kameyama 97]
M. Kameyama. Information Extraction across Linguistic Boundaries. In *AAAI Spring Symposium on Cross-Language Text and Speech Processing*, Stanford University, 1997.
- [Kammerer *et al.* 09]
Y. Kammerer, R. Nairn, P. Pirolli, and E. H. Chi. Signpost from the masses: Learning effects in an exploratory social tag search browser. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 625–634, New York, NY, USA, 2009. ACM.
- [Kamp 81]
H. Kamp. A theory of truth and semantic representation. In *Formal Methods in The Study of Language*. J. Groenendijk, Janssen, T, and Stokhof, M, 1981.
- [Kamprath *et al.* 98]
C. Kamprath, E. Adolphson, T. Mitamura, and E. Nyberg. Controlled Language for Multilingual Document Production: Experience with Caterpillar Technical English. In *Second International Workshop on Controlled Language Applications (CLAW '98)*, Pittsburgh, 1998.
- [Kan & McKeown 02]
M.-Y. Kan and K. R. McKeown. Boosted Wrapper Induction. In *Proceedings of INLG*, pages 1–8, Harriman, New York, USA, 2002.
- [Kano *et al.* 09]
Y. Kano, L. McCrohon, S. Ananiadou, and J. Tsujii. Integrated NLP Evaluation System for Plug-gable Evaluation Metrics with Extensive Interoperable Toolkit. In *Software engineering, testing, and quality assurance for natural language processing workshop (SETQA-NLP), NAACL-HLT*, pages 22–30, 2009.
- [Kanthak *et al.* 00]
S. Kanthak, A. Sixtus, S. Molau, R. Schlüter, and H. Ney. Fast search for large vocabulary speech recognition. *VerbMobil: Foundations of Speech-to-Speech Translation*, pages 63–78, November 2000.
- [Kapanipathi *et al.* 11]
P. Kapanipathi, F. Orlandi, A. Sheth, and A. Passant. Personalized Filtering of the Twitter Stream. In *2nd workshop on Semantic Personalized Information Management at ISWC 2011*, 2011.
- [Kaplan & Bresnan 82]
R. Kaplan and J. Bresnan. Lexical functional grammar: A formal system for grammatical representation. In J. Bresnan, editor, *The Mental Representation of Grammatical Relations*, pages 173–281. MIT Press, Cambridge MA, 1982.
- [Karamanis *et al.* 08]
N. Karamanis, R. Seal, I. Lewin, P. McQuilton, A. Vlachos, C. Gasperin, R. Drysdale, and T. Briscoe. Natural language processing in aid of flybase curators. *BMC Bioinformatics*, 9(1):193, 2008.

- [Karanastasi *et al.* 07]
 A. Karanastasi, A. Zotos, and S. Christodoulakis. The OntoNL Framework for Natural Language Interface Generation and a Domain-Specific Application. In *Digital Libraries: Research and Development*, pages 228–237. Springer Berlin / Heidelberg, 2007.
- [Karkaletsis 95]
 E. Karkaletsis. *Terminological Knowledge Bases and their Exploitation in Multilingual Software Applications*. Unpublished PhD thesis, Department of Informatics, University of Athens, 1995. In Greek.
- [Karkaletsis *et al.* 92]
 E. Karkaletsis, N. Fecos, and C. D. Spyropoulos. System for exploiting knowledge from natural language texts. Technical report of the Research Programme PABE-88 BE 224, January 1992.
- [Karkaletsis *et al.* 94]
 E. Karkaletsis, C. Spyropoulos, and G. Vouros. A knowledge based approach for organising terminological data and generating messages in software applications. In *Proceedings of ‘Language Engineering on the Information Highway’ Workshop*. Santorini, September 1994.
- [Karkaletsis *et al.* 95a]
 E. Karkaletsis, C. Spyropoulos, and G. Vouros. The use of terminological knowledge bases in software localisation. In *Lecture Notes in AI (LNAI)*, number 898. Springer-Verlag, March 1995.
- [Karkaletsis *et al.* 95b]
 E. Karkaletsis, C. Spyropoulos, G. Vouros, and C. Halatsis. Organisation and exploitation of terminological knowledge in software localisation. *TermNet News (TNN) The Journal for International Cooperation in Terminology*, March 1995.
- [Karkaletsis *et al.* 99]
 V. Karkaletsis, C. Spyropoulos, and G. Petasis. Named Entity Recognition from Greek texts: the GIE Project. In S.Tzafestas, editor, *Advances in Intelligent Systems: Concepts, Tools and Applications*, pages 131–142. Kluwer Academic Publishers, 1999.
- [Karlgrén & Sahlgren 01]
 J. Karlgrén and M. Sahlgren. From words to understanding. In Y. Uesaka, P. Kanerva, and H. Asoh, editors, *Foundations of Real-World Intelligence*, pages 294–308. 2001.
- [Karlgrén 94]
 D. Karlgrén, J. and Cutting. Recognising text genres with simple metrics using discriminant analysis. In *Proceedings of COLING 94*, Kyoto, Japan, 1994.
- [Karlgrén 98]
 J. Karlgrén. Stylistic experiments for information retrieval. In T. Strzalkowski, editor, *Natural Language Information Retrieval*. Kluwer, 1998.
- [Karlsson & Karttunen 96]
 F. Karlsson and L. Karttunen. Sub-sentential processing. In R. Cole, editor, *Survey of the State of the Art in Human Language Technology*. Cambridge University Press, 1996.
- [Karov & Edelman 96]
 Y. Karov and S. Edelman. Learning similarity-based word sense disambiguation from sparse data. In E. Ejerhed and I. Dagan, editors, *Proceedings of the Fourth Workshop on Very Large Corpora*, Copenhagen, 1996.
- [Karov & Edelman 98]
 Y. Karov and S. Edelman. Similarity-based word sense disambiguation. *Computational Linguistics*, 24(1):41–60, 1998.

- [Karttunen 84]
L. Karttunen. Features and values. *Proceedings of the Tenth International Conference on Computational Linguistics*, 1984.
- [Kaski 98]
S. Kaski. Dimensionality reduction by random mapping. In *Proceedings of International Joint Conference on Neural Networks (volume 1)*, pages 413–418, 1998.
- [Katz & Lin 03]
B. Katz and J. Lin. Selectively using relations to improve precision in question answering. In *Proceedings of the EACL-2003 Workshop on Natural Language Processing for Question Answering, April 2003*, 2003.
- [Katz 87]
S. Katz. Estimation of probabilities from sparse data for the language model component of a speech recognizer. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 35(3):400–401, 1987.
- [Kaufmann & Bernstein 07]
E. Kaufmann and A. Bernstein. How Useful are Natural Language Interfaces to the Semantic Web for Casual End-users? In *Proceedings of the Forth European Semantic Web Conference (ESWC 2007)*, Innsbruck, Austria, June 2007.
- [Kaufmann 06]
E. Kaufmann. Talking to the semantic web—query interfaces to ontologies for the casual user. In *5th International Semantic Web Conference (ISWC)*, 2006.
- [Kaufmann et al. 06]
E. Kaufmann, A. Bernstein, and R. Zumstein. Querix: A Natural Language Interface to Query Ontologies Based on Clarification Dialogs. In *5th International Semantic Web Conference (ISWC 2006)*, pages 980–981. Springer, November 2006.
- [Kay 79]
M. Kay. Functional grammar. *Proceedings of the Fifth Annual Meeting of the Berkeley Linguistic Society*, pages 142–158, 1979.
- [Kay 97a]
M. Kay. It’s Still the Proper Place. *Machine Translation*, 12:3–23, 1997.
- [Kay 97b]
M. Kay. The Proper Place of Men and Machines in Language Translation. *Machine Translation*, 12:3–23, 1997. Originally appeared as a Xerox PARC Working Paper in 1980.
- [Kay 97c]
M. Kay. The Proper Place of Men and Machines in Language Translation. *Machine Translation*, 12:3–23, 1997. Originally appeared as a Xerox PARC Working Paper in 1980.
- [Kay et al. 87]
S. R. Kay, A. Fiszbein, and L. A. Opler. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia bulletin*, 13(2):261–276, 1987. Cited by 8221.
- [Kay et al. 94]
M. Kay, J. Gawron, and P. Norvig. *Verbmobil, A Translation System for Face-to-Face Dialog*. CSLI, Stanford, CA, 1994.
- [Kazai et al. 04]
G. Kazai, M. Lalmas, and A. Vries. The Overlapping problem in Content-Oriented XML Retrieval Evaluation. In *Proceedings of the 27th International conference on Research and development in information retrieval*, pages 72–79, Sheffield, UK, 2004.

- [Kazi & Ravin 00]
Z. Kazi and Y. Ravin. Who’s who? Identifying concepts and entities across multiple documents. In *Proceedings of the 33rd Hawaii International Conference on System Sciences*, volume 3, Hawaii, USA, Jan 2000.
- [Kedzie *et al.* 15]
C. Kedzie, K. McKeown, and F. Diaz. Predicting salient updates for disaster summarization. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics*, volume 1, pages 1608–1617. Association for Computational Linguistics, 2015.
- [Keijola 99]
M. Keijola. BRIEFS – Gaining Information of Value in Dynamical Business Environments. <http://www.tuta.hut.fi/briefs>, 1999.
- [Keil 79]
F. C. Keil. *Semantic and conceptual development: an ontological perspective*. Harvard University Press, Cambridge, MA, USA, 1979.
- [Keller 84]
W. Keller. A program for analysing questions. M.a. thesis, cognitive studies programme, University of Sussex, 1984.
- [Keller 86]
W. Keller. Nested cooper storage: The proper treatment of quantification in ordinary noun phrases. Unpublished Paper, Cognitive Studies Programme, 1986.
- [Kemps-Snijders *et al.* 09]
M. Kemps-Snijders, N. Bel, and P. Wittenburg. Proposal for a CLARIN European demonstrator. Technical report, CLARIN Consortium, September 2009.
- [Kendall *et al.* 02]
E. Kendall, M. Dutra, and D. McGuinness. Towards a Commercial Ontology Development Environment. In *International Semantic Web Conference (ISWC), Late Breaking Topics*, 2002.
- [Kennedy & Boguraev 96]
C. Kennedy and B. Boguraev. Anaphora for everyone: Pronominal anaphora resolution without a parser. In *Proceedings of the 16th International Conference on Computational Linguistics, COLING-96*, pages 113–118, Copenhagen, 1996.
- [Kenny *et al.* 93]
P. Kenny, R. Hollan, V. N. Gupta, M. Lennig, P. Mermelstein, and D. O’Shaughnessy. A*-admissible heuristics for rapid lexical access. 1:59–58, 1993.
- [Kergl *et al.* 14]
D. Kergl, R. Roedler, and S. Seeber. On the endogenesis of Twitter’s Spritzer and Gardenhose sample streams. In *Advances in Social Networks Analysis and Mining (ASONAM), 2014 IEEE/ACM International Conference on*, pages 357–364. IEEE, 2014.
- [Kershaw *et al.* 96a]
D. J. Kershaw, M. M. Hochberg, and A. J. Robinson. Context-dependent classes in a hybrid recurrent network-hmm speech recognition system. In *Advances in Neural Information Processing Systems*, volume 8. MIT Press, 1996.
- [Kershaw *et al.* 96b]
D. J. Kershaw, A. J. Robinson, and S. J. Renals. The 1995 ABBOT hybrid connectionist-hmm large-vocabulary recognition system. In *Proceedings of ARPA Spoken Language Systems Technology Workshop*, 1996.

- [Kessler *et al.* 97]
 B. Kessler, G. Nunberg, and H. Schütze. Automatic detection of text genre. In *Proceedings of 35th Annual Meeting of the Association for Computational Linguistics*, Madrid, Spain, 1997.
- [Kessler *et al.* 10]
 J. S. Kessler, M. Eckert, L. Clark, and N. Nicolov. The 2010 icwsm jdpa sentiment corpus for the automotive domain. In *4th International AAAI Conference on Weblogs and Social Media Data Workshop Challenge (ICWSM-DWC 2010)*, 2010.
- [Keysers *et al.* 02]
 D. Keysers, F. J. Och, and H. Ney. Efficient Maximum Entropy Training for Statistical Object Recognition. In *Informatiktage 2002 der Gesellschaft für Informatik*, pages 342–345, Bad Schussenried, Germany, November 2002.
- [Khabiri *et al.* 11]
 E. Khabiri, J. Caverlee, and C.-F. Hsu. Summarizing user-contributed comments. In *International AAAI Conference on Weblogs and Social Media*, 2011.
- [Khanna *et al.* 10]
 S. Khanna, A. Ratan, J. Davis, and W. Thies. Evaluating and improving the usability of Mechanical Turk for low-income workers in India. In *Proceedings of the first ACM symposium on computing for development*. ACM, 2010.
- [Khosravi 99]
 H. Khosravi. *Extracting Pragmatic Content From Email*. Unpublished PhD thesis, University of Sheffield, UK, 1999.
- [Kiefer *et al.* 07]
 C. Kiefer, A. Bernstein, and J. Tappolet. Analyzing software with isparql. In *Proceedings of the 3rd ESWC International Workshop on Semantic Web Enabled Software Engineering (SWESE)*. Springer, June 2007.
- [Kieniewicz & Wallis 12]
 J. Kieniewicz and M. Wallis. User requirements. Technical Report <http://gate.ac.uk/projects/envilod/EnviLOD-WP2-User-Requirements.pdf>, EnviLOD project deliverable, 2012.
- [Kieniewicz *et al.* 11]
 J. Kieniewicz, A. Sudlow, and E. Newbold. Coordinating improved environmental information access and discovery: Innovations in sharing environmental observations and information. In W. Pillman, S. Schade, and P. Smits, editors, *Proceedings of the 25th International EnviroInfo Conference*, 2011.
- [Kilgarriff & Evans 95]
 A. Kilgarriff and R. Evans. Standards and How to Do Lexical Research. In *Proceedings of the Language Engineering Conference, London*, 1995.
- [Kilgarriff 92]
 A. Kilgarriff. *Polysemy*. Unpublished PhD thesis, University of Sussex, 1992.
- [Kilgarriff 93a]
 A. Kilgarriff. Dictionary word sense distinctions: An enquiry into their nature. *Computers and the Humanities*, pages 356–387, 1993.
- [Kilgarriff 93b]
 A. Kilgarriff. Dictionary word sense distinctions: An enquiry into their nature. *Computers and the Humanities*, 26:365–387, 1993.

- [Kilgarriff 97a]
A. Kilgarriff. Evaluating word sense disambiguation programs: progress report. In *Proc. of SALT Workshop in Speech and Language Technology*, pages 114–120, Sheffield University, UK, 1997.
- [Kilgarriff 97b]
A. Kilgarriff. Foreground and background lexicons and word sense disambiguation for information extraction. In *Proc. of International Workshop on Lexically Driven Information Extraction*, pages 51–62, Frascati, Italy, 1997.
- [Kilgarriff 97c]
A. Kilgarriff. "I don't believe in word senses". *Computing and the Humanities*, 31(2):91–113, 1997.
- [Kim & Hovy 05]
S. Kim and E. Hovy. Identifying Opinion Holders for Question Answering in Opinion Texts. Proceedings of AAAI-05 workshop on Question Answering in Restricted Domains. Pittsburgh, Pennsylvania, 2005.
- [Kim & Hovy 06]
S. Kim and E. Hovy. Extracting Opinions Expressed in Online News Media Text with Opinion Holders and Topics. In *Proceedings of the Workshop on Sentiment and Subjectivity in Text at the joint COLING-ACL conference*, 2006.
- [Kim & Moldovan 95a]
J. Kim and D. Moldovan. Acquisition of linguistic patterns for knowledge-based information extraction. *IEEE Transactions on Knowledge and Data Engineering*, 7(5):713–724, 1995.
- [Kim & Moldovan 95b]
J.-T. Kim and D. Moldovan. Acquisition of linguistic patterns for knowledge-based information extraction. *IEEE Transactions on Knowledge and Data Engineering.*, 7(5), 1995.
- [Kim & Shim 11]
Y. Kim and K. Shim. Twitobi: A recommendation system for twitter using probabilistic modeling. In *Proceedings of the 2011 IEEE 11th International Conference on Data Mining, ICDM '11*, pages 340–349, Washington, DC, USA, 2011. IEEE Computer Society.
- [Kim & Yu 11]
D. Kim and H. Yu. Figure text extraction in biomedical literature. *PLoS ONE*, 6(1):e15338, 01 2011.
- [Kim & Zhai 09]
H. D. Kim and C. Zhai. Generating comparative summaries of contradictory opinions in text. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management, CIKM '09*, pages 385–394, 2009.
- [Kim *et al.* 02]
J. Kim, I. Kang, and K. Choi. Unsupervised named entity classification models and their ensembles. In *Proceedings of the 19th International Conference on Computational Linguistics (COLING'02)*, Taipei, Taiwan, 2002.
- [Kincaid 10]
J. Kincaid. EdgeRank: The secret sauce that makes Facebook's news feed tick, April 2010.
- [Kinderen & Gordijn 08]
S. Kinderen and J. Gordijn. e3service: an ontological approach for deriving multi-supplier it-service bundles from consumer needs. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS-41)*, pages 318–318. IEEE Computer Society, January 2008.

- [King 03]
M. King. Living up to standards. In *Proceedings of the EACL 2003 Workshop on Evaluation Initiatives in Natural Language Processing*, Budapest, Hungary, 2003.
- [Kingston 94]
J. Kingston. Linking knowledge acquisition with CommonKADS knowledge representation. Technical Report AIAI-TR-156, University of Edinburgh, Edinburgh, UK, 1994.
- [Kintsch 98]
W. Kintsch. *Comprehension: A paradigm for cognition*. Cambridge University Press, New York, 1998.
- [Kirkpatrick *et al.* 83]
S. Kirkpatrick, C. Gelatt, and M. Vecchi. Optimisation by simulated annealing. *Science*, 220(4598):671–680, 1983.
- [Kirkpatrick *et al.* 10]
B. Kirkpatrick, G. P. Strauss, L. Nguyen, B. A. Fischer, D. G. Daniel, A. Cienfuegos, and S. R. Marder. The brief negative symptom scale: Psychometric properties. *Schizophrenia Bulletin*, 37(2):300–305, June 2010. Cited by 0000.
- [Kiryakov 03]
A. Kiryakov. Ontology and Reasoning in MUMIS: Towards the Semantic Web. Technical Report CS-03-03, Department of Computer Science, University of Sheffield, 2003. <http://gate.ac.uk/gate/doc/papers.html>.
- [Kiryakov 06]
A. Kiryakov. OWLIM: balancing between scalable repository and light-weight reasoner. In *Proceedings of the 15th International World Wide Web Conference (WWW2006), 23–26 May 2010*, Edinburgh, Scotland, 2006.
- [Kiryakov *et al.* 01]
A. Kiryakov, K. Simov, and M. Dimitrov. OntoMap – the Guide to the Upper-Level. In *Proceedings of the International Semantic Web Working Symposium (SWWS)*, Stanford University, California, USA, 2001.
- [Kiryakov *et al.* 04]
A. Kiryakov, B. Popov, D. Ognyanoff, D. Manov, A. Kirilov, and M. Goranov. Semantic annotation, indexing and retrieval. *Journal of Web Semantics*, 1(2):671–680, 2004.
- [Kiryakov *et al.* 09]
A. Kiryakov, D. Ognyanoff, R. Velkov, Z. Tashev, and I. Peikov. Ldsr: Materialized reason-able view to the web of linked data. In *OWL: Experiences and Directions workshop (OWLED 2009)*, 2009.
- [Kiss & Strunk 06]
T. Kiss and J. Strunk. Unsupervised multilingual sentence boundary detection. *Computational Linguistics*, 32(4):485–525, 2006.
- [Kittredge & Lehrberger 82a]
R. Kittredge and J. Lehrberger, editors. *Sublanguage: Studies of Language in Restricted Semantic Domains*. Walter de Gruyter, Berlin, 1982.
- [Kittredge & Lehrberger 82b]
R. Kittredge and J. Lehrberger, editors. *Sublanguage: Studies of Language in Restricted Semantic Domains*. Walter de Gruyter, 1982.
- [Kittur *et al.* 08]
A. Kittur, E. H. Chi, and B. Suh. Crowdsourcing user studies with mechanical turk. In *Proceedings of the 26th Conference on Human Factors in Computing Systems*, pages 453–456, 2008.

- [Kittur *et al.* 11]
 A. Kittur, B. Smus, S. Khamkar, and R. E. Kraut. Crowdforge: crowdsourcing complex work. In *Proceedings of the 24th ACM Symposium on User Interface Software and Technology*, pages 43–52, 2011.
- [Kiyono & Tsujii 94]
 M. Kiyono and J.-I. Tsujii. Hypothesis selection in grammar acquisition. In *Proc. of 15th International Conference on Computational Linguistics (COLING)*, Kyoto, Japan, 1994.
- [Klampanos *et al.* 10]
 I. A. Klampanos, H. Wu, T. Roelleke, and H. Azzam. Logic-based retrieval: Technology for content-oriented and analytical querying of patent data. In H. Cunningham, A. Hanbury, and S. Rüger, editors, *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference)*. LNCS volume number: 6107, Lecture Notes in Computer Science, Vienna, Austria, May 2010. Springer.
- [Klein & Manning 03]
 D. Klein and C. Manning. Accurate unlexicalized parsing. In *Proceedings of the 41st Meeting of the Association for Computational Linguistics*, 2003.
- [Klein & Noy 03]
 M. Klein and N. F. Noy. A Component-Based Framework for Ontology Evolution. In *Workshop on Ontologies and Distributed Systems, IJCAI*, 2003.
- [Klein *et al.* 02]
 M. Klein, D. Fensel, A. Kiryakov, and D. Ognyanov. Ontology Versioning and Change Detection on the Web. In *13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02)*, pages 197–212, Siguenza, Spain, 2002.
- [Kleinberg 98]
 J. Kleinberg. Authoritative sources in a hyperlinked environment. In *Proc. ACM-SIAM Symposium on Discrete Algorithms*, 1998.
- [Klovstad & Mondshein 74]
 J. W. Klovstad and L. F. Mondshein. The CASPERS linguistic analysis system. In *IEEE Symposium on Speech Recognition*, pages 234–240, Carnegie-Mellon University, 1974.
- [Klyne & Carrol 04]
 G. Klyne and J. J. Carrol. Resource description framework (rdf): Concepts and abstract syntax. W3C recommendation – 10 feb 2004, W3C, 2004. Available at <http://www.w3.org/TR/rdf-concepts/>.
- [Klyne & Carroll 04]
 G. Klyne and J. Carroll. Resource description framework (RDF): Concepts and abstract syntax. W3C recommendation, W3C, 2004. Available at <http://www.w3.org/TR/rdf-concepts/>.
- [Knapp 97]
 P. Knapp. *Virtual Grammar: Writing as Affect/Effect*. Unpublished PhD thesis, University of Technology, Sydney, <http://www.arch.apana.org.au/p.knapp/chapter4.pdf>, 1997.
- [Knight & Marcu 00]
 K. Knight and D. Marcu. Statistics-based summarization – step one: Sentence compression. In *AAAI/IAAI*, pages 703–710, Austin, Texas, 2000.
- [Knott *et al.* 96]
 A. Knott, C. Mellish, J. Oberlander, and M. O’Donnell. Sources of flexibility in dynamic hypertext generation. In *Proceedings of the 8th International Workshop on Natural Language Generation (INLG’96)*, pages 151 – 160, 1996.

- [Knublauch 05]
H. Knublauch. Ramblings on Agile Methodologies and Ontology-Driven Software Development. In *Workshop on Semantic Web Enabled Software Engineering (SWESE)*, Galway, Ireland, 2005.
- [Knuth 93]
D. Knuth. *The Stanford GraphBase, A Platform for Combinatorial Computing*. ACM Press, New York, 1993.
- [Kobsa 90]
A. Kobsa. Modelling the user’s conceptual knowledge in bgp-ms, a user modelling shell system. *Computational Intelligence*, 6:193–208, 1990.
- [Kogut & Holmes 01]
P. Kogut and W. Holmes. AeroDAML: Applying Information Extraction to Generate DAML Annotations from Web Pages. In *First International Conference on Knowledge Capture (K-CAP 2001), Workshop on Knowledge Markup and Semantic Annotation*, Victoria, B.C., 2001.
- [Kohavi & Sommerfield 96]
R. Kohavi and D. Sommerfield. MLC++ – Machine Learning Library in C++. <http://www.sgi.com/Technology/mlc>, Silicon Graphics Inc., 1996.
- [Kohlschütter *et al.* 10]
C. Kohlschütter, P. Fankhauser, and W. Nejdl. Boilerplate Detection using Shallow Text Features. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, 2010.
- [Kohonen *et al.* 81]
T. Kohonen, E. Oja, and P. Lehtio. Storage and processing of information in distributed associative memory systems. In G. Hinton and J. Anderson, editors, *Parallel Models of Associative Memory*. Erlbaum, Hillsdale, N.J., 1981.
- [Kokkinakis & Johansson-Kokkinakis 99a]
D. Kokkinakis and S. Johansson-Kokkinakis. A Cascaded Finite-State Parser for Syntactic Analysis of Swedish. Technical report, Department of Swedish, University of Göteborg, Göteborg, 1999.
- [Kokkinakis & Johansson-Kokkinakis 99b]
D. Kokkinakis and S. Johansson-Kokkinakis. Cascaded finite-state parser for syntactic analysis of swedish. Technical Report GU-ISS-99-2, Dept. of Swedish, Göteborg University, 1999. <http://svenska.gu.se/svedk/publications.html>.
- [Kokkinakis 98]
D. Kokkinakis. AVENTINUS, GATE and Swedish Lingware. In *Proceedings of the 11th NODALIDA Conference*, pages 22–33, Copenhagen, 1998.
- [Kokkinakis *et al.* 00]
D. Kokkinakis, G. Magnusdottir, and A. Samiotou. The MUMIS Domain Model (Second Draft). Technical report, ESTeam, 2000.
- [Koller *et al.* 09]
A. Koller, K. Striegnitz, D. Byron, J. Cassell, R. Dale, S. Dalzel-Job, J. Oberlander, and J. Moore. Validating the web-based evaluation of nlg systems. In *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*, pages 301–304, 2009.
- [Koller *et al.* 10]
A. Koller, K. Striegnitz, A. Gargett, D. Byron, J. Cassell, R. Dale, J. Moore, and J. Oberlander. Report on the second nlg challenge on generating instructions in virtual environments (give-2). In *Proceedings of the 6th International Natural Language Generation Conference*, pages 243–250, 2010.

- [Kolodner 93]
J. Kolodner. *Case-Based Reasoning*. Morgan Kaufmann Publishers Inc., San Mateo, California, 1993.
- [Komatani & Kawahara 00]
K. Komatani and T. Kawahara. Generating Effective Confirmation and Guidance Using Two-Level Confidence Measures for Dialogue Systems. In *Int. Conf. on Spoken Language Processing (ICSLP)*, volume 2, pages 648–651, Beijing, China, October 2000.
- [Koning *et al.* 95]
J. Koning, M. Stefanini, and Y. Deamzeau. DAI Interaction Protocols as Control Strategies in a Natural Language Processing System. In *Proceedings of IEEE Conference on Systems, Man and Cybernetics*, 1995.
- [Konstantinova 14]
N. Konstantinova. Review of relation extraction methods: What is new out there? In *Analysis of Images, Social Networks and Texts*, pages 15–28. Springer, 2014.
- [Körner *et al.* 10]
C. Körner, D. Benz, A. Hotho, M. Strohmaier, and G. Stumme. Stop thinking, start tagging: Tag semantics emerge from collaborative verbosity. In *Proceedings of the 19th International Conference on World Wide Web (WWW'10)*, pages 521–530, 2010.
- [Koroleva & Röhler 12]
K. Koroleva and A. B. Röhler. Reducing information overload: Design and evaluation of filtering & ranking algorithms for social networking sites. In *20th European Conference on Information Systems, ECIS 2012, Barcelona, Spain, June 10-13, 2012*, page 12, 2012.
- [Kowalski & Kuehner 71]
R. Kowalski and D. Kuehner. Linear resolution with selection function. *Artificial Intelligence*, 2, 1971.
- [Kowalski 75]
R. Kowalski. A proof procedure using connection graphs. *J.ACM*, 22(4), 1975.
- [Kowalski 79]
R. Kowalski. *Logic and Problem Solving*. North-Holland, New York, 1979.
- [Krahmer *et al.* 01]
E. Krahmer, M. Swerts, M. Theune, and M. Weegels. Error Detection in Spoken Human-Machine Interaction. *International Journal of Speech Technology*, 4(1):19–30, 2001.
- [Krause & Smeddinck 11]
M. Krause and J. Smeddinck. Human computation games: a survey. In *Proceedings of 19th European Signal Processing Conference (EUSIPCO 2011)*, 2011.
- [Krause 80]
J. Krause. Natural Language Access to Information Systems. An Evaluation Study of its Acceptance by End Users. *Information Systems*, 5:297–319, 1980.
- [Krieger & Schaefer 94]
H.-U. Krieger and U. Schaefer. *TDL*—a type description language for constraint-based grammars. In *Proceedings of the 15th International Conference on Computational Linguistics, COLING-94*, pages 893–899, 1994.
- [Kring 13]
A. M. Kring. The clinical assessment interview for negative symptoms (CAINS): final development and validation. *American Journal of Psychiatry*, 170(2):165, February 2013. Cited by 0007.

- [Kripke 72]
S. A. Kripke. *Naming and necessity*. Springer, 1972.
- [Krishnamurthy *et al.* 08]
B. Krishnamurthy, P. Gill, and M. Arlitt. A few chirps about Twitter. In *Proceedings of the first workshop on Online social networks*, pages 19–24. ACM, 2008.
- [Krotov *et al.* 94a]
A. Krotov, R. Gaizauskas, and Y. Wilks. Acquiring a stochastic context-free grammar from the penn treebank. In *Proceedings of the Third International Conference of the Cognitive Science of Natural Language Processing (CSNLP94)*, pages 79–86, Dublin, July 1994.
- [Krotov *et al.* 94b]
A. Krotov, R. Gaizauskas, and Y. Wilks. Acquiring a stochastic context-free grammar from the penn treebank. In *Third International Conference of the Cognitive Science of Natural Language Processing (CSNLP94)*. Dublin, July 1994.
- [Krovetz & Croft 92]
R. Krovetz and W. B. Croft. Lexical ambiguity and information retrieval. *ACM Transactions on Information Systems*, 10(2):115–141, 1992.
- [Krovetz 92]
R. Krovetz. Sense-linking in a machine readable dictionary. In *Proc. of the 30th Annual Meeting of the Association for Computational Linguistics*, pages 330–332, 1992.
- [Krovetz 93]
R. Krovetz. Viewing morphology as an inference process. In *Proc. of 16th Annual International ACM/SIGIR Conference on Research & Development in Information Retrieval*, pages 191–203, 1993.
- [Krug 09]
S. Krug. *Don't make me think: A common sense approach to web usability*. Pearson Education, 2009.
- [Krupka 95]
G. Krupka. Description of the SRA System as used for MUC-6. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, pages 221–236. Morgan Kaufmann, California, 1995.
- [Krupka *et al.* 92]
G. Krupka, P. Jacobs, L. Rau, L. Childs, and I. Sider. GE NLTOOLSET: Description of the System as used for MUC-4. In *Proceedings of the Fourth Message Understanding Conference (MUC-4)*, pages 177–185. Morgan Kaufmann, California, 1992.
- [Ku *et al.* 06a]
L.-W. Ku, Y.-T. Liang, and H.-H. Chen. Opinion extraction, summarization and tracking in news and blog Corpora. In *Proceedings of AAAI-2006 Spring Symposium on Computational Approaches to Analyzing Weblogs*, 2006.
- [Ku *et al.* 06b]
L.-W. Ku, Y.-T. Liang, and H.-H. Chen. Opinion extraction, summarization and tracking in news and blog corpora. In *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs*, pages 100–107, 2006.
- [Kubala *et al.* 94]
F. Kubala, A. Anastasakos, J. Makhoul, L. Nguyen, R. Schwartz, and G. Zavaliagkos. Comparative Experiments on Large Vocabulary Speech Recognition. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 1, pages 561–564, Adelaide, 1994.

- [Kudo & Matsumoto 00]
T. Kudo and Y. Matsumoto. Use of Support Vector Learning for Chunk Identification. In *Proceedings of Sixth Conference on Computational Natural Language Learning (CoNLL-2000)*, 2000.
- [Kudoh & Matsumoto 00]
T. Kudoh and Y. Matsumoto. Japanese Dependency Structure Analysis Based on Support Vector Machines. In *2000 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora*, 2000.
- [Kuhn & de Mori 92]
R. Kuhn and R. de Mori. A cache-based natural language model for speech recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:570–583, 1992.
- [Kuhn & de Mori 95]
R. Kuhn and R. de Mori. The application of semantic classification trees to natural language understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:449–460, 1995.
- [Kuhn & deMori 98]
R. Kuhn and R. de Mori. Sentence Interpretation. In R. de Mori, editor, *Spoken Dialogues with Computers*, chapter 14, pages 485–522. Academic Press, London, 1998.
- [Kuhn 06a]
T. Kuhn. In *Attempto Controlled English as an Ontology Language*. REVERSE -Reasoning on the Web with Rules and Semantics, 2006.
- [Kuhn 06b]
T. Kuhn. Attempto Controlled English as ontology language. In Bry and Schwertel [Bry & Schwertel 06].
- [Kullback & Leibler 51]
S. Kullback and R. A. Leibler. On information and sufficiency. *The annals of mathematical statistics*, pages 79–86, 1951.
- [Kumar *et al.* 16]
S. Kumar, R. West, and J. Leskovec. Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. In *WWW*, pages 591–602. ACM, 2016.
- [Kunath & Weinberger 10]
S. A. Kunath and S. H. Weinberger. The wisdom of the crowd’s ear: speech accent rating and annotation with amazon mechanical turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, pages 168–171, 2010.
- [Kuno & Oettinger 62]
S. Kuno and A. Oettinger. Multiple path syntactic analyser. In *Information Processing 1962, proceedings IFIP, Popplewell (ed.)*, Amsterdam, 1962. North-Holland.
- [Kuo *et al.* 09]
Y.-l. Kuo, J.-C. Lee, K.-y. Chiang, R. Wang, E. Shen, C.-w. Chan, and J. Y.-j. Hsu. Community-based game design: experiments on social games for commonsense data collection. In *Proceedings of the ACM SIGKDD Workshop on Human Computation, HCOMP ’09*, pages 15–22, New York, NY, USA, 2009. ACM.
- [Kurt 01]
H. Kurt. On-line new event detection and tracking in a multi-resource environment. Unpublished M.Sc. thesis, Institute of Science and Engineering of Bilkent University, Turkey, 2001.

- [Kwak *et al.* 10]
H. Kwak, C. Lee, H. Park, and S. Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web, WWW '10*, pages 591–600, New York, NY, USA, 2010. ACM.
- [Kwak *et al.* 11]
H. Kwak, H. Chun, and S. Moon. Fragile online relationship: A first look at unfollow dynamics in twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '11*, pages 1091–1100, New York, NY, USA, 2011. ACM.
- [Laclavik & Maynard 09]
M. Laclavik and D. Maynard. Motivating intelligent email in business: an investigation into current trends for email processing and communication research. In *Proceedings of Workshop on Emails in e-Commerce and Enterprise Context, 11th IEEE Conference on Commerce and Enterprise Computing*, Vienna, Austria, 2009.
- [Laclavik *et al.* 07]
M. Laclavik, M. Seleng, E. Gatial, Z. Balogh, and L. Hluchy. Ontology based Text Annotation – OnTeA. In *Proceedings of the 2007 conference on Information Modelling and Knowledge Bases XVIII*, pages 311–315, Amsterdam, The Netherlands, The Netherlands, February 2007. IOS Press.
- [Laclavik *et al.* 08]
M. Laclavik, M. Seleng, and L. Hluchy. Towards large scale semantic annotation built on mapreduce architecture. In M. Bubak, G. van Albada, J. Dongarra, and P. Sloot, editors, *Computational Science – ICCS 2008*, volume 5103 of *Lecture Notes in Computer Science*, pages 331–338. Springer Berlin / Heidelberg, 2008.
- [Lafferty & Berger 99]
J. Lafferty and A. Berger. Information Retrieval as Statistical Translation. In *Proceedings of the 22nd International Conference on Research and Development in Information retrieval (SIGIR'99)*, Carnegie Mellon University, Pittsburgh, PA 15213, 1999.
- [Lafferty *et al.* 01]
J. Lafferty, A. McCallum, and F. Pereira. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In *Proceedings of the Eighteenth International Conference on Machine Learning*, pages 282–289, San Francisco: Morgan Kaufmann, 2001.
- [Lai 10]
P. Lai. Extracting Strong Sentiment Trends from Twitter. <http://nlp.stanford.edu/courses/cs224n/2011/reports/patlai.pdf>, 2010.
- [Lal & Ruger 02]
P. Lal and S. Ruger. Extract-based summarization with simplification. In *Proceedings of the ACL 2002 Automatic Summarization / DUC 2002 Workshop*, 2002. <http://www.doc.ic.ac.uk/~srueger/pr-p.lal-2002/duc02-final.pdf>.
- [Lal 02]
P. Lal. Text summarisation. Unpublished M.Sc. thesis, Imperial College, London, 2002.
- [Lamel *et al.* 94]
L. Lamel, F. Schiel, A. Fourcin, and J. Mariani. The translingual english database (ted)'. In *Proc., ICSLP*, Yokohama, 1994.
- [Lampos *et al.* 10]
V. Lampos, T. De Bie, and N. Cristianini. Flu detector-tracking epidemics on twitter. In *Machine Learning and Knowledge Discovery in Databases*, pages 599–602. Springer, 2010.

- [Lamos *et al.* 13]
V. Lamos, D. Preotiu-Pietro, and T. Cohn. A user-centric model of voting intention from social media. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, pages 993–1003. Association for Computational Linguistics, 2013.
- [Landauer & Dumais 40]
T. K. Landauer and S. T. Dumais. A solution to plato’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104:1997, 211-240.
- [Landauer & Hirakawa 95]
J. Landauer and M. Hirakawa. Visual AWK: A Model for Text Processing by Demonstration. In *Proceedings VL’95 11th International IEEE Symposium on Visual Languages, Darmstadt*. IEEE Computer Society Press, 1995.
- [Landis & Koch 77]
J. R. Landis and G. G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, pages 159–174, 1977.
- [Landman & Morzycki 02]
M. Landman and M. Morzycki. Event-kinds and the representation of manner. In *Proceedings of the Western Conference in Linguistics (WECOL)*, volume 11, 2002.
- [Landman 85]
F. Landman. The realist theory of meaning. *Language and Philosophy* 8, pages 35–51, 1985.
- [Lane 99]
D. Lane. *Hyperstat*. 2nd edition, Dec 1999.
- [Lang 88]
B. Lang. Parsing incomplete sentences. *Proceedings of COLING 88*, 1988.
- [Langkilde & Knight 98]
I. Langkilde and K. Knight. Generation that exploits corpus-based statistical knowledge. In *COLING-ACL ’98, Proceedings of the Conference*, pages 704 – 710, Montreal, Canada, August 1998. Association for Computational Linguistics.
- [Langkilde-Geary 02]
I. Langkilde-Geary. An empirical verification of coverage and correctness for a general-purpose sentence generator. In *Proceedings of the International Natural Language Generation Conference (INLG)*, 2002.
- [Langley *et al.* 02]
C. Langley, A. Lavie, L. Levin, D. Wallace, D. Gates, and K. Peterson. Spoken Language Parsing Using Phrase-Level Grammars and Trainable Classifiers. In *Proceedings of the Workshop Speech-to-Speech Translation*, pages 15–22, July 2002.
- [Laniado & Mika 10]
D. Laniado and P. Mika. Making sense of twitter. In P. Patel-Schneider, Y. Pan, P. Hitzler, P. Mika, L. Zhang, J. Pan, I. Horrocks, and B. Glimm, editors, *The Semantic Web – ISWC 2010*, volume 6496 of *Lecture Notes in Computer Science*, pages 470–485. Springer Berlin / Heidelberg, 2010.
- [Lappin & Leass 94]
S. Lappin and H. Leass. An algorithm for pronominal anaphora resolution. *Computational Linguistics*, 20:535–561, 1994.
- [Lari & Young 90]
K. Lari and S. J. Young. The estimation of stochastic context-free grammars using the Inside-Outside algorithm. *Computer Speech and Language*, 4:35–56, 1990.

- [Larkin & Simon 87]
 J. Larkin and H. Simon. Why a diagram is (sometimes) worth then thousand words. *Cognitive Science*, 11:65–99, 1987.
- [Lascarides & Asher 91]
 A. Lascarides and N. Asher. Discourse relations and defeasible knowledge. In *Proceedings of the 29th Conference of the Association for Computational Linguistics*, pages 55–62, 1991.
- [Lassila & Swick 99]
 O. Lassila and R. Swick. Resource Description Framework (RDF) Model and Syntax Specification. Technical Report 19990222, W3C Consortium, <http://www.w3.org/TR/REC-rdf-syntax/>, 1999.
- [Lau et al. 93]
 R. Lau, R. Rosenfeld, and S. Roukos. Trigger-based language models: A maximum entropy approach. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 2, pages 45–48, Minneapolis MN, 1993.
- [Lauriston 95]
 A. Lauriston. Criteria for measuring term recognition. In *Proc. of 7th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 17–22, 1995.
- [Lauriston 96]
 A. Lauriston. *Automatic term recognition: performance of linguistic and statistical learning techniques*. Unpublished PhD thesis, UMIST, Manchester, UK, 1996.
- [Lavelli et al. 02]
 A. Lavelli, F. Pianesi, E. Maci, I. Prodanof, L. Dini, and G. Mazzini. SiSSA: An infrastructure for developing nlp applications. In *Proceedings of 3rd Language Resources and Evaluation Conference (LREC'2002)*, Gran Canaria, Spain, 2002.
- [Lavelli et al. 04a]
 A. Lavelli, M. Califf, F. Ciravegna, D. Freitag, C. Giuliano, N. Kushmerick, and L. Romano. A Critical Survey of the Methodology for IE Evaluation. In *Proceedings of the Language Resources and Evaluation Conference (LREC 2004)*, Lisbon, Portugal, 2004.
- [Lavelli et al. 04b]
 A. Lavelli, M. Califf, F. Ciravegna, D. Freitag, C. Giuliano, N. Kushmerick, and L. Romano. IE evaluation: Criticisms and recommendations. In *Proceedings of Workshop on Adaptive Text Extraction and Mining (ATEM 2004)*, San Jose, California, USA, 2004.
- [Lavoie & Rambow 97]
 B. Lavoie and O. Rambow. A fast and portable realizer for text generation systems. In *ANLP'97 [ANLP'97 97]*, pages 265–68.
- [Lavoie et al. 97]
 B. Lavoie, O. Rambow, and E. Reiter. Customizable Descriptions of Object-Oriented Models. In *Proceedings of the 5th Conference on Applied Natural Language Processing (ANLP-97)*, March 1997.
- [Lawrence 01a]
 S. Lawrence. Access to Scientific Literature. *The Nature Yearbook of Science and Technology*, pages 86–88, 2001.
- [Lawrence 01b]
 S. Lawrence. Online or Invisible? *Nature*, 411(6837):521, 2001.
- [Laws et al. 11]
 F. Laws, C. Scheible, and H. Schütze. Active learning with amazon mechanical turk. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1546–1556, 2011.

- [Lawson 97]
V. Lawson. The terms and arts of patentese: wolves in sheep’s clothing. In S. Wright and G. Budin, editors, *Handbook of Terminology Management*, volume 1: Basic Aspects of Terminology Management, chapter 2.1.3, pages 171–183. John Benjamins, Amsterdam, 1997.
- [Lawson et al. 10]
N. Lawson, K. Eustice, M. Perkowitz, and M. Yetisgen-Yildiz. Annotating large email datasets for named entity recognition with mechanical turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, pages 71–79, 2010.
- [Leacock et al. 93]
C. Leacock, G. Towell, and E. Voorhees. Corpus-Based Statistical Sense Resolution. In *Proceedings of ARPA Human Language Technology Workshop*, pages 260–265, 1993.
- [Lee & Wilks 96a]
M. Lee and Y. Wilks. An ascription-based approach to speech acts. In *Proceedings of the 16th Conference on Computational Linguistics (COLING-96)*, Copenhagen, 1996.
- [Lee & Wilks 96b]
M. Lee and Y. Wilks. Speech acts on demand. In *Proceedings of the 5th International Pragmatics Conference*, Mexico City, 1996.
- [Lee & Wilks 97]
M. Lee and Y. Wilks. Eliminating deceptions and mistaken belief to infer conversational implicature. In *IJCAI-97 workshop on Conflict, Cooperation and Collaboration in Dialogue Systems*, Tokyo, 1997.
- [Lee 88]
K. F. Lee. *Large Vocabulary Speaker-Independent Continuous Speech Recognition: The SPHINX System*. Unpublished PhD thesis, School of Computer Science, Carnegie Mellon University, 1988.
- [Lee 89]
K. Lee. *Automatic Speech Recognition, the Development of the SPHINX System*. Kluwer, Dordrecht, Netherlands, 1989.
- [Lee et al. 04]
Y. Lee, H. Ng, and T. Chia. Supervised Word Sense Disambiguation with Support Vector Machines and Multiple Knowledge Sources. In *Proceedings of SENSEVAL-3: Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text*, pages 137–140, 2004.
- [Lee et al. 11]
K. Lee, B. D. Eoff, and J. Caverlee. Seven months with the devils: A long-term study of content polluters on twitter. In L. A. Adamic, R. A. Baeza-Yates, and S. Counts, editors, *Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, July 17-21, 2011*. The AAAI Press, 2011.
- [Leek 97]
T. R. Leek. Information Extraction Using Hidden markov Models. Technical report, University of California, San Diego, USA, 1997.
- [Lees & Chomsky 57]
R. B. Lees and N. Chomsky. Syntactic structures. *Language*, 33(3 Part 1):375–408, 1957.
- [Leginus et al. 15]
M. Leginus, L. Derczynski, and P. Dolog. Enhanced information access to social streams through word clouds with entity grouping. In *Proceedings of the 11th International Conference on Web Information Systems and Technologies (WEBIST)*, 2015.

- [Lehmann & Völker 14]
 J. Lehmann and J. Völker. *Perspectives on Ontology Learning*, volume 18. IOS Press, 2014.
- [Lehnert 77]
 W. Lehnert. Human computational question answering. *Cognitive Science*, 1:47–73, 1977.
- [Lehnert *et al.* 92]
 W. Lehnert, C. Cardie, D. Fisher, J. McCarthy, and E. Riloff. University of Massachusetts: Description of the CIRCUS system as used for MUC-4. In *Proceedings of the Fourth Message Understanding Conference (MUC-4)*, pages 282–288. Morgan Kaufmann, 1992.
- [Lehrberger & Bourbeau 88]
 J. Lehrberger and L. Bourbeau. Machine translation: linguistic characteristics of mt systems and general methodology of evaluation. In *Linguisticae Investigationes Supplementa – Studies in French and German Linguistics*, pages 132–185. John Benjamins, 1988.
- [Lehrberger 82]
 J. Lehrberger. Automatic translation and the concept of sublanguage. In R. Kittredge and J. Lehrberger, editors, *Sublanguage: Studies of Language in Restricted Semantic Domains*. Walter de Gruyter, 1982.
- [Lehtola *et al.* 99]
 A. Lehtola, J. Tenni, C. Bounsaythip, and K. Jaaranen. Controlled Languages as the Basis for Multilingual Catalogues on the WWW. In J. Roger, B. Stanford-Smith, and P. Kidd, editors, *Business and Work in the Information Society: New Technologies and Applications*, pages 207–213. IOS-Press, Amsterdam, 1999.
- [Lei *et al.* 06a]
 Y. Lei, V. Uren, and E. Motta. SemSearch: a Search Engine for the Semantic Web. In *Managing Knowledge in a World of Networks*, pages 238–245. Springer Berlin / Heidelberg, 2006.
- [Lei *et al.* 06b]
 Y. Lei, V. Uren, and E. Motta. Semsearch: A search engine for the semantic web. In S. Staab and V. Svátek, editors, *Managing Knowledge in a World of Networks*, volume 4248 of *Lecture Notes in Computer Science*, pages 238–245. Springer Berlin, Heidelberg, 2006.
- [Leidner 03]
 J. Leidner. Current Issues in Software Engineering for Natural Language Processing. In *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, 2003.
- [Lenat & Feigenbaum 87]
 D. B. Lenat and E. A. Feigenbaum. On the thresholds of knowledge. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence (IJCAI-87)*, pages 1173–1182. Milan, Italy, 1987.
- [Lenat & Guha 90]
 D. Lenat and R. Guha. *Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project*. Addison-Wesley, Reading, MA, 1990.
- [Lenat 83]
 D. Lenat. The role of heuristics in learning by discovery: Three case studies. In J. Carbonell, R. Michalski, and T. Mitchell, editors, *Machine Learning: An Artificial Intelligence Approach*. Tioga Pub. Co., Palo Alto, CA, 1983.
- [Lenat 95]
 D. Lenat. CYC: A Large-Scale Investment in Knowledge Infrastructure. *Communications of the ACM*, 38(11), November 1995.

- [Lenat *et al.* 86]
D. B. Lenat, M. Prakash, and M. Shepherd. Cyc: Using common sense knowledge to overcome brittleness and knowledge acquisition bottlenecks. *AI Magazine*, 7(7):65–85, 1986.
- [Lenat *et al.* 95]
D. Lenat, G. Miller, and T. Yokoi. CYC, WordNet, and EDR: Critiques and Responses. *Communications of the ACM*, 38(11), November 1995.
- [Lerman *et al.* 06]
K. Lerman, A. Plangprasopchok, and C. Knoblock. Automatically labeling the inputs and outputs of web services. In *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI-06)*, Boston, July 2006.
- [Leser & Hakenberg 05]
U. Leser and J. Hakenberg. What makes a gene name? named entity recognition in the biomedical literature. *Briefings in Bioinformatics*, 6(4):357–369, 2005.
- [Lesk 86a]
M. Lesk. Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone. In *Proceedings of ACM SIGDOC Conference*, pages 24–26, Toronto, Ontario, 1986.
- [Lesk 86b]
M. Lesk. Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone. In *Proceedings of ACM SIGDOC Conference*, pages 24–26, Toronto, Ontario, 1986.
- [Leskovec *et al.* 09]
J. Leskovec, L. Backstrom, and J. M. Kleinberg. Meme-tracking and the dynamics of the news cycle. In *KDD*, pages 497–506. ACM, 2009.
- [Lessig 04]
L. Lessig. *Free Culture*. Penguin, New York, 2004.
- [Leung & Zue 88]
H. C. Leung and V. W. Zue. Some phonetic recognition experiments using artificial neural nets. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 422–425, New York, 1988.
- [Leung & Zue 90]
H. C. Leung and V. W. Zue. Phonetic classification using multi-layer perceptrons. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 525–528, Albuquerque, 1990.
- [Levenshtein 66]
V. I. Levenshtein. Binary Codes capable of correcting deletions, insertions and reversals. *Soviet Phys. Dokl.*, 10:707–710, 1966.
- [Levesque 86]
H. Levesque. Making believers out of computers. *Artificial Intelligence*, 30, pages 81–108, 1986.
- [Levin & Pieraccini 95a]
E. Levin and R. Pieraccini. CHRONUS - The Next Generation. In *Proceedings of ARPA Spoken Language Systems Technology Workshop*, pages 269–271, January 1995.
- [Levin & Pieraccini 95b]
E. Levin and R. Pieraccini. Concept-Based Spontaneous Speech Understanding System. In *European Conference on Speech Communication and Technology (EUROSPEECH)*, volume 2, pages 555–558, Madrid, Spain, September 1995.

- [Levine 79]
 B. Levine. The automated inference of tree systems. Phd thesis, computer science department, Oregon State University, 1979.
- [Levinson & Wittenburg 01]
 S. Levinson and P. Wittenburg. Language as cultural heritage - promoting research and public awareness on the internet. In J. Renn, editor, *ECHO – an infrastructure to bring European Cultural Heritage Online*, pages 104–111. Max Planck Institute for the History of Science, 2001.
- [Levinson *et al.* 83a]
 S. E. Levinson, L. R. Rabiner, and M. M. Sondhi. An introduction to the application of the theory of probabilistic processes of a Markov process to automatic speech recognition. *Bell System Technical Journal*, 62(4):1035–1074, 1983.
- [Levinson *et al.* 83b]
 S. E. Levinson, L. R. Rabiner, and M. M. Sondhi. On the application of vector quantization and hidden markov models to speaker independent isolated word recognition. *Bell System Technical Journal*, 62(4), 1983.
- [Levoy *et al.* 00]
 M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk. The Digital Michelangelo Project: 3D Scanning of Large Statues. In K. Akeley, editor, *Siggraph 2000, Computer Graphics Proceedings*, pages 131–144. ACM Press / ACM SIGGRAPH / Addison Wesley Longman, 2000.
- [Levy *et al.* 97]
 D. Levy, R. Catizone, B. Battacharia, A.Krotov, and Y. Wilks. CONVERSE: A conversational companion. In *Proceedings of the 1st International Workshop on Human-Computer Conversation*, Bellagio, Italy, 1997.
- [Lewandowsky *et al.* 12]
 S. Lewandowsky, U. K. Ecker, C. M. Seifert, N. Schwarz, and J. Cook. Misinformation and its correction continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3):106–131, 2012.
- [Lewis *et al.* 04]
 D. D. Lewis, Y. Yang, T. G. Rose, and F. Li. Rcv1: A new benchmark collection for text categorization research. *Journal of Machine Learning Research*, 5(Apr):361–397, 2004.
- [Li & Bontcheva 08]
 Y. Li and K. Bontcheva. Adapting support vector machines for f-term-based classification of patents. *ACM Transactions on Asian Language Information Processing*, 7(2):7:1–7:19, 2008.
- [Li & Cunningham 08]
 Y. Li and H. Cunningham. Geometric and Quantum Methods for Information Retrieval. *SIGIR Forum*, 42(2):22–32, 2008.
- [Li & McCallum 03]
 W. Li and A. McCallum. Rapid Development of Hindi Named Entity Recognition using Conditional Random Fields and Feature Induction. *Special issue of ACM Transactions on Asian Language Information Processing: Rapid Development of Language Capabilities: The Surprise Languages*, 2003.
- [Li & Shawe-Taylor 03]
 Y. Li and J. Shawe-Taylor. The SVM with Uneven Margins and Chinese Document Categorization. In *Proceedings of The 17th Pacific Asia Conference on Language, Information and Computation (PACLIC17)*, Singapore, Oct. 2003.

- [Li & Shawe-Taylor 06]
 Y. Li and J. Shawe-Taylor. Using KCCA for Japanese-English Cross-language Information Retrieval and Document Classification. *Journal of Intelligent Information Systems*, 27(2):117–133, 2006.
- [Li & Shawe-Taylor 07]
 Y. Li and J. Shawe-Taylor. Advanced Learning Algorithms for Cross-language Patent Retrieval and Classification. *Information Processing and Management*, 43(5):1183–1199, 2007.
- [Li 11]
 A. Li. Comparing public cloud providers. *IEEE Internet Computing*, 15(2):50–53, 2011.
- [Li et al. 98]
 L. Li, D. Dahl, L. Norton, M. Linebarger, and D. Chen. A Test Environment for Natural Language Understanding Systems. In *Proceedings of the 36th ACL and the 17th COLING (ACL-COLING '98)*, pages 763–767, Montreal, 1998.
- [Li et al. 02]
 Y. Li, H. Zaragoza, R. Herbrich, J. Shawe-Taylor, and J. Kandola. The Perceptron Algorithm with Uneven Margins. In *Proceedings of the 9th International Conference on Machine Learning (ICML-2002)*, pages 379–386, 2002.
- [Li et al. 04]
 Y. Li, K. Bontcheva, and H. Cunningham. An SVM Based Learning Algorithm for Information Extraction. Machine Learning Workshop, Sheffield, 2004. <http://gate.ac.uk/sale/ml-ws04/mlw2004.pdf>.
- [Li et al. 05a]
 Y. Li, K. Bontcheva, and H. Cunningham. Using Uneven Margins SVM and Perceptron for Information Extraction. In *Proceedings of Ninth Conference on Computational Natural Language Learning (CoNLL-2005)*, 2005.
- [Li et al. 05b]
 Y. Li, C. Miao, K. Bontcheva, and H. Cunningham. Perceptron Learning for Chinese Word Segmentation. In *Proceedings of Fourth SIGHAN Workshop on Chinese Language processing (Sighan-05)*, pages 154–157, Korea, 2005.
- [Li et al. 05c]
 Y. Li, K. Bontcheva, and H. Cunningham. SVM Based Learning System For Information Extraction. In J. Winkler, M. Niranjan, and N. Lawrence, editors, *Deterministic and Statistical Methods in Machine Learning: First International Workshop, 7–10 September, 2004*, volume 3635 of *Lecture Notes in Computer Science*, pages 319–339, Sheffield, UK, 2005. Springer.
- [Li et al. 06]
 Y. Li, K. Bontcheva, and H. Cunningham. Perceptron-like learning for ontology based information extraction. Technical report, University of Sheffield, Sheffield, UK, 2006.
- [Li et al. 07a]
 Y. Li, K. Bontcheva, and H. Cunningham. Hierarchical, Perceptron-like Learning for Ontology Based Information Extraction. In *16th International World Wide Web Conference (WWW2007)*, pages 777–786, May 2007.
- [Li et al. 07b]
 Y. Li, K. Bontcheva, and H. Cunningham. Cost Sensitive Evaluation Measures for F-term Patent Classification. In *The First International Workshop on Evaluating Information Access (EVIA 2007), 15 May 2007*, pages 44–53, Tokyo, Japan, May 2007.

- [Li *et al.* 07c]
 Y. Li, K. Bontcheva, and H. Cunningham. Experiments of opinion analysis on the corpora MPQA and NTCIR-6. In *Proceedings of the Sixth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access*, pages 323–329, May 2007.
- [Li *et al.* 07d]
 Y. Li, K. Bontcheva, and H. Cunningham. SVM Based Learning System for F-term Patent Classification. In *Proceedings of the Sixth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access*, pages 396–402, May 2007.
- [Li *et al.* 09a]
 Y. Li, K. Bontcheva, and H. Cunningham. Adapting SVM for Data Sparseness and Imbalance: A Case Study on Information Extraction. *Natural Language Engineering*, 15(2):241–271, 2009.
- [Li *et al.* 09b]
 Y. Li, K. Bontcheva, and H. Cunningham. Adapting SVM for Data Sparseness and Imbalance: A Case Study on Information Extraction. *Natural Language Engineering*, 15(2):241–271, 2009.
- [Li *et al.* 12]
 C. Li, J. Weng, Q. He, Y. Yao, A. Datta, A. Sun, and B.-S. Lee. Twiner: named entity recognition in targeted twitter stream. In *Proc. of the 35th ACM Conf. on Research and Development in Information Retrieval*, pages 721–730. ACM, 2012.
- [Liang 05]
 P. Liang. Semi-supervised learning for natural language. Unpublished M.Sc. thesis, Massachusetts Institute of Technology, 2005.
- [Liben-Nowell *et al.* 05]
 D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins. Geographic routing in social networks. *Proceedings of the National Academy of Sciences*, 102(33):11623–11628, August 2005.
- [Lieberman *et al.* 07]
 H. Lieberman, D. Smith, and A. Teeters. Common consensus: A web-based game for collecting commonsense goals. In *Proceedings of IUI’07*, 2007.
- [Liebrecht *et al.* 13]
 C. Liebrecht, F. Kunneman, and A. van den Bosch. The perfect solution for detecting sarcasm in tweets# not. *WASSA 2013*, page 29, 2013.
- [Likert 32]
 R. Likert. A technique for the measurement of attitudes. *Archives of Psychology*, 22(140):1–55, 1932.
- [Lin & Dyer 10]
 J. Lin and C. Dyer. *Data-Intensive Text Processing with MapReduce*, volume 3. Morgan& Claypool, San Francisco, USA, 2010.
- [Lin 95]
 D. Lin. Description of the PIE System as used for MUC-6. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, pages 113–126, San Francisco, 1995. Morgan Kaufmann, California.
- [Lin 97]
 D. Lin. Using syntactic dependency as local context to resolve word sense ambiguity. In *Proc. of ACL/EACL*, pages 64–71, Madrid, Spain, 1997.

- [Lin 98a]
D. Lin. Extracting collocations from text corpora. In *Proc. of 1st Workshop on Computational Terminology, Computerm '98*, pages 57–63, Montreal, Canada, 1998. COLING-ACL.
- [Lin 98b]
D. Lin. An information-theoretic definition of similarity. In *Proceedings of the Fifteenth International Conference on Machine Learning, ICML '98*, pages 296–304, San Francisco, CA, USA, 1998. Morgan Kaufmann Publishers Inc.
- [Lin 04a]
C.-Y. Lin. Rouge: A package for automatic evaluation of summaries. In *Text Summarization Branches Out: Proceedings of the ACL-04 Workshop*, pages 74–81, Barcelona, Spain, July 2004. Association for Computational Linguistics.
- [Lin 04b]
C.-Y. Lin. Rouge: A package for automatic evaluation of summaries. In *Text Summarization Branches Out: Proceedings of the ACL-04 Workshop*, pages 74–81, Barcelona, Spain, 2004.
- [Lin 08]
J. Lin. Scalable language processing algorithms for the masses: a case study in computing word co-occurrence matrices with mapreduce. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 419–428, 2008.
- [Lin et al. 06]
W. Lin, T. Wilson, J. Wiebe, and A. Hauptmann. Which Side are You on? Identifying Perspectives at the Document and Sentence Levels. In *Proceedings of the Tenth Conference on Computational Natural Language Learning (CoNLL-X)*, pages 109–116, 2006.
- [Linckels & Meinel 07]
S. Linckels and C. Meinel. Semantic Interpretation of Natural Language User Input to Improve Search in Multimedia Knowledge Base. *it - Information Technologies*, 49(1):40–48, 2007.
- [Liporace 82]
L. A. Liporace. Maximum likelihood estimation for multivariate observations of Markov sources. *IEEE Transactions on Information Theory*, IT-28:729–734, 1982.
- [Litkowski 99]
K. C. Litkowski. Towards a meaning-full comparison of lexical resources. In *Proceedings of the Association for Computational Linguistics Special Interest Group on the Lexicon*, College Park, MD, 1999.
- [Litman et al. 99]
D. J. Litman, M. A. Walker, and M. S. Kearns. Automatic Detection of Poor Speech Recognition at the Dialogue Level. In *Proceedings of ACL*, pages 309–316, University of Maryland, June 1999.
- [Liu & Luo 07]
C.-Y. Liu and S.-Y. Luo. Investigation of Carbon Nanotubes Using the F-term Code of Japanese Patent Information. *Data Science Journal*, 6(Supplement):S255–S260, 2007.
- [Liu et al. 03]
H. Liu, H. Lieberman, and T. Selker. A model of textual affect sensing using real-world knowledge. In *Proceedings of the 8th International Conference on Intelligent User Interfaces*, pages 125–132, 2003.
- [Liu et al. 05]
B. Liu, M. Hu, and J. Cheng. Opinion observer: analyzing and comparing opinions on the web. In *Proceedings of the 14th international conference on World Wide Web (WWW '05)*, pages 342–351, New York, NY, USA, 2005. ACM.

- [Liu *et al.* 11a]
 X. Liu, S. Zhang, F. Wei, and M. Zhou. Recognizing named entities in tweets. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 359–367, 2011.
- [Liu *et al.* 11b]
 X. Liu, K. Li, M. Zhou, and Z. Xiong. Collective semantic role labeling for tweets with clustering. In *IJCAI*, volume 11, pages 1832–1837. Citeseer, 2011.
- [Liu *et al.* 12]
 X. Liu, M. Zhou, F. Wei, Z. Fu, and X. Zhou. Joint inference of named entity recognition and normalization for tweets. In *Proceedings of the Association for Computational Linguistics*, pages 526–535, 2012.
- [Liu *et al.* 15]
 X. Liu, A. Nourbakhsh, Q. Li, R. Fang, and S. Shah. Real-time rumor debunking on twitter. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, CIKM '15*, pages 1867–1870, New York, NY, USA, 2015. ACM.
- [Lleida & Rose 96]
 E. Lleida and R. C. Rose. Efficient Decoding and Training Procedures for Utterance Verification in Continuous Speech Recognition. In *Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 1, pages 507–510, Atlanta, Georgia, May 1996.
- [Llorens *et al.* 10a]
 H. Llorens, E. Saquete, and B. Navarro. TIPSem (English and Spanish): Evaluating CRFs and Semantic Roles in TempEval-2. In *Proceedings of SemEval-2010*, pages 284–291. ACL, 2010.
- [Llorens *et al.* 10b]
 H. Llorens, E. Saquete, and B. Navarro-Colorado. TimeML events recognition and classification: learning CRF models with semantic roles. In *Proceedings of the 23rd International Conference on Computational Linguistics*, pages 725–733. Association for Computational Linguistics, 2010.
- [Llorens *et al.* 11]
 H. Llorens, E. Saquete, and B. Navarro. Syntax-motivated context windows of morpho-lexical features for recognizing time and event expressions in natural language. In *Natural Language Processing and Information Systems*, pages 295–299. Springer, 2011.
- [Llorens *et al.* 12a]
 H. Llorens, L. Derczynski, R. J. Gaizauskas, and E. Saquete. TIMEN: An open temporal expression normalisation resource. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC), Istanbul, Turkey*, pages 3044–3051, 2012.
- [Llorens *et al.* 12b]
 H. Llorens, L. Derczynski, R. J. Gaizauskas, and E. Saquete. TIMEN: An open temporal expression normalisation resource. In *LREC*, pages 3044–3051, 2012.
- [Llorens *et al.* 13]
 H. Llorens, E. Saquete, and B. Navarro-Colorado. Applying semantic knowledge to the automatic processing of temporal expressions and events in natural language. *Information Processing & Management*, 49(1):179–197, 2013.
- [Lloyd 87]
 J. Lloyd. *Foundations of Logic Programming*. Springer-Verlag, Berlin, second edition, 1987.
- [Lloyd *et al.* 09]
 K. Lloyd, M. Cella, M. Tanenblatt, and A. Coden. Analysis of clinical uncertainties by health professionals and patients: an example from mental health. *BMC Medical Informatics and Decision Making*, 9(1):34, 2009.

- [Lombard *et al.* 02]
M. Lombard, J. Snyder-Duch, and C. C. Bracken. Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, 28:587–604, 2002.
- [Loper & Bird 02a]
E. Loper and S. Bird. NLTK: The Natural Language Toolkit. In *ACL Workshop on Effective Tools and Methodologies in Teaching NLP*, 2002.
- [Loper & Bird 02b]
E. Loper and S. Bird. Nltk: The natural language toolkit. *CoRR*, cs.CL/0205028, 2002.
- [Lopez & Motta 04]
V. Lopez and E. Motta. Ontology-driven Question Answering in AquaLog. In *NLDB 2004 (9th International Conference on Applications of Natural Language to Information Systems)*, Manchester, UK, 2004.
- [Lopez 10]
P. Lopez. Automatic extraction and resolution of bibliographical references in patent documents. In H. Cunningham, A. Hanbury, and S. Rüger, editors, *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference)*. LNCS volume number: 6107, Lecture Notes in Computer Science, Vienna, Austria, May 2010. Springer.
- [Lopez *et al.* 06]
V. Lopez, E. Motta, and V. S. Uren. Poweraqua: Fishing the semantic web. In *ESWC*, pages 393–410, 2006.
- [Lopez *et al.* 07a]
V. Lopez, E. Motta, V. Uren, and M. Sabou. State of the art on Semantic Question Answering. Technical report, Knowledge Media Institute, May 2007.
- [Lopez *et al.* 07b]
V. Lopez, V. Uren, E. Motta, and M. Pasin. AquaLog: An Ontology-driven Question Answering System for Organizational Semantic Intranets. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(2):72–105, June 2007.
- [Lösch & Müller 11]
U. Lösch and D. Müller. Mapping microblog posts to encyclopedia articles. *Lecture Notes in Informatics*, 192(150), 2011.
- [Lotan *et al.* 11]
G. Lotan, E. Graeff, M. Ananny, D. Gaffney, I. Pearce, and danah boyd. The arab spring| the revolutions were tweeted: Information flows during the 2011 tunisian and egyptian revolutions. *International Journal of Communication*, 5(0), 2011.
- [Louis & Nenkova 13]
A. Louis and A. Nenkova. Automatically assessing machine summary content without a gold standard. *Computational Linguistics*, 39(2):267–300, 2013.
- [Loukis 07]
E. N. Loukis. An ontology for g2g collaboration in public policy making, implementation and evaluation. *Artificial Intelligence and Law*, 15(1):19–48, 2007.
- [Loveland 78]
D. Loveland. *Automated Theorem Proving: A Logical Basis*. North-Holland-Elsevier, Amsterdam, 1978.
- [Lovins 68]
J. B. Lovins. Development of a stemming algorithm. *Mechanical Translation and Computational Linguistics*, 11:22–31, 1968.

- [Lowe & Tipping 96]
D. Lowe and M. Tipping. Feed-forward neural networks and topographic mappings for exploratory data analysis. *Neural Computing and Applications*, 4:83–95, 1996.
- [Lowe 04]
D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, January 2004.
- [Lowerre & Reddy 80]
B. Lowerre and R. Reddy. The Harpy speech understanding system. In W. A. Lea, editor, *Trends in Speech Recognition*. Prentice Hall, New York, 1980.
- [LRE10]
Proceedings of the seventh international conference on Language Resources and Evaluation (LREC), Valetta, Malta, May 2010.
- [LREC-1 98]
Conference on Language Resources Evaluation (LREC-1), Granada, Spain, 1998.
- [LREC-2 00]
Second Conference on Language Resources Evaluation (LREC-2), Athens, 2000.
- [Lu et al. 09]
Y. Lu, C. Zhai, and N. Sundaresan. Rated aspect summarization of short comments. In *Proceedings of the 18th International Conference on World Wide Web*, pages 131–140, 2009.
- [Luhn 57]
H. P. Luhn. A statistical approach to mechanized encoding and searching of literary information. *IBM Journal of Research and Development*, 1:309–317, 1957.
- [Luhn 58]
H. P. Luhn. The automatic creation of literature abstracts. *IBM J. Res. Dev.*, 2(2):159–165, April 1958.
- [Lui & Baldwin 11]
M. Lui and T. Baldwin. Cross-domain feature selection for language identification. In *Proceedings of 5th International Joint Conference on Natural Language Processing*, pages 553–561, November 2011.
- [Lui & Baldwin 12a]
M. Lui and T. Baldwin. langid.py: An off-the-shelf language identification tool. In *The 50th Annual Meeting of the Association for Computational Linguistics, Proceedings of the System Demonstrations, July 10, 2012, Jeju Island, Korea*, pages 25–30. The Association for Computer Linguistics, 2012.
- [Lui & Baldwin 12b]
M. Lui and T. Baldwin. langid.py: An off-the-shelf language identification tool. In *ACL*, pages 25–30. The Association for Computer Linguistics, 2012.
- [Luís & deMatos 09]
T. Luís and D. M. de Matos. High-performance high-volume layered corpora annotation. In *Proceedings of the Third Linguistic Annotation Workshop*, pages 99–107, 2009.
- [Luk 95]
A. Luk. Statistical sense disambiguation with relatively small corpora using dictionary definitions. In *Proceedings of ACL 95*, pages 181–188, 1995.
- [Lukasik et al. 15]
M. Lukasik, T. Cohn, and K. Bontcheva. Classifying tweet level judgements of rumours in social media. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, EMNLP 2015, Lisbon, Portugal, September 17-21, 2015*, pages 2590–2595, 2015.

- [Lukasik *et al.* 16a]
M. Lukasik, K. Bontcheva, T. Cohn, A. Zubiaga, M. Liakata, and R. Procter. Using gaussian processes for rumour stance classification in social media. *CoRR*, abs/1609.01962, 2016.
- [Lukasik *et al.* 16b]
M. Lukasik, P. K. Srijith, D. Vu, K. Bontcheva, A. Zubiaga, and T. Cohn. Hawkes processes for continuous time sequence classification: an application to rumour stance classification in twitter. In *ACL*, 2016.
- [Lund & Burgess 96]
K. Lund and C. Burgess. Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, and Computers*, 28:203–208, 1996.
- [LuperFoy *et al.* 98]
S. LuperFoy, D. Loehr, D. Duff, K. Miller, F. Reeder, and L. Harper. An Architecture for Dialogue Management, Context Tracking, and Pragmatic Adaptation in Spoken Dialogue Systems. In *Proceedings of the 36th ACL and the 17th COLING (ACL-COLING '98)*, pages 794–801, Montreal, 1998.
- [Lupu & Hanbury 13]
M. Lupu and A. Hanbury. Patent retrieval. *Foundations and Trends in Information Retrieval*, 7(1):1–97, 2013.
- [Lux-Pogodalla & Juan 06]
V. Lux-Pogodalla and E. S. Juan. Query refinement by multiword term expansions and semantic synonymy. In *Proceedings of The 1st International Conference on Multidisciplinary Information Sciences and Technologies (InScit2006)*, 2006.
- [Lynch 97]
C. Lynch. Medical terminology management. In S. Wright and G. Budin, editors, *Handbook of Terminology Management*, volume 1: Basic Aspects of Terminology Management, chapter 2.1.2, pages 160–170. John Benjamins, Amsterdam, 1997.
- [Lytinen & Gershman 86a]
S. L. Lytinen and A. Gershman. Atrans: Automatic processing of money transfer messages. In *Proceedings of the 5th National Conference on Artificial Intelligence (AAAI-86)*, pages 1089–1093. Trento, Italy, 1986.
- [Lytinen & Gershman 86b]
S. Lytinen and A. Gershman. ATRANS: Automatic processing of money transfer messages. In *Proceedings of the Fifth National Conference on Artificial Intelligence (AAAI-86)*, pages 1089–1093, 1986.
- [M. Ortuño and P. Carpena and P. Bernaola-Galván and E. Muñoz and A. M. Somoza 02]
M. Ortuño and P. Carpena and P. Bernaola-Galván and E. Muñoz and A. M. Somoza. Keyword detection in natural languages and dna. *Europhysics Letters*, 57(5):759–764, 2002.
- [Ma *et al.* 02]
X. Ma, H. Lee, S. Bird, and K. Maeda. Models and tools for collaborative annotation. In *Proceedings of 3rd Language Resources and Evaluation Conference (LREC'2002)*, Gran Canaria, Spain, 2002.
- [Ma *et al.* 15]
J. Ma, W. Gao, Z. Wei, Y. Lu, and K.-F. Wong. Detect rumors using time series of social context information on microblogging websites. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, CIKM '15*, pages 1751–1754, New York, NY, USA, 2015. ACM.

- [MacGregor 91]
 R. MacGregor. The evolving technology of classification-based knowledge representation systems. In J. Sowa, editor, *Principles of Knowledge Representation*. Morgan Kaufmann, California, 1991.
- [Macherey & Ney 02]
 K. Macherey and H. Ney. Scoring Criteria for Tree based Dialogue Course Management. In *ISCA Tutorial and Research Workshop Multi-Modal Dialogue in Mobile Environments*, Kloster Irsee, Germany, June 2002.
- [Macherey et al. 01]
 K. Macherey, F. J. Och, and H. Ney. Natural Language Understanding Using Statistical Machine Translation. In *European Conference on Speech Communication and Technology (EUROSPEECH)*, volume 3, pages 2205–2208, Aalborg, Denmark, September 2001.
- [Macherey et al. 02]
 W. Macherey, J. Viechtbauer, and H. Ney. Probabilistic Retrieval Based On Document Representations. In *Int. Conf. on Spoken Language Processing (ICSLP)*, pages 1481–1484, Denver, Colorado, September 2002.
- [Macherey et al. 03]
 W. Macherey, J. Viechtbauer, and H. Ney. Probabilistic Aspects in Spoken Document Retrieval. *EURASIP Journal on Applied Signal Processing, Special issue on Unstructured Information Management from Multimedia Data Sources*, (2):115–127, February 2003.
- [Mackie et al. 14]
 S. Mackie, R. McCreddie, C. Macdonald, and I. Ounis. Comparing algorithms for microblog summarisation. In E. Kanoulas, M. Lupu, P. Clough, M. Sanderson, M. Hall, A. Hanbury, and E. Toms, editors, *Information Access Evaluation. Multilinguality, Multimodality, and Interaction*, volume 8685 of *Lecture Notes in Computer Science*, pages 153–159. Springer International Publishing, 2014.
- [Macleod & Grishman 94]
 C. Macleod and R. Grishman. Complex syntax reference manual. New York University, 1994.
- [Macleod et al. 02]
 C. Macleod, N. Ide, and R. Grishman. The American National Corpus: Standardized Resources for American English. In *Proceedings of 2nd Language Resources and Evaluation Conference (LREC)*, pages 831–836, Athens, Greece, 2002.
- [MacWhinney 99]
 B. MacWhinney. *The CHILDES Project: Tools for Analysing Talk (second ed.)*. Lawrence Erlbaum, Hillsdale, N.J., 1999.
- [Madnani et al. 10]
 N. Madnani, J. Boyd-Graber, and P. Resnik. Measuring transitivity using untrained annotators. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, pages 188–194, 2010.
- [Maeda & Strassel 04]
 K. Maeda and S. Strassel. Annotation Tools for Large-Scale Corpus Development: Using AGTK at the Linguistic Data Consortium. In *Proceedings of 4th Language Resources and Evaluation Conference (LREC’2004)*, 2004.
- [Maeda et al. 08]
 K. Maeda, H. Lee, S. Medero, J. Medero, R. Parker, and S. Strassel. Annotation Tool Development for Large-Scale Corpus Creation Projects at the Linguistic Data Consortium. In *Proceedings of the Sixth International Language Resources and Evaluation (LREC’08)*, 2008.

- [Maedche & Staab 00]
A. Maedche and S. Staab. Semi-automatic Engineering of Ontologies from Text. In *Proceedings of the 12th International Conference on Software Engineering and Knowledge Engineering*, 2000.
- [Maedche & Staab 01]
A. Maedche and S. Staab. Comparing Ontologies – Similarity Measures and a Comparison Study. Technical Report Internal Report No. 408, AIFB, University of Karlsruhe, 2001.
- [Maedche & Staab 02]
A. Maedche and S. Staab. Measuring similarity between ontologies. In *EKAW '02: Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web*, pages 251–263, London, UK, 2002. Springer-Verlag.
- [Maedche & Staab 04]
A. Maedche and S. Staab. Ontology learning. In *Handbook on ontologies*, pages 173–190. Springer, 2004.
- [Maedche 02a]
A. Maedche. *Ontology Learning for the Semantic Web*. Kluwer Academic Publishers, Amsterdam, 2002.
- [Maedche 02b]
A. Maedche. *Ontology Learning for the Semantic Web*. Kluwer Academic Publishers, Amsterdam, 2002.
- [Magadur & Tabuteau 96]
J.-Y. Magadur and G. Tabuteau. Semantic disambiguation in an information retrieval system. In *NLP+IA 96*, pages 148–154, Moncton, Canada, 1996.
- [Magerman 94]
D. Magerman. *Natural Language Parsing as Statistical Pattern Recognition*. Unpublished PhD thesis, Department of Computer Science, Stanford University, CA, 1994.
- [Magerman 95a]
D. Magerman. Statistical Decision Tree Models for Parsing. In *Proceedings of ACL*, 1995.
- [Magerman 95b]
D. M. Magerman. Statistical decision-tree models for parsing. In *Proc. of the 33th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 276–283, Cambridge, MA, June 1995.
- [Magnani & Rossi 11]
M. Magnani and L. Rossi. The ML-model for multi-layer social networks. In *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on*, pages 5–12. IEEE, 2011.
- [Mahesh 96]
K. Mahesh. Ontology development for mt: Ideology and methodology. Technical report, CRL New Mexico State University, 1996.
- [Mahesh *et al.* 66]
K. Mahesh, S. Nirenburg, J. Cowie, and D. Farwell. An Assessment of Cyc for Natural Language Processing. Technical Report MCCS Report, New Mexico State University, 1966.
- [Mahesh *et al.* 97]
K. Mahesh, S. Nirenburg, S. Beale, E. Viegas, V. Raskin, and B. Onyshkevych. Word sense disambiguation: Why have statistics when we have these numbers. In *Proceedings of the 7th International Conference on Theoretical and Methodological Issues in Machine Translation*, pages 151–159, July 1997.

- [Mahesh *et al.* 99]
K. Mahesh, J. Kud, and P. Dixon. Oracle at TREC8: A Lexical Approach. In *Proceedings of the Eighth Text Retrieval Conference (TREC-8)*, 1999.
- [Mahmud *et al.* 12]
J. Mahmud, J. Nichols, and C. Drews. Where is this tweet from? Inferring home locations of Twitter users. In *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, Dublin, Ireland, 2012.
- [Majocchi & Strange 07]
A. Majocchi and R. Strange. The FDI Location Decision: does Liberalisation Matter? *Transactional Corporation Review*, 2007. To Appear.
- [Makhoul *et al.* 85]
J. Makhoul, S. Roucos, and H. Gish. Vector quantization in speech coding. *Proceedings of the IEEE*, 73:1551–1588, 1985.
- [Makhoul *et al.* 00]
J. Makhoul, F. Kubala, T. Leek, D. Liu, and L. Nguyen. Speech and language technologies for audio indexing and retrieval. *Proc., IEEE*, 88(8), 2000.
- [Makins 91]
M. Makins, editor. *Collins English Dictionary, 3rd Edition*. Harper Collins, 1991.
- [Makita *et al.* 03]
M. Makita, S. Higuchi, A. Fujii, and T. Ishikawa. A system for Japanese/English/Korean multilingual patent retrieval. In *Proceedings of Machine Translation Summit IX (online at <http://www.amtaweb.org/summit/MTSummit/papers.html>)*, Sept. 2003.
- [Malhotra 75]
A. Malhotra. *Design criteria for a Knowledge based English Language System for Management: An Experimental Analysis*. Unpublished PhD thesis, Massachusetts Institute of Technology, Alfred P. Sloan School of Management, 1975.
- [Malioutov *et al.* 07]
I. Malioutov, A. Park, R. Barzilay, and J. Glass. Making sense of sound: Unsupervised topic segmentation over acoustic input. In *Proceedings, ACL*, 2007.
- [Maloney & Niv 98]
J. Maloney and M. Niv. Tagarab: A fast, accurate arabic name recogniser using high precision morphological analysis. In *Proceedings of the ACL Workshop on Computational Approaches to Semitic Languages*, Montreal, Canada, 1998.
- [Mani & Bloedorn 98]
I. Mani and E. Bloedorn. Machine learning of generic and user-focused summarization. In *Proceedings of the Fifteenth National/Tenth Conference on Artificial Intelligence/Innovative Applications of Artificial Intelligence, AAAI '98/IAAI '98*, pages 820–826, Menlo Park, CA, USA, 1998. American Association for Artificial Intelligence.
- [Mani & Bloedorn 99]
I. Mani and E. Bloedorn. Summarizing similarities and differences among related documents. *Information Retrieval*, 1(1):35–67, 1999.
- [Mani 01]
I. Mani. *Automatic Summarization*. John Benjamins Publishing Company, 2001.

- [Mann & Matthiessen 83]
W. C. Mann and C. M. Matthiessen. Nigël: A systemic grammar for text generation. Technical Report ISI/RR-83-105, Information Sciences Institute, February 1983. 4676 Admiralty Way, Marina del Rey, CA.
- [Mann & Yarowsky 03a]
G. Mann and D. Yarowsky. Unsupervised personal name disambiguation. In *Proceedings of CoNLL*, 2003.
- [Mann & Yarowsky 03b]
G. S. Mann and D. Yarowsky. Unsupervised personal name disambiguation. In W. Daelemans and M. Osborne, editors, *Proceedings of the 7th Conference on Natural Language Learning (CoNLL-2003)*, pages 33–40. Edmonton, Canada, May 2003.
- [Mann 83]
W. C. Mann. An overview of the PENMAN text generation system. In *Proceedings of the National Conference on Artificial Intelligence*, pages 261–265. AAAI, August 1983. Also appears as USC/Information Sciences Institute, RR-83-114.
- [Manna & Waldinger 77]
Z. Manna and R. Waldinger. *Studies in Automatic Programming Logic*. North-Holland, New York, 1977.
- [Manna 74]
Z. Manna. *Mathematical Theory of Computation*. McGraw-Hill, New York, 1974.
- [Manning & Schütze 99a]
C. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing*. MIT press, Cambridge, MA, 1999. Supporting materials available at <http://www.sultry.arts.usyd.edu.au/~fsnlp/>.
- [Manning & Schütze 99b]
C. D. Manning and H. Schütze. Evaluation measures. In *Foundations of statistical natural language processing*, chapter 8.1, pages 267–271. Cambridge, MA, MIT Press, 1999.
- [Manning & Schütze 99c]
C. D. Manning and H. Schütze. *Foundations of statistical natural language processing*. Cambridge, MA, MIT Press, 1999. chapter 10.
- [Manning *et al.* 08]
C. D. Manning, P. Raghavan, and H. Schütze. *Introduction to information retrieval*. Cambridge University Press, New York, NY, 2008.
- [Manning *et al.* 14]
C. D. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. J. Bethard, and D. McClosky. The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 55–60, 2014.
- [Manov & Popov 05]
D. Manov and B. Popov. D2.6.1 Massive Automatic Annotation. Technical report, SEKT EU Project Deliverable, 2005.
- [Manov *et al.* 03]
D. Manov, A. Kiryakov, B. Popov, K. Bontcheva, and D. Maynard. Experiments with geographic knowledge for information extraction. In *Workshop on Analysis of Geographic References, HLT/NAACL’03*, Edmonton, Canada, 2003. <http://gate.ac.uk/sale/hlt03/paper03.pdf>.

- [Manthey & Bry 88]
R. Manthey and F. Bry. Satchmo: A Theorem Prover Implemented in PROLOG. In *Proceedings, 9th International Conference on Automated Deduction (CADE-9)(LNCS-310)*, pages 415–434, 1988.
- [Marcelo *et al.* 10]
M. Marcelo, P. Barbara, and C. Carlos. Twitter under crisis: Can we trust what we rt? In *1st Workshop on Social Media Analytics (SOMA)*, 2010.
- [Marchionini 06]
G. Marchionini. Exploratory search: From finding to understanding. *Commun. ACM*, 49(4):41–46, 2006.
- [Marcu 97]
D. Marcu. From Discourse Structures to Text Summaries. In *The Proceedings of the ACL’97/EACL’97 Workshop on Intelligent Scalable Text Summarization*, pages 82–88, Madrid, Spain, July 11 1997.
- [Marcus *et al.* 93a]
M. Marcus, B. Santorini, and M. Barcinkiewicz. Building a Large Annotated Corpus of English: The Penn Treebank. *Computational Linguistics*, 19(2):313–330, 1993.
- [Marcus *et al.* 93b]
M. Marcus, B. Santorini, and M. A. Marcinkiewicz. Building a large annotated corpus of English: The Penn Tree Bank. *Computational Linguistics*, 19(2):313–330, 1993.
- [Marcus *et al.* 93c]
M. Marcus, B. Santorini, and M. Marcinkiewicz. Building a Large Annotated Corpus of English: The Penn Treebank. *Computational Linguistics*, 19(2):313–330, 1993.
- [Marcus *et al.* 93d]
P. Marcus, Mitchell, B. Santorini, and M. A. Marcinkewitz. Building a large annotated corpus of English: the Penn Treebank. *Computational Linguistics*, 19(2), 1993.
- [Marcus *et al.* 94]
M. P. Marcus, B. Santorini, and M. A. Marcinkiewicz. Building a large annotated corpus of english: The Penn Treebank. *Computational Linguistics*, 19(2):313–330, 1994.
- [Marcus *et al.* 95]
M. Marcus, G. Kim, M. Marcinkiewicz, R. MacIntyre, A. Bies, M. Ferguson, K.Katz, and B. Schasberger. The Penn Treebank: Annotating Predicate Argument Structure. Distributed on The Penn Treebank Release 2 CD-ROM by the Linguistic Data Consortium, 1995.
- [Marcus *et al.* 11]
A. Marcus, M. S. Bernstein, O. Badar, D. R. Karger, S. Madden, and R. C. Miller. TwitInfo: Aggregating and visualizing microblogs for event exploration. In *Proceedings of the 2011 Conference on Human Factors in Computing Systems (CHI)*, pages 227–236, 2011.
- [Marder *et al.* 11]
S. R. Marder, D. G. Daniel, L. Alphs, A. G. Awad, and R. S. E. Keefe. Methodological issues in negative symptom trials. *Schizophrenia Bulletin*, 37(2):250–254, January 2011. Cited by 0010.
- [Marge *et al.* 10]
M. Marge, S. Banerjee, and A. I. Rudnicky. Using the amazon mechanical turk to transcribe and annotate meeting speech for extractive summarization. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, pages 99–107, 2010.

- [Marinova *et al.* 08]
 Z. Marinova, F. Amardeilh, K. Georgiev, and T. Francart. User tools. Technical Report D3.4.1, TAO Project Deliverable, 2008. <http://www.tao-project.eu/resources/publicdeliverables/d3-4-1-final.pdf>.
- [Markel & Gray, Jr. 76]
 J. D. Markel and A. H. Gray, Jr. *Linear prediction of speech*. Springer-Verlag, Berlin-Heidelberg-New York, 1976.
- [Markert & Nissim 02a]
 K. Markert and M. Nissim. Metonymy resolution as a classification task. In *Proceedings of EMNLP2002*, 2002.
- [Markert & Nissim 02b]
 K. Markert and M. Nissim. Towards a corpus annotated for metonymies: the case of location names. In *Proceedings of LREC 2002*, 2002.
- [Markowitz *et al.* 86]
 J. Markowitz, T. Ahlswede, and M. Evens. Semantically significant patterns in dictionary definitions. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, pages 112–119, New York, 1986.
- [Marrero *et al.* 09]
 M. Marrero, S. Sanchez-Cuadrado, J. Lara, and G. Andreadakis. Evaluation of Named Entity Extraction Systems. *Advances in Computational Linguistics, Research in Computing Science*, 41:47–58, 2009.
- [Marsh & Perzanowski 98]
 E. Marsh and D. Perzanowski. Muc-7 evaluation of ie technology: Overview of results. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)*. http://www.itl.nist.gov/iaui/894.02/related_projects/muc/index.html, 1998.
- [Marshall & Bandar 99]
 P. Marshall and Z. Bandar. Working towards connectionist modelling of term formation. In *Proc. of Computer Intelligence Theory and Applications 6th Fuzzy Days International Conference*, Lecture Notes in Computer Science, Dortmund, Germany, 1999. Springer-Verlag.
- [Marshall *et al.* 04]
 A. Marshall, D. McDonald, H. Chen, and W. Chung. EBizPort: Collecting and Analysing Business Intelligence Information. *Journal of the American Society for Information Science and Technology*, 55(10):873–891, 2004.
- [Martelli & Montanari 82]
 A. Martelli and U. Montanari. An Efficient Unification Algorithm. *ACM Transactions on Programming Languages and Systems*, 4(2), April, 1982.
- [Martin & Eklund 99]
 P. Martin and P. Eklund. Embedding knowledge in Web documents. *Computer Networks (Amsterdam, Netherlands: 1999)*, 31(11–16):1403–1419, 1999.
- [Martin 01]
 W. Martin. An archive for all of europe. In *Proceedings of the ACL 2001 Workshop on Sharing Tools and Resources*, pages 11–14, 2001.
- [Martin *et al.* 00]
 S. C. Martin, H. Ney, and C. Hamacher. Maximum entropy language modeling and the smoothing problem. *IEEE Trans. on Speech and Audio Processing*, 8:626–632, September 2000.

- [Martín *et al.* 08a]
 J. Martín, G. Herrero, K. Bontcheva, D. Damjanovic, and F. Cerbah. Case study reports on evaluating the methodology. Technical Report D1.3.1, TAO Project Deliverable, 2008. <http://www.tao-project.eu/resources/publicdeliverables/d1-2.pdf>.
- [Martín *et al.* 08b]
 J. Martín, G. Herrero, A. Capellini, T. Francart, F. Amardeilh, and Z. Marinova. Architecture and integration requirements and specifications. Technical Report D5.2, TAO Project Deliverable, 2008. <http://www.tao-project.eu/resources/publicdeliverables/d5-2.pdf>.
- [Mascaro & Goggins 12]
 C. Mascaro and S. P. Goggins. Twitter as virtual town square: Citizen engagement during a nationally televised republican primary debate. In *APSA 2012 Annual Meeting Paper*, 2012.
- [Mason 98]
 O. Mason. The CUE Corpus Access Tool. In *Workshop on Distributing and Accessing Linguistic Resources*, pages 20–27, Granada, Spain, 1998. <http://www.dcs.shef.ac.uk/~hamish/dalr/>.
- [Masterman 56]
 M. Masterman. The potentialities of a mechanical thesaurus. *Machine Translation*, 1956.
- [Masuch *et al.* 99]
 M. Masuch, B. Freudenberg, B. Ludowici, S. Kreiker, and T. Strothotte. Virtual Reconstruction of Medieval Architecture. In *Proceedings of Eurographics'99*, pages 87–90, September 1999.
- [Mathes 04]
 A. Mathes. Folksonomies: Cooperative classification and communication through shared metadata. In *Computer Mediated Communication, LIS590CMC (Doctoral Seminar)*, Graduate School of Library and Information Science, University of Illinois Urbana-Champaign, 2004.
- [May *et al.* 03]
 J. May, A. Brunstein, P. Natarajan, and R. Weischedel. Rapid Development of Hindi Named Entity Recognition using Conditional Random Fields and Feature Induction. *Special issue of ACM Transactions on Asian Language Information Processing: Rapid Development of Language Capabilities: The Surprise Languages*, 2003.
- [Maybury 95]
 M. Maybury. Research in multimedia parsing and generation. *Artificial Intelligence Review: Special Issue on the Integration of Natural Language and Vision Processing*, 9(2-3):103–127, 1995.
- [Maybury 97]
 M. Maybury, editor. *Intelligent Multimedia Information Retrieval*. MIT Press, Cambridge, MA, 1997.
- [Maybury 01]
 M. Maybury. Personalcasting – Tailored Broadcast News. In *Workshop on Personalization in Future TV, User Modeling 2001*, Sonthofen, Germany, 2001.
- [Mayfield *et al.* 03]
 J. Mayfield, P. McNamee, and C. Piatko. Named Entity Recognition Using Hundreds of Thousands of Features. In *Proceedings of CoNLL-2003*, pages 184–187. Edmonton, Canada, 2003.
- [Maynard & Ananiadou 98a]
 D. Maynard and S. Ananiadou. Acquiring contextual information for term disambiguation. In *Proc. of 1st Workshop on Computational Terminology, Computerm '98*, pages 86–90, Montreal, Canada, 1998.

- [Maynard & Ananiadou 98b]
D. Maynard and S. Ananiadou. Term disambiguation using the UMLS medical thesaurus. In *Proc. of 1st Computational Linguistics UK Research Colloquium (CLUK2)*, University of Sunderland, UK, 1998.
- [Maynard & Ananiadou 98c]
D. Maynard and S. Ananiadou. Term sense disambiguation using a domain-specific thesaurus. In *Proc. of 1st International Conference on Language Resources and Evaluation (LREC)*, pages 681–687, Granada, Spain, 1998.
- [Maynard & Ananiadou 99a]
D. Maynard and S. Ananiadou. Identifying contextual information for term extraction. In *Proc. of 5th International Congress on Terminology and Knowledge Engineering (TKE '99)*, pages 212–221, Innsbruck, Austria, 1999.
- [Maynard & Ananiadou 99b]
D. Maynard and S. Ananiadou. Incorporating linguistic information for multi-word term extraction. In *Proc. of 2nd Computational Linguistics UK Research Colloquium (CLUK2)*, University of Essex, UK, 1999.
- [Maynard & Ananiadou 99c]
D. Maynard and S. Ananiadou. A linguistic approach to context clustering. In *Proc. of Natural Language Processing Pacific Rim Symposium (NLPRS)*, pages 346–351, Beijing, China, 1999.
- [Maynard & Ananiadou 99d]
D. Maynard and S. Ananiadou. Term extraction using a similarity-based approach. In *Recent Advances in Computational Terminology*. John Benjamins, 1999.
- [Maynard & Ananiadou 00a]
D. Maynard and S. Ananiadou. Creating and using domain-specific ontologies for terminological applications. In *Proc. of 2nd International Conference on Language Resources and Evaluation (LREC)*, Athens, Greece, 2000.
- [Maynard & Ananiadou 00b]
D. Maynard and S. Ananiadou. Identifying terms by their family and friends. In *Proc. of 18th International Conference on Computational Linguistics (COLING)*, Saarbrücken, Germany, 2000.
- [Maynard & Ananiadou 00c]
D. Maynard and S. Ananiadou. Identifying terms by their family and friends. In *Proc. of 18th International Conference on Computational Linguistics (COLING)*, Saarbrücken, Germany, 2000.
- [Maynard & Ananiadou 00d]
D. Maynard and S. Ananiadou. Terminological acquaintance: the importance of contextual information in terminology. In *Proc. of NLP2000 Workshop on Computational Terminology for Medical and Biological Applications*, Patras, Greece, 2000.
- [Maynard & Bontcheva 14]
D. Maynard and K. Bontcheva. Natural Language Processing. In J. Lehmann and J. Voelker, editors, *Perspectives of Ontology Learning*. IOS Press, 2014.
- [Maynard & Bontcheva 15]
D. Maynard and K. Bontcheva. Understanding climate change tweets: an open source toolkit for social media. In *Proceedings of EnviroInfo*, Copenhagen, Denmark, 2015.
- [Maynard & Bontcheva 16]
D. Maynard and K. Bontcheva. Challenges of Evaluating Sentiment Analysis Tools on Social Media. In *Proceedings of LREC 2016*, Portoroz, Slovenia, 2016.

- [Maynard & Cunningham 03]
D. Maynard and H. Cunningham. Multilingual Adaptations of a Reusable Information Extraction Tool. In *Proceedings of the Demo Sessions of EACL'03*, Budapest, Hungary, 2003. ACL.
- [Maynard & Funk 11a]
D. Maynard and A. Funk. Automatic detection of political opinions in tweets. In *Proceedings of MSM 2011: Making Sense of Microposts Workshop at 8th Extended Semantic Web Conference*, Heraklion, Greece, 2011.
- [Maynard & Funk 11b]
D. Maynard and A. Funk. Automatic detection of political opinions in tweets. In D. F. Raúl García-Castro and G. Antoniou, editors, *The Semantic Web: ESWC 2011 Selected Workshop Papers, Lecture Notes in Computer Science*. Springer, 2011.
- [Maynard & Greenwood 12]
D. Maynard and M. A. Greenwood. Large Scale Semantic Annotation, Indexing and Search at The National Archives. In *Proceedings of LREC 2012*, Turkey, 2012.
- [Maynard & Greenwood 14]
D. Maynard and M. A. Greenwood. Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis. In *Proceedings of LREC 2014*, Reykjavik, Iceland, 2014.
- [Maynard & Hare 15]
D. Maynard and J. Hare. Entity-based opinion mining from text and multimedia. In *Advances in Social Media Analysis*, pages 65–86. Springer, 2015.
- [Maynard & Scharl 14]
D. Maynard and A. Scharl. Text analytics tools for environmental information extraction. Technical Report D2.2.1, DecarboNet Project Deliverable, 2014.
- [Maynard & Scharl 16]
D. Maynard and A. Scharl. Text analytics tools for environmental information extraction v.2. Technical Report D2.2.2, DecarboNet Project Deliverable, 2016.
- [Maynard 97]
D. G. Maynard. "With". Unpublished M.Sc. thesis, University of Manchester, UK, 1997.
- [Maynard 99]
D. Maynard. *Term Recognition Using Combined Knowledge Sources*. Unpublished PhD thesis, Department of Computing and Mathematics, Manchester Metropolitan University, UK, 1999.
- [Maynard 00]
D. G. Maynard. *Term Recognition Using Combined Knowledge Sources*. Unpublished PhD thesis, Manchester Metropolitan University, UK, 2000.
- [Maynard 03]
D. Maynard. Multi-source and multilingual information extraction. *Expert Update*, 2003.
- [Maynard 05]
D. Maynard. Benchmarking ontology-based annotation tools for the semantic web. In *UK e-Science Programme All Hands Meeting (AHM2005) Workshop on Text Mining, e-Research and Grid-enabled Language Technology*, Nottingham, UK, 2005.
- [Maynard 08]
D. Maynard. Benchmarking textual annotation tools for the semantic web. In *Proc. of 6th International Conference on Language Resources and Evaluation (LREC)*, Marrakech, Morocco, 2008.

- [Maynard 14]
D. Maynard. Challenges in Analysing Social Media. In A. Duşa, D. Nelle, G. Stock, and G. G. Wagner, editors, *Facing the Future: European Research Infrastructures for the Humanities and Social Sciences*. SCIVERO Verlag, Berlin, 2014.
- [Maynard 15]
D. Maynard. Environmental opinion extraction. Technical Report D2.3.1, DecarboNet Project Deliverable, 2015.
- [Maynard 16]
D. Maynard. Environmental opinion extraction v.2. Technical Report D2.3.2, DecarboNet Project Deliverable, 2016.
- [Maynard *et al.* 00]
D. Maynard, H. Cunningham, K. Bontcheva, R. Catizone, G. Demetriou, R. Gaizauskas, O. Hamza, M. Hepple, P. Herring, B. Mitchell, M. Oakes, W. Peters, A. Setzer, M. Stevenson, V. Tablan, C. Ursu, and Y. Wilks. A Survey of Uses of GATE. Technical Report CS-00-06, Department of Computer Science, University of Sheffield, 2000.
- [Maynard *et al.* 01]
D. Maynard, V. Tablan, C. Ursu, H. Cunningham, and Y. Wilks. Named Entity Recognition from Diverse Text Types. In *Recent Advances in Natural Language Processing 2001 Conference*, pages 257–274, Tzigov Chark, Bulgaria, 2001.
- [Maynard *et al.* 02a]
D. Maynard, K. Bontcheva, H. Saggion, H. Cunningham, and O. Hamza. Using a Text Engineering Framework to Build an Extendable and Portable IE-based Summarisation System. In *Proceedings of the ACL Workshop on Text Summarisation*, pages 19–26, Philadelphia, Pennsylvania, 2002. ACM.
- [Maynard *et al.* 02b]
D. Maynard, K. Bontcheva, H. Saggion, H. Cunningham, and O. Hamza. Using a Text Engineering Framework to Build an Extendable and Portable IE-based Summarisation System. In *Proceedings of the ACL Workshop on Text Summarisation*, 2002.
- [Maynard *et al.* 02c]
D. Maynard, H. Cunningham, K. Bontcheva, and M. Dimitrov. Adapting A Robust Multi-Genre NE System for Automatic Content Extraction. In *Proceedings of the Tenth International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA 2002)*, 2002.
- [Maynard *et al.* 02d]
D. Maynard, H. Cunningham, K. Bontcheva, and M. Dimitrov. Adapting a robust multi-genre NE system for automatic content extraction. In *Proceedings of the 10th International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA '02)*, Varna, Bulgaria, Sep 2002.
- [Maynard *et al.* 02e]
D. Maynard, H. Cunningham, and R. Gaizauskas. Named entity recognition at sheffield university. In H. Holmboe, editor, *Nordic Language Technology – Arbog for Nordisk Sprogteknologisk Forskningsprogram 2002-2004*, pages 141–145. Museum Tusculanums Forlag, 2002.
- [Maynard *et al.* 02f]
D. Maynard, V. Tablan, H. Cunningham, C. Ursu, H. Saggion, K. Bontcheva, and Y. Wilks. Architectural Elements of Language Engineering Robustness. *Journal of Natural Language Engineering – Special Issue on Robust Methods in Analysis of Natural Language Data*, 8(2/3):257–274, 2002.
- [Maynard *et al.* 03a]
D. Maynard, K. Bontcheva, and H. Cunningham. From information extraction to content extraction. Submitted to EACL'2003, 2003.

- [Maynard *et al.* 03b]
D. Maynard, K. Bontcheva, and H. Cunningham. Towards a semantic extraction of Named Entities. In *Recent Advances in Natural Language Processing*, Bulgaria, 2003.
- [Maynard *et al.* 03c]
D. Maynard, K. Bontcheva, and H. Cunningham. Towards a semantic extraction of named entities. In Angelova *et al.* [Angelova *et al.* 03], pages 255–261. <http://gate.ac.uk/sale/ranlp03/ranlp03.pdf>.
- [Maynard *et al.* 03d]
D. Maynard, V. Tablan, K. Bontcheva, and H. Cunningham. Rapid customisation of an Information Extraction system for surprise languages. *Special issue of ACM Transactions on Asian Language Information Processing: Rapid Development of Language Capabilities: The Surprise Languages*, 2:295–300, 2003.
- [Maynard *et al.* 03e]
D. Maynard, V. Tablan, K. Bontcheva, H. Cunningham, and Y. Wilks. Muse: a multi-source entity recognition system. *Submitted to Computers and the Humanities*, 2003.
- [Maynard *et al.* 03f]
D. Maynard, V. Tablan, K. Bontcheva, H. Cunningham, and Y. Wilks. Multi-source entity recognition – an information extraction system for diverse text types. Research Memorandum CS–03–02, Department of Computer Science, University of Sheffield, April 2003.
- [Maynard *et al.* 03g]
D. Maynard, V. Tablan, and H. Cunningham. NE recognition without training data on a language you don’t speak. In *ACL Workshop on Multilingual and Mixed-language Named Entity Recognition: Combining Statistical and Symbolic Models*, Sapporo, Japan, 2003.
- [Maynard *et al.* 04a]
D. Maynard, K. Bontcheva, and H. Cunningham. Automatic Language-Independent Induction of Gazetteer Lists. In *Proceedings of 4th Language Resources and Evaluation Conference (LREC’04)*, Lisbon, Portugal, 2004. ELRA.
- [Maynard *et al.* 04b]
D. Maynard, H. Cunningham, A. Kourakis, and A. Kokossis. Ontology-Based Information Extraction in hTechSight. In *First European Semantic Web Symposium (ESWS 2004)*, Heraklion, Crete, 2004.
- [Maynard *et al.* 04c]
D. Maynard, M. Yankova, N. Aswani, and H. Cunningham. Automatic Creation and Monitoring of Semantic Metadata in a Dynamic Knowledge Portal. In *Proceedings of the 11th International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA 2004)*, Varna, Bulgaria, 2004.
- [Maynard *et al.* 05]
D. Maynard, M. Yankova, A. Kourakis, and A. Kokossis. Ontology-based information extraction for market monitoring and technology watch. In *ESWC Workshop "End User Aspects of the Semantic Web"*, Heraklion, Crete, 2005.
- [Maynard *et al.* 06a]
D. Maynard, W. Peters, and Y. Li. Metrics for evaluation of ontology-based information extraction. In *WWW 2006 Workshop on Evaluation of Ontologies for the Web (EON)*, Edinburgh, Scotland, 2006.
- [Maynard *et al.* 06b]
D. Maynard, W. Peters, M. Sabou, and P. Cimiano. Prototypes of language dependent tools for evaluation. Technical Report D2.3.6, KnowledgeWeb Deliverable, 2006.

- [Maynard *et al.* 07a]
D. Maynard, S. Dasiopolou, S. Costache, K. Eckert, H. Stuckenschmidt, M. Dzbor, and S. Handschuh. Benchmarking of annotation tools. Technical Report D1.2.2.1.3, KnowledgeWeb Deliverable, 2007.
- [Maynard *et al.* 07b]
D. Maynard, W. Peters, M. d’Aquin, and M. Sabou. Change management for metadata evolution. In *ESWC International Workshop on Ontology Dynamics (IWOD)*, Innsbruck, Austria, June 2007.
- [Maynard *et al.* 07c]
D. Maynard, W. Peters, M. D’Aquin, M. Sabou, and N. Aswani. Dynamics of metadata. Technical Report D1.5.1, NeOn Project Deliverable, 2007.
- [Maynard *et al.* 07d]
D. Maynard, H. Saggion, M. Yankova, K. Bontcheva, and W. Peters. Natural language technology for information integration in business intelligence. In W. Abramowicz, editor, *10th International Conference on Business Information Systems*, Poland, 25-27 April 2007. <http://gate.ac.uk/sale/bis07/musing-bis07-final.pdf>.
- [Maynard *et al.* 07e]
D. Maynard, H. Saggion, M. Yankova, K. Bontcheva, and W. Peters. Natural Language Technology for Information Integration in Business Intelligence. In *10th International Conference on Business Information Systems (BIS-07)*, Poznan, Poland, 25-27 April 2007.
- [Maynard *et al.* 08a]
D. Maynard, W. Peters, S. Angeletou, and M. D’Aquin. Implementation of metadata evolution. Technical Report D1.5.2, NeOn Project Deliverable, 2008.
- [Maynard *et al.* 08b]
D. Maynard, W. Peters, and Y. Li. Evaluating evaluation metrics for ontology-based applications: Infinite reflection. In *Proc. of 6th International Conference on Language Resources and Evaluation (LREC)*, Marrakech, Morocco, 2008.
- [Maynard *et al.* 08c]
D. Maynard, Y. Li, and W. Peters. NLP Techniques for Term Extraction and Ontology Population. In P. Buitelaar and P. Cimiano, editors, *Bridging the Gap between Text and Knowledge - Selected Contributions to Ontology Learning and Population from Text*. IOS Press, 2008.
- [Maynard *et al.* 09a]
D. Maynard, N. Aswani, W. Peters, F. Zablith, and M. D’Aquin. Advanced methods for change propagation between networked ontologies and metadata. Technical Report D1.5.3, NeOn Project Deliverable, 2009.
- [Maynard *et al.* 09b]
D. Maynard, N. Aswani, W. Peters, F. Zablith, and M. D’Aquin. Advanced methods for change propagation between networked ontologies and metadata. Technical Report D1.5.3, NeOn Project Deliverable, 2009.
- [Maynard *et al.* 09c]
D. Maynard, A. Funk, and W. Peters. NLP-based support for ontology lifecycle development. In *CK 2009 – ISWC Workshop on Workshop on Collaborative Construction, Management and Linking of Structured Knowledge*, Washington, USA, October 2009.
- [Maynard *et al.* 09d]
D. Maynard, A. Funk, and W. Peters. SPRAT: a tool for automatic semantic pattern-based ontology population. In *International Conference for Digital Libraries and the Semantic Web*, Trento, Italy, September 2009.

- [Maynard *et al.* 09e]
D. Maynard, A. Funk, and W. Peters. Using Lexico-Syntactic Ontology Design Patterns for ontology creation and population. In *WOP 2009 – ISWC Workshop on Ontology Patterns*, Washington, USA, October 2009.
- [Maynard *et al.* 10]
D. Maynard, N. Aswani, J. Voelker, S. Schenk, and F. Zablith. Bottom-up evolution of networked ontologies from metadata. Technical Report D1.5.4, NeOn Project Deliverable, 2010.
- [Maynard *et al.* 12]
D. Maynard, K. Bontcheva, and D. Rout. Challenges in developing opinion mining tools for social media. In *Proceedings of @NLP can u tag #usergeneratedcontent?! Workshop at LREC 2012*, Turkey, 2012.
- [Maynard *et al.* 13]
D. Maynard, D. Dupplaw, and J. Hare. Multimodal Sentiment Analysis of Social Media. In *Proceedings of Thirty-third SGA International Conference on Artificial Intelligence (AI-2013 Workshop on Social Media Analysis)*, Cambridge, UK, 2013.
- [Maynard *et al.* 15a]
D. Maynard, M. A. Greenwood, I. Roberts, G. Windsor, and K. Bontcheva. Real-time social media analytics through semantic annotation and linked open data. In *Proceedings of WebSci*, Oxford, UK, 2015.
- [Maynard *et al.* 15b]
D. Maynard, G. Gossen, M. Fisichella, and A. Funk. Should I care about your opinion? Detection of opinion interestingness and dynamics in social media. *Journal of Future Internet*, 2015.
- [Maynard *et al.* 16a]
D. Maynard, K. Bontcheva, and I. Augenstein. *Natural Language Processing for the Semantic Web*. Morgan and Claypool, 2016.
- [Maynard *et al.* 16b]
D. Maynard, J. Petrak, A. Funk, and L. Derczynski. Multilingual content processing methods. Technical Report D3.1, COMRADES Project Deliverable, 2016.
- [Maynard *et al.* ss]
D. Maynard, I. Roberts, M. A. Greenwood, D. Rout, and K. Bontcheva. A Framework for Real-time Semantic Social Media Analysis. *Web Semantics: Science, Services and Agents on the World Wide Web*, In Press.
- [McArthur 81]
T. McArthur. *Longman Lexicon of Contemporary English*. Longman, London, 1981.
- [McCallum *et al.* 00a]
A. McCallum, D. Freitag, and F. Pereira. Maximum Entropy Markov Models for Information Extraction and Segmentation. In *Proceedings of the Seventeenth International Conference on Machine Learning*, San Francisco: Morgan Kaufmann, 2000.
- [McCallum *et al.* 00b]
A. McCallum, D. Freitag, and F. Pereira. Maximum entropy markov models for information extraction and segmentation. In *Proceedings of the Seventeenth International Conference on Machine Learning*, pages 591–598. Citeseer, 2000.
- [McClelland & Rumelhart 81]
J. McClelland and D. Rumelhart. An interactive activation model of context effect in letter perception: Part 1, an account of the basic findings. *Psychological Review*, 88(5):375–407, 1981.

- [McClelland & Rumelhart 86]
 J. McClelland and D. Rumelhart. *Parallel Distributed Processing*. MIT Press, Cambridge, MA, 1986.
- [McCormick 02]
 S. McCormick. Open Lexicon Interchange Format. Technical report, OLIF Consortium, <http://www.olif.net>, 2002.
- [McCreadie *et al.*]
 R. McCreadie, C. Macdonald, and I. Ounis. Identifying top news using crowdsourcing. *Information Retrieval*, pages 1–31. 10.1007/s10791-012-9186-z.
- [McCreadie *et al.* 12]
 R. McCreadie, C. Macdonald, and I. Ounis. Identifying top news using crowdsourcing. *Information Retrieval*, pages 1–31, 2012. 10.1007/s10791-012-9186-z.
- [McCulloch *et al.* 90]
 N. McCulloch, W. Ainsworth, and R. Linggard. Multi-Layer Perceptrons Applied to Speech Technology. In *Speech and Language Processing*. Chapman and Hall, London, 1990.
- [McDermott & Doyle 80]
 D. McDermott and J. Doyle. Non-monotonic logic i. *Artificial Intelligence*, 13, 1980.
- [McDonald & Pereira 05]
 R. McDonald and F. Pereira. Identifying Gene and Protein Mentions in Text Using Conditional Random Fields. *BMC Bioinformatics*, 6(Suppl 1):S6, 2005.
- [McDonald 83]
 D. McDonald. Natural Language Generation as a Computational Problem: an Introduction. In M. Brady and R. Berwick, editors, *Computational Models of Discourse*, Cambridge, MA, 1983. MIT Press.
- [McDonald 93]
 D. D. McDonald. Internal and external evidence in the identification and semantic categorisation of proper names. In *Proceedings of SIGLEX workshop on Acquisition of Lexical Knowledge from Text*, pages 32–43. Ohio, U.S.A., June 1993.
- [McDonald *et al.* 90]
 J. E. McDonald, A. T. Plate, and R. W. Schvaneveldt. Using pathfinder to extract semantic information from text. In R. W. Schvaneveldt, editor, *Pathfinder Associative Networks: Studies in Knowledge Organization*. Ablex, Norwood, NJ, 1990.
- [McDonald *et al.* 04]
 R. T. McDonald, R. S. Winters, M. Mandel, Y. Jin, P. S. White, and F. Pereira. An entity tagger for recognizing acquired genomic variations in cancer literature. *Bioinformatics*, 20(17):3249–3251, 2004.
- [McDowell & Cafarella 06]
 L. K. McDowell and M. Cafarella. Ontology-Driven Information Extraction with OntoSyphon. In *5th Internal Semantic Web Conference (ISWC'06)*. Springer, 2006.
- [McEnery 97]
 A. McEnery. Multilingual Corpora – Current Practice and Future Trends. In *13th ASLIB Machine Translation Conference*, pages 75–86, London, 1997.
- [McEnery *et al.* 99]
 A. McEnery, S. Botley, and A. Wilson, editors. *Multilingual Corpora: Teaching and Research*. Rodopi, Amsterdam, 1999.

- [McEnergy *et al.* 00]
 A. McEnergy, P. Baker, R. Gaizauskas, and H. Cunningham. EMILLE: Building a Corpus of South Asian Languages. *Vivek, A Quarterly in Artificial Intelligence*, 13(3):23–32, 2000.
- [Mcgraw *et al.* 09]
 I. McGraw, A. Gruenstein, and A. Sutherland. A self-labeling speech corpus: Collecting spoken words with an online educational game, 2009.
- [Mcgraw *et al.* 10]
 I. McGraw, C. ying Lee, L. Hetherington, S. Seneff, and J. R. Glass. Collecting voices from the cloud, 2010.
- [McInnes *et al.* 89]
 F. R. McInnes, Y. Ariki, and A. A. Wrench. Enhancement and optimisation of a speech recognition front end based on hidden Markov models. In *Proceedings European Conference on Speech Communication and Technology*, volume 2, pages 461–464, 1989.
- [McKay *et al.* 96]
 D. McKay, J. Pastor, and R. McEntire. An Architecture for Information Agents. In *Proceedings of the 3rd International Conference on AI Planning Systems*, 1996.
- [McKay *et al.* 08]
 J. D. McKay, R. J. Hung, V. Gaborieau, A. Chabrier, G. Byrnes, and et al. Lung cancer susceptibility locus at 5p15.33. *Nature Genetics*, 40(12):1404–1406, 2008.
- [McKelvie & Mikheev 98]
 D. McKelvie and A. Mikheev. Indexing SGML files using LT NSL. LT Index documentation, from <http://www.ltg.ed.ac.uk/>, 1998.
- [McKelvie *et al.* 97a]
 D. McKelvie, C. Brew, and H. Thompson. Using SGML as a Basis for Data-Intensive NLP. In *Proceedings of the fifth Conference on Applied Natural Language Processing (ANLP-97)*, Washington, DC, 1997.
- [McKelvie *et al.* 97b]
 D. McKelvie, C. Brew, and H. Thompson. Using sgml as a basis for data-intensive nlp. In *Fifth Conference on Applied Natural Language Processing (ANLP97)*, Washington, DC, 1997.
- [McKelvie *et al.* 98]
 D. McKelvie, C. Brew, and H. Thompson. Using SGML as a Basis for Data-Intensive Natural Language Processing. *Computers and the Humanities*, 31(5):367–388, 1998.
- [McKelvie *et al.* 01]
 D. McKelvie, A. Isard, A. Mengel, M. Moeller, M. Grosse, and M. Klein. The MATE Workbench - an annotation tool for XML coded speech corpora. *Speech Communication*, 33(1-2):97–112, 2001.
- [McKeown & Radev 95]
 K. R. McKeown and D. R. Radev. Generating summaries of multiple news articles. In *Proceedings, 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 74–82, Seattle, Washington, July 1995.
- [McKeown 85]
 K. R. McKeown. *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*. Cambridge University Press, 1985.
- [McKeown *et al.* 90]
 K. McKeown, M. Elhadad, Y. Fukumuto, J. Lim, C. Lombardi, J. Robin, and F. Smadja. Natural Language Generation in COMET. In R. Dale, C. Mellish, and M. Zock, editors, *Current Research in Natural Language Generation*, London, 1990. Academic Press.

- [McKeown *et al.* 95]
K. R. McKeown, J. Robin, and K. Kukich. Generating concise natural language summaries. *Information Processing & Management*, 31(5):702–733, 1995.
- [McKeown *et al.* 98]
K. McKeown, D. Jordan, and H. V. Generating patient-specific summaries of on-line literature. In *Intelligent Text Summarization. Papers from the 1998 AAAI Spring Symposium. Technical Report SS-98-06*, pages 34–43, Stanford (CA), USA, March 23-25 1998. The AAAI Press.
- [McKeown *et al.* 99]
K. McKeown, J. Klavans, V. Hatzivassiloglou, R. Barzilay, and E. Eskin. Towards multidocument summarization by reformulation: Progress and prospects. In *AAAI/IAAI*, pages 453–460, 1999.
- [Mcnamee & Mayfield 04]
P. Mcnamee and J. Mayfield. Character n-gram tokenization for european language text retrieval. *Information Retrieval*, 7(1):73–97, 2004.
- [McNaught *et al.* 00]
J. McNaught, W. Black, F. Rinaldi, E. Bertino, A. Brasher, D. Deavin, B. Catania, D. Silvestri, B. Armani, A. Persidis, G. Semerano, F. Esposito, V. Candela, G. Zarri, and L. Gilardoni. Integrated document and knowledge management for the knowledge-based enterprise. In *Proceedings of the 3rd International Conference on the practical application of Knowledge Management*. The paractical application company, 2000.
- [McRoy 92a]
S. McRoy. Using multiple knowledge sources for word sense disambiguation. *Computational Linguistics*, 18(1):1–30, 1992.
- [McRoy 92b]
S. W. McRoy. Using multiple knowledge sources for word sense disambiguation. *Computational Linguistics*, 1992.
- [McWhirter 95]
J. McWhirter. VGrep: A Graphical Tool for the Exploration of Textual Documents. In *CHI '95: Proceedings 1995 Conference on Human Factors in Computing Systems*. ACM, 1995.
- [Medlock & Briscoe 07]
B. Medlock and T. Briscoe. Weakly supervised learning for hedge classification in scientific literature. In *Proceedings of the ACL 2007*, 2007.
- [Meier *et al.* 01]
D. Meier, C. Tautz, R. Traphoner, M. Wissen, and J. Ziegler. Building ontologies for knowledge management applications in group sessions. In *Proceedings of K-CAP 2001 Workshop*, Victoria BC, Canada, 2001.
- [Meier *et al.* 14]
F. Meier, D. Elswiler, and M. L. Wilson. More than liking and bookmarking? Towards understanding twitter favouriting behaviour. In *Proceedings of the International AAAI Conference on Weblogs and Social Media*, 2014.
- [Meij *et al.* 12]
E. Meij, W. Weerkamp, and M. de Rijke. Adding semantics to microblog posts. In *Proc. of the Fifth Int. Conf. on Web Search and Data Mining (WSDM)*, 2012.
- [Meili *et al.* 14]
C. Meili, R. Hess, M. Fernandez, and G. Burel. Earth hour report. Technical Report D6.2.1, DecarboNet Project Deliverable, 2014.

- [Mellebeek *et al.* 10]
 B. Mellebeek, F. Benavent, J. Grivolla, J. Codina, M. R. Costa-jussà, and R. Banchs. Opinion mining of spanish customer comments with non-expert annotations on mechanical turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 114–121, 2010.
- [Mellish & Evans 04]
 C. Mellish and R. Evans. Implementation Architectures for Natural Language Generation. *Natural Language Engineering*, 2004.
- [Mellish & Scott 99]
 C. Mellish and D. Scott. Workshop preface. In *Proceedings of the AISB'99 Workshop on Reference Architectures and Data Standards for NLP*. The Society for the Study of Artificial Intelligence and Simulation of Behaviour, April 1999.
- [Mellish *et al.* 94]
 C. Mellish, P. Whitelock, and G. Ritchie. Techniques in natural language processing 1: Module workbook. Unpublished lecture notes used in Department of Artificial Intelligence, 1994.
- [Mellish *et al.* 98]
 C. Mellish, A. Knott, J. Oberlander, and M. O'Donnell. Experiments using Stochastic Search for Text Planning. In *Proceedings of the ninth International Workshop on Natural Language Generation*, pages 98–107, Niagra-on-the-Lake, Ontario, Canada, 1998.
- [Mellish *et al.* 04]
 C. Mellish, D. Scott, L. Cahill, R. Evans, D. Paiva, and M. Reape. A Reference Architecture for Generation Systems. *Natural Language Engineering*, 2004.
- [Melster *et al.* 98]
 R. Melster, A. Diaz, and B. Groth. SCORE-The virtual museum, development of a distributed, object-oriented system for 3D real-time visualization. Technical report, TU Berlin, October 1998.
- [Melucci 08]
 M. Melucci. Towards modeling implicit feedback with quantum entanglement. In *Proceedings of the Second Quantum Interaction Symposium (QI-2008)*. College Publications, 2008.
- [Melucci 10]
 M. Melucci. An investigation of quantum interference in information retrieval. In H. Cunningham, A. Hanbury, and S. Rüger, editors, *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference)*. LNCS volume number: 6107, Lecture Notes in Computer Science, Vienna, Austria, May 2010. Springer.
- [Memodata]
 Memodata. The Dicologique conceptual dictionary. <http://www.memodata.com>, ?
- [Mendelson 79]
 E. Mendelson. *Introduction to Mathematical Logic*. Van Nostrand, New York, 1979.
- [Mendes *et al.* 10a]
 P. N. Mendes, A. Passant, and P. Kapanipathi. Twarql: Tapping into the wisdom of the crowd. In *Proceedings of the 6th International Conference on Semantic Systems, I-SEMANTICS '10*, pages 45:1–45:3, 2010.
- [Mendes *et al.* 10b]
 P. N. Mendes, A. Passant, P. Kapanipathi, and A. P. Sheth. Linked open social signals. In *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT '10*, pages 224–231, Washington, DC, USA, 2010. IEEE Computer Society.

- [Mendes *et al.* 11a]
P. N. Mendes, M. Jakob, A. García-Silva, and C. Bizer. DBpedia Spotlight: Shedding light on the web of documents. In *Proc. of I-SEMANTICS*, pages 1–8, 2011.
- [Mendes *et al.* 11b]
P. N. Mendes, M. Jakob, A. García-Silva, and C. Bizer. DBpedia Spotlight: Shedding light on the web of documents. In *Proceedings of the 7th International Conference on Semantic Systems (I-Semantics)*, 2011.
- [Mendes *et al.* 12]
P. N. Mendes, M. Jakob, and C. Bizer. Dbpedia: A multilingual cross-domain knowledge base. In *LREC*, pages 1813–1817, 2012.
- [Meng]
X. Meng. Tutorial on MapReduce Framework in NLTK. Technical report.
- [Meng *et al.* 04]
H. Meng, J. Shawe-Taylor, S. Szedmak, and J. Farquhar. Support Vector Machine to Synthesis Kernels. Machine Learning Workshop, Sheffield, 2004.
- [Merchant *et al.* 96]
R. Merchant, M. Okurowski, and N. Chinchor. The Multi Lingual Entity Task (MET) Overview. In *Advances in Text Processing – TIPSTER Programme Phase II*. DARPA, Morgan Kaufmann, California, 1996.
- [Merlo 93]
P. Merlo. For an incremental computation of intrasentential coreference. In *the 13th International Joint Conference on Artificial Intelligence*, pages 1216–1221. Morgan Kaufmann, California, 1993.
- [Merriam-Webster 99]
Merriam-Webster. Hypertext Webster Dictionary, 1999. <http://www.fin.gov.nt.ca/webster.htm>.
- [Metropolis *et al.* 53]
N. Metropolis, R. A., M. Rosenbluth, A. Teller, and E. Teller. *J. Chem. Phys*, 21:1087, 1953.
- [Meurs *et al.* 11]
M.-J. Meurs, C. Murphy, N. Naderi, I. Morgenstern, C. Cantu, S. Semarjit, G. Butler, J. Powlowski, A. Tsang, and R. Witte. Towards evaluating the impact of semantic support for curating the fungus scientific literature. In *The 3rd Canadian Semantic Web Symposium (CSWS2011)*, Vancouver, British Columbia, Canada, 08/2011 2011.
- [Meyer & Gurevych 12]
C. M. Meyer and I. Gurevych. Wiktionary: a new rival for expert-built lexicons? Exploring the possibilities of collaborative lexicography. In *Electronic Lexicography*. Oxford University Press, 2012.
- [Meyer 01]
I. Meyer. Extracting knowledge-rich contexts for terminography. In *Recent Advances in Computational Terminology*, pages 279–302. John Benjamins, 2001.
- [Meyer *et al.* 99]
I. Meyer, K. Mackintosh, C. Barrière, and T. Morgan. Conceptual sampling for terminographical corpus analysis. In *Proc. of 5th International Congress on Terminology and Knowledge Engineering (TKE '99)*, pages 256–267, Innsbruck, Austria, 1999.
- [Meyer *et al.* 11]
B. Meyer, K. Bryan, Y. Santos, and B. Kim. TwitterReporter: Breaking news detection and visualization through the geo-tagged Twitter network. In *Proceedings of the ISCA 26th International Conference on Computers and Their Applications*, pages 84–89, 2011.

- [Meystre *et al.* 08]
 S. Meystre, G. Savova, K. Kipper-Schuler, and J. Hurdle. Extracting information from textual documents in the electronic health record: A review of recent research. *IMIA Yearbook of Medical Informatics*, pages 128–144, 2008.
- [Michalski 83]
 R. Michalski. A theory and methodology of inductive learning. In J. Carbonell, R. Michalski, and T. Mitchell, editors, *Machine Learning: An Artificial Intelligence Approach*. Tioga Pub. Co., Palo Alto, CA, 1983.
- [Michelson & Knoblock 07]
 M. Michelson and C. Knoblock. Unsupervised information extraction from unstructured, ungrammatical data sources on the World Wide Web. *International Journal on Document Analysis and Recognition*, 10(3):211–226, 2007.
- [Michelson & Macskassy 10]
 M. Michelson and S. A. Macskassy. Discovering users’ topics of interest on Twitter: A first look. In *Proceedings of the Fourth Workshop on Analytics for Noisy Unstructured Text Data*, AND ’10, pages 73–80, 2010.
- [Michiels & Noel 82]
 A. Michiels and J. Noel. Approaches to Thesaurus Production. In *Proceedings of the 9th International Conference on Computational Linguistics (COLING-82)*, pages 227–232, Prague, Czechoslovakia, 1982.
- [Michiels *et al.* 80]
 A. Michiels, J. Mullenders, and J. Noel. Exploiting a Large Data Base by Longman. In *Proceedings of the 8th International Conference on Computational Linguistics (COLING-80)*, pages 374–382, Tokyo, Japan, 1980.
- [Mihalcea & Tarau 04a]
 R. Mihalcea and P. Tarau. TextRank: Bringing order into text. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 404–411, 2004.
- [Mihalcea & Tarau 04b]
 R. Mihalcea and P. Tarau. TextRank: Bringing order into text. In *EMNLP*, 2004.
- [Mika 07]
 P. Mika. Ontologies are us: A unified model of social networks and semantics. *Journal of Web Semantics*, 5(1):5–15, 2007.
- [Mikheev & Finch 97]
 A. Mikheev and S. Finch. A Workbench for Finding Structure in Text. In *Fifth Conference on Applied NLP (ANLP-97)*, Washington, DC, 1997.
- [Mikheev *et al.* 99a]
 A. Mikheev, C. Grover, and M. Moens. XML Tools and Architecture for Named Entity Recognition. *Journal of Markup Languages: Theory and Practice*, 1(3):89–113, 1999.
- [Mikheev *et al.* 99b]
 A. Mikheev, M. Moens, and C. Grover. Named Entity recognition without gazetteers. In *Proceedings of the Ninth Conference of the European Chapter of the Association for Computational Linguistics (EACL’99)*, pages 1–8, 1999.
- [Mikheev *et al.* 99c]
 A. Mikheev, M. Moens, and C. Grover. Named Entity Recognition without Gazetteers. In *Proceedings of EACL*. Bergen, Norway, 1999.

- [Miller 90]
G. A. Miller. WordNet: An on-line lexical database. *International Journal of Lexicography*, 3(4):235–312, 1990.
- [Miller 95]
G. Miller. WordNet: a Lexical Database for English. *Communications of the ACM*, Volume 38(Number 11), November 1995.
- [Miller *et al.* 90a]
G. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. Miller. Five papers on WordNet. Technical Report CSL Report 43, Cognitive Science Laboratory. Princeton University, 1990.
- [Miller *et al.* 90b]
G. A. Miller, R. Beckwith, C. Felbaum, D. Gross, G. A. Miller, C. Miller, R. Beckwith, C. Felbaum, D. Gross, and M. Miller, C. Minsky. Five papers on WordNet. 1990.
- [Miller *et al.* 94a]
S. Miller, R. Bobrow, R. Ingria, and R. Schwartz. Hidden Understanding Models of Natural Language. In *Proceedings of the Association of Computational Linguistics*, pages 25–32, June 1994.
- [Miller *et al.* 94b]
S. Miller, R. Bobrow, R. Schwartz, and R. Ingria. Statistical Language Processing using Hidden Understanding Models. In *Proceedings of the Human Language Technology Workshop*, pages 278–282, March 1994.
- [Miller *et al.* 95a]
E. Miller, M. Kado, M. Hirakawa, and T. Ichikawa. HI-VISUAL as a User-Customizable Visual Programming Environment. In *Proceedings VL’95 11th International IEEE Symposium on Visual Languages, Darmstadt*. IEEE Computer Society Press, 1995.
- [Miller *et al.* 95b]
S. Miller, M. Bates, R. Bobrow, R. Ingria, J. Makhoul, and R. Schwartz. Recent progress in hidden understanding models. In *Proceedings of the Spoken Language Systems Workshop*, pages 276–279, Austin TX, 1995.
- [Miller *et al.* 98a]
D. Miller, T. Leek, and R. Schwartz. BBN at TREC7: Using Hidden Markov Models for Information Retrieval. In *NIST Special Publication 500-242: The Seventh Text REtrieval Conference (TREC 7)*, page 133ff, BBN Technologies, Cambridge MA, USA, November 1998.
- [Miller *et al.* 98b]
S. Miller, M. Crystal, H. Fox, L. Ramshaw, R. Schwartz, R. Stone, and R. Weischedel. Description of the BBN System Used for MUC-7. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)*. http://www.itl.nist.gov/iaui/894.02/related_projects/muc/index.html, 1998.
- [Milne & Witten 08a]
D. Milne and I. H. Witten. Learning to link with Wikipedia. In *Proc. of the 17th Conf. on Information and Knowledge Management (CIKM)*, pages 509–518, 2008.
- [Milne & Witten 08b]
D. Milne and I. H. Witten. An effective, low-cost measure of semantic relatedness obtained from wikipedia links. In *In Proceedings of AAAI 2008*, 2008.
- [Milosavljevic 99]
M. Milosavljevic. *The Automatic Generation of Comparison in Descriptions of Entities*. Unpublished PhD thesis, Macquarie University, 1999.

- [Milosavljevic *et al.* 96]
M. Milosavljevic, A. Tulloch, and R. Dale. Text Generation in a Dynamic Hypertext Environment. In *Proceedings of 19th Australian Computer Science Conference*, Melbourne, 1996.
- [Mima *et al.* 99]
H. Mima, S. Ananiadou, and J. Tsujii. A web-based integrated knowledge mining aid system using term-oriented natural language processing. In *Proc. of Natural Language Processing Pacific Rim Symposium (NLPRS)*, Beijing, China, 1999. 13-18.
- [Minker *et al.* 99]
W. Minker, A. Waibel, and J. Mariani. *Stochastically-Based Semantic Analysis*. Kluwer Academic Publishers, Dordrecht, 1999.
- [Minock 05]
M. Minock. A Phrasal Approach to Natural Language Interfaces over Databases. In *Proceedings of the 10th International Conference on Applications of Natural Language to Information Systems (NLDB-2005)*, pages 333–336, June 2005.
- [Minsky 75]
M. Minsky. A Framework for Representing Knowledge. In P. H. Winston, editor, *The Psychology of Computer Vision*, pages 211–277. McGraw-Hill, London, 1975.
- [Minsky 85]
M. Minsky. A framework for representing knowledge. In R. Brachman and H. Levesque, editors, *Readings in Knowledge Representation*. Morgan Kaufmann, California, Los Altos CA, 1985.
- [Mishne 06]
G. Mishne. AutoTag: A collaborative approach to automated tag assignment for weblog posts. In *Proceedings of the 15th International Conference on World Wide Web*, pages 953–954, 2006.
- [Mitamura & Nyberg 95]
T. Mitamura and E. Nyberg. Controlled English for Knowledge-Based MT: Experience with the KANT System. In *TMI-95*, 1995.
- [Mitamura *et al.* 14]
T. Mitamura, E. Hovy, and M. Palmer, editors. *Proceedings of the Second Workshop on EVENTS: Definition, Detection, Coreference, and Representation*. Association for Computational Linguistics, Baltimore, Maryland, USA, June 2014.
- [Mitchell 97a]
B. Mitchell. Named Entity Recognition in German: the identification and classification of certain proper names. Unpublished M.Sc. thesis, Dept. of Computer Science, University of Sheffield, 1997. <http://www.dcs.shef.ac.uk/campus/dcscd/projects/bm.pdf>.
- [Mitchell 97b]
T. Mitchell. *Machine Learning*. McGraw-Hill, New York, 1997.
- [Mitchell 97c]
T. Mitchell. *Machine Learning*. McGraw-Hill International Editions, 1997.
- [Mithun & Kosseim 09]
S. Mithun and L. Kosseim. Summarizing blog entries versus news texts. In *Proceedings of the Workshop on Events in Emerging Text Types*, pages 1–8, Borovets, Bulgaria, September 2009.
- [Mitkov 95]
R. Mitkov. Language Engineering on the Highway: New Perspectives for the Multilingual Society. In *Proceedings of NLPRS*, 1995.

- [Mitkov 96]
R. Mitkov. Language Engineering: towards a clearer picture. In *Proceedings of the International Conference on Mathematical Linguistics (ICML '96)*, 1996.
- [Mitkov 98]
R. Mitkov. Robust Anaphora Resolution with Limited Knowledge. In *Proceedings of COLING'98/ACL'98*, 1998.
- [MITRE 02]
MITRE. Galaxy Communicator. <http://communicator.sourceforge.net/>, 2002.
- [Miyao *et al.* 06]
Y. Miyao, T. Ohta, K. Masuda, Y. Tsuruoka, K. Yoshida, T. Ninomiya, and J. Tsujii. Semantic Retrieval for the Accurate Identification of Relational Concepts in Massive Textbases. In *Proceedings of COLING/ACL 2006*, pages 1017–1024, 2006.
- [Mizerski 82]
R. Mizerski. An attribution explanation of the disproportionate influence of unfavourable information. *Journal of Consumer Research*, 9:301–310, 1982.
- [Moen 98]
W. E. Moen. The CIMI Profile Release 1.0H – A Z39.50 Profile for Cultural Heritage Information. Technical report, Consortium for the Computer Interchange of Museum Information (CIMI), University of North Texas, Denton, TX 76201, November 1998.
- [Moens *et al.* 14]
M.-F. Moens, J. Li, and T.-S. Chua. *Mining user generated content*. CRC Press, 2014.
- [Moghaddam & Ester 13]
S. Moghaddam and M. Ester. The flda model for aspect-based opinion mining: addressing the cold start problem. In *Proceedings of the 22nd international conference on World Wide Web*, pages 909–918. International World Wide Web Conferences Steering Committee, 2013.
- [Moghaddam & Popowich 10]
S. Moghaddam and F. Popowich. Opinion polarity identification through adjectives. *CoRR*, abs/1011.4623, 2010.
- [Mohler & Mihalcea 09]
M. Mohler and R. Mihalcea. Text-to-text semantic similarity for automatic short answer grading. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, EACL '09, pages 567–575, Stroudsburg, PA, USA, 2009. Association for Computational Linguistics.
- [Moldovan *et al.* 02]
D. Moldovan, S. Harabagiu, R. Girju, P. Morarescu, F. Lacatusu, A. Novischi, A. Badulescu, and O. Bolohan. Lcc tools for question answering. In E. M. Voorhees and L. P. Buckland, editors, *The Eleventh Text Retrieval Conference (TREC 2002)*. Department of Commerce, National Institute of Standards and Technology, 2002.
- [Moldovan *et al.* 07]
D. Moldovan, C. Clark, and M. Bowden. Lymba's PowerAnswer 4 in TREC 2007. In E. M. Voorhees and L. P. Buckland, editors, *Proceedings of the The Sixteenth Text Retrieval Conference (TREC 2007)*, Gaithersburg, Maryland, US, 2007.
- [Moltedo & Vitulano 98]
L. Moltedo and D. Vitulano. An Automatic Classification of Materials Degradation. In *Proceedings of WSCG'98*, Plzen, Chzeck Republic, February 1998.

- [Montague 74]
 R. Montague. The proper treatment of quantification in ordinary english. In R. Thomason, editor, *Formal Philosophy: Selected Papers of Richard Montague*. Yale University Press, New Haven, Conn., 1974.
- [Montemagni & Vanderwende 93]
 S. Montemagni and L. Vanderwende. Structural patterns vs. string patterns for extracting semantic information from dictionaries. In K. Jensen, G. Heidhorn, and S. Richardson, editors, *Natural Language Processing: the PLNLP Approach*, chapter 12, pages 149–159. Kluwer Academic Publishers, Dordrecht, 1993.
- [Montemagni 95]
 S. Montemagni. *Subject and Object in Italian Sentence Processing*. Unpublished PhD thesis, UMIST, Manchester, UK, 1995.
- [Montes 01]
 J. Montes. Consumer entertainment software - industry trends. In B. Stanford-Smith and E. Chozza, editors, *E-Work and E-Commerce*, pages –7. IOS Press, Amsterdam, 2001.
- [Moore & Dowding 91]
 R. Moore and J. Dowding. Efficient bottom-up parsing. In *Proceedings of the DARPA Speech and Natural Language Workshop*, pages 200–203, 1991.
- [Moore & Paris 93]
 J. D. Moore and C. L. Paris. Planning texts for advisory dialogs: capturing intentional and rhetorical information. *Computational Linguistics*, 19(4):651 – 694, December 1993.
- [Moore & Swartout 91]
 J. D. Moore and W. R. Swartout. A reactive approach to explanation: Taking the user’s feedback into account. In C. L. Paris, W. R. Swartout, and W. C. Mann, editors, *Natural language generation in artificial intelligence and computational linguistics*. Kluwer Academic Publishers, July 1991. Presented at the Fourth International Workshop on Natural Language Generation. Santa Catalina Island, CA, July, 1988.
- [Moore 85a]
 R. Moore. A formal theory of knowledge and action. In J. Hobbs and R. Moore, editors, *Formal Theories of the Common Sense World*. Ablex Pub. Corp., Norwood, N.J., 1985.
- [Moore 85b]
 R. Moore. The role of logic in knowledge representation and commonsense reasoning. In R. Brachman and H. Levesque, editors, *Readings in Knowledge Representation*. Morgan Kaufmann, California, Los Altos, CA, 1985.
- [Moore 95a]
 G. Moore. *Crossing the Chasm: Marketing and Selling High-Tech Products to Mainstream Customers*. HarperCollins, New York, 1995.
- [Moore 95b]
 J. D. Moore. *Participating in Explanatory Dialogues*. MIT Press, Cambridge, MA, 1995.
- [Morel-Guillemaz *et al.* 90]
 A. Morel-Guillemaz, R. Baud, and J. Scherrer. Proximity processing of medical texts. In *Proceedings of MIE 1990*, pages 625–630. Springer-Verlag, 1990.
- [Morgan 96]
 R. Morgan. An architecture for user defined information extraction. Technical Report 8/96, dept. Computer Science, University of Durham, 1996.

- [Morgan *et al.* 95]
 R. Morgan, R. Garigliano, P. Callaghan, S. Poria, M. Smith, A. Urbanowicz, R. Collingham, M. Costantino, and C. Cooper. Description of the LOLITA System as used for MUC-6. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, pages 71–86, San Francisco, 1995. Morgan Kaufmann, California.
- [Morgan *et al.* 03]
 A. Morgan, L. Hirschman, A. Yeh, and M. Colosimo. Gene Name Extraction Using FlyBase Resources. In *Proc. of ACL 2003 Workshop on Natural Language Processing in Biomedicine*, pages 1–8, Sapporo, Japan, 2003.
- [Morgan *et al.* 04]
 A. A. Morgan, L. Hirschman, M. Colosimo, A. S. Yeh, and J. B. Colombe. Gene name identification and normalization using a model organism database. *J. of Biomedical Informatics*, 37(6):396–410, 2004.
- [Morik *et al.* 99]
 K. Morik, P. Brockhausen, and T. Joachims. Combining statistical learning with a knowledgebased approach - a case study in intensive care monitoring. In *Proceedings of the 16th International Conference on Machine Learning (ICML-99)*, pages 268–277, San Francisco, 1999.
- [Morris & Hirst 91]
 J. Morris and G. Hirst. Lexical cohesion computed by thesaural relations as an indicator of the structure of text. *Computational Linguistics*, 17(1):21–48, 1991.
- [Morris *et al.* 12]
 R. R. Morris, M. Dontcheva, and E. M. Gerber. Priming for better performance in microtask crowdsourcing environments. *Internet Computing, IEEE*, 16(5):13–19, 2012.
- [Morris *et al.* 13]
 R. R. Morris, M. Dontcheva, A. Finkelstein, and E. Gerber. Affect and creative performance on crowdsourcing platforms. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, pages 67–72. IEEE, 2013.
- [Moschitti *et al.* 08]
 A. Moschitti, D. Pighin, and R. Basili. Tree kernels for semantic role labeling. *Computational Linguistics*, 34(2):193–224, 2008.
- [Mostafa 13]
 M. M. Mostafa. More than words: Social networks’ text mining for consumer brand sentiments. *Expert Systems with Applications*, 40(10):4241 – 4251, 2013.
- [Motta *et al.* 99]
 E. Motta, S. B. Shum, and J. Domingue. Case studies in ontology-driven document enrichment. In *Proceedings of the 12th Banff Knowledge Acquisition Workshop*, Banff, Alberta, Canada, 1999.
- [Motta *et al.* 02]
 E. Motta, M. Vargas-Vera, J. Domingue, M. Lanzoni, A. Stutt, and F. Ciravegna. MnM: Ontology Driven Semi-Automatic and Automatic Support for Semantic Markup. In *13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02)*, pages 379–391, Siguenza, Spain, 2002.
- [Mountain Data Systems 88]
 Mountain Data Systems. Acronym finder. <http://www.AcronymFinder.com>, 1988.
- [Mouslim 86]
 M. Mouslim. *Conception et R’realisation du Noyau d’un syst’eme de gestion de banques de donn’ees terminologiques (Syst’eme CEZEAU)*. Unpublished PhD thesis, Universit’e de Clermont II, France, 1986.

- [Muggleton 94]
S. Muggleton. Recent advances in inductive logic programming. In *Proc. 7th Annual ACM Workshop on Computer Learning Theory*, pages 3–11, New York, 1994. ACM Press.
- [Muggleton *et al.* 97]
S. Muggleton, J. Cussens, D. Page, and A. Srinivasan. Using inductive logic programming for natural language processing. In *Proceedings of in ECML*, Workshop Notes on Empirical Learning of Natural Language, pages 25–34. Springer-Verlag, 1997.
- [Muljadi & Hideaki 05]
H. Muljadi and T. Hideaki. Semantic wiki as an integrated content and metadata management system. In *Poster Session at the International Semantic Web Conference (ISWC'05)*, 2005.
- [Mullaly *et al.* 10]
A. Mullaly, C. Gagné, T. Spalding, and K. Marchak. Examining ambiguous adjectives in adjective-noun phrases: Evidence for representation as a shared core-meaning. *The Mental Lexicon*, 5(1):87–114, 2010.
- [Müller & Strube 06]
C. Müller and M. Strube. Multi-level annotation of linguistic data with MMAX2. In S. Braun, K. Kohn, and J. Mukherjee, editors, *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*, pages 197–214. Peter Lang, Frankfurt a.M., Germany, 2006.
- [Müller 06]
C. Müller. Representing and accessing multi-level annotations in mmax2. In *Proceedings of the 5th Workshop on NLP and XML: Multi-Dimensional Markup in Natural Language Processing*, pages 73–76, 2006.
- [Müller *et al.* 04a]
H.-M. Müller, E. E. Kenny, and P. W. Sternberg. Textpresso: An ontology-based information retrieval and extraction system for biological literature. *PLoS Biol*, 2(11):e309, 09 2004.
- [Müller *et al.* 04b]
H.-M. Müller, E. E. Kenny, and P. W. Sternberg. Textpresso: An ontology-based information retrieval and extraction system for biological literature. *PLoS Biol*, 2(11):e309, 09 2004.
- [Müller *et al.* 10]
B. Müller, R. Klinger, H. Gurulingappa, H.-T. Mevissen, M. Hofmann-Apitius, J. Fluck, and C. Friedrich. Abstracts versus full texts and patents: A quantitative analysis of biomedical entities. In H. Cunningham, A. Hanbury, and S. Rüger, editors, *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference)*. LNCS volume number: 6107, Lecture Notes in Computer Science, Vienna, Austria, May 2010. Springer.
- [Munro *et al.* 10]
R. Munro, S. Bethard, V. Kuperman, V. T. Lai, R. Melnick, C. Potts, T. Schnoebelen, and H. Tily. Crowdsourcing and language studies: the new generation of linguistic data. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 122–130, 2010.
- [Murata *et al.* 07]
M. Murata, T. Kanamaru, T. Shirado, and H. Isahara. Using the k-nearest neighbor method and smart weighting in the patent document categorization subtask at ntcir-6. In *Proceedings of the Sixth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access*, pages 407–413, 2007.
- [Murnane *et al.* 13]
E. L. Murnane, B. Haslhofer, and C. Lagoze. Reslve: leveraging user interest to improve entity

- disambiguation on short text. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 1275–1284, 2013.
- [Murveit *et al.* 93]
H. Murveit, J. Butzberger, V. Digalakis, and M. Weintraub. Large-vocabulary dictation using SRI’s DECIPHER speech recognition system: Progressive search techniques. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 2, pages 319–322, 1993.
- [Naaman *et al.* 10a]
M. Naaman, J. Boase, and C. Lai. Is it really about me?: Message content in social awareness streams. In *Proceedings of the 2010 ACM conference on Computer Supported Cooperative Work*, pages 189–192. ACM, 2010.
- [Naaman *et al.* 10b]
M. Naaman, J. Boase, and C. Lai. Is it really about me?: Message content in social awareness streams. In *ACM CSCW*, 2010.
- [Naaman *et al.* 10c]
M. Naaman, J. Boase, and C.-H. Lai. Is it really about me?: message content in social awareness streams. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work, CSCW ’10*, pages 189–192, New York, NY, USA, 2010. ACM.
- [Nádas 84]
A. Nádas. Estimation of probabilities in the language model of the IBM speech recognition system. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32:859–861, 1984.
- [Nádas 85]
A. Nádas. On Turing’s formula for word probabilities. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33:1414–1416, 1985.
- [Nadeau & Sekine 07]
D. Nadeau and S. Sekine. A survey of named entity recognition and classification. *Linguisticae Investigationes*, 30(1):3–26, 2007.
- [Naderi *et al.* 11]
N. Naderi, T. Kappler, C. Baker, and R. Witte. OrganismTagger: Detection, normalization, and grounding of organism entities in biomedical documents. *Bioinformatics*, 27(19):2721–2729, 2011.
- [Nadkarni *et al.* 11]
P. M. Nadkarni, L. Ohno-Machado, and W. W. Chapman. Natural language processing: an introduction. *J Am Med Inform Assoc*, 18(5):544–551, 2011.
- [Nagarajan & Gamon 11]
M. Nagarajan and M. Gamon, editors. *LSM ’11: Proceedings of the Workshop on Languages in Social Media*, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics.
- [Nagarajan *et al.* 09]
M. Nagarajan, K. Gomadam, A. Sheth, A. Ranabahu, R. Mutharaju, and A. Jadhav. Spatio-temporal-thematic analysis of citizen sensor data: Challenges and experiences. In *Web Information Systems Engineering*, pages 539–553, 2009.
- [Nakagawa & Mori 98]
H. Nakagawa and T. Mori. Nested collocation and compound noun for term extraction. In *Proc. of 1st Workshop on Computational Terminology, Computerm ’98*, pages 64–70, Montreal, Canada, 1998.

- [Nakagawa *et al.* 01]
T. Nakagawa, T. Kudoh, and Y. Matsumoto. Unknown Word Guessing and Part-of-Speech Tagging Using Support Vector Machines. In *Proceedings of the Sixth Natural Language Processing Pacific Rim Symposium*, 2001.
- [Nakayama 08]
K. Nakayama. Wikipedia mining for triple extraction enhanced by co-reference resolution. In *SDOW 2008 - "Social Data on the Web" Workshop at ISWC 2008*, Karlsruhe, Germany, October 2008.
- [Namaki *et al.* 13]
M. Namaki, B. Bakhshinategh, M. Noorhoseini, and M. Dehghan. Expressmind: Recommending contents in an anonymous social network. In *Computer and Knowledge Engineering (ICCKE), 2013 3th International eConference on*, pages 278–282, Oct 2013.
- [Nanba & Okumura 00]
H. Nanba and M. Okumura. Producing more readable extracts by revising them. In *Proceedings of the 18th Conference on Computational Linguistics - Volume 2*, COLING '00, pages 1071–1075, Stroudsburg, PA, USA, 2000. Association for Computational Linguistics.
- [Narayanan *et al.* 09]
R. Narayanan, B. Liu, and A. Choudhary. Sentiment analysis of conditional sentences. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1*, pages 180–189. Association for Computational Linguistics, 2009.
- [Nat97]
National Library of Medicine, U.S. Dept. of Health and Human Services. *UMLS Knowledge Sources*, 8th edition, January 1997.
- [Nat98]
National Library of Medicine, U.S. Dept. of Health and Human Services. *UMLS Knowledge Sources*, 9th edition, January 1998.
- [Naveed *et al.* 11]
N. Naveed, T. Gottron, J. Kunegis, and A. C. Alhadi. Bad news travel fast: A content-based analysis of interestingness on twitter. In *Proceedings of the 3rd International Web Science Conference*, WebSci '11, pages 8:1–8:7, New York, NY, USA, 2011. ACM.
- [Navigli & Velardi 04]
R. Navigli and P. Velardi. Learning Domain Ontologies from Document Warehouses and Dedicated Websites. *Computational Linguistics*, 4(2), 2004.
- [Nebel 90]
B. Nebel. *Reasoning and Revision in Hybrid Representation Systems*. Springer-Verlag, Berlin, 1990.
- [Neff *et al.* 03]
M. S. Neff, R. J. Byrd, and B. K. Boguraev. The talent system: Textract architecture and data model. In *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 1–8, 2003.
- [Neff *et al.* 04]
M. S. Neff, R. J. Byrd, and B. K. Boguraev. The Talent System: Textract Architecture and Data Model. *Natural Language Engineering*, 2004.
- [Negri & Mehdad 10]
M. Negri and Y. Mehdad. Creating a bi-lingual entailment corpus through translations with mechanical turk: \$100 for a 10-day rush. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 212–216, 2010.

- [Negri *et al.* 11]
M. Negri, L. Bentivogli, Y. Mehdad, D. Giampiccolo, and A. Marchetti. Divide and conquer: crowd-sourcing the creation of cross-lingual textual entailment corpora. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 670–679, 2011.
- [Nelson & McEvoy 07]
D. Nelson and C. McEvoy. Entangled Associative Structures and Context. In *Proceedings of the AAAI Spring Symposium on Quantum Interaction*. AAAI Press, 2007.
- [Nelson 97]
T. Nelson. Embedded Markup Considered Harmful. In D. Connolly, editor, *XML: Principles, Tools and Techniques*, pages 129–134. O’Reilly, Cambridge, MA, 1997.
- [Nelson *et al.* 95]
S. Nelson, N. Olson, L. Fuller, M. Tuttle, W. Cole, and D. Sherertz. Identifying concepts in medical knowledge. In *Proc. of 8th World Congress on Medical Informatics (MEDINFO)*, pages 33–36, 1995.
- [Nenkova & McKeown 11]
A. Nenkova and K. McKeown. Automatic summarization. *Foundations and Trends in Information Retrieval*, 5(2-3):103–233, 2011.
- [Nenkova & McKeown 12]
A. Nenkova and K. McKeown. A survey of text summarization techniques. In C. C. Aggarwal and C. Zhai, editors, *Mining Text Data*, pages 43–76. Springer US, 2012.
- [Nenkova & Passonneau 04]
A. Nenkova and R. Passonneau. Evaluating content selection in summarization: The pyramid method. In *Proceedings of HLT-NAACL*, pages 145–152, 2004.
- [Neto *et al.* 02]
J. Neto, A. Freitas, and C. Kaestner. Automatic text summarization using a machine learning approach. In G. Bittencourt and G. Ramalho, editors, *Advances in Artificial Intelligence*, volume 2507 of *Lecture Notes in Computer Science*, pages 205–215. Springer Berlin Heidelberg, 2002.
- [Netter & Pianesi 97]
K. Netter and F. Pianesi. Preface. In *Proceedings of the Workshop on Computational Environments for Grammar Development and Linguistic Engineering*, pages iii–v, Madrid, 1997.
- [Netter 98]
K. Netter. Pop-eye and olive. human language as the medium for cross-lingual multimedia information retrieval. Technical report, Language Technology Lab. DFKI GmbH., 1998.
- [Neubig *et al.* 11]
G. Neubig, Y. Matsubayashi, M. Hagiwara, and K. Murakami. Safety information mining – what can NLP do in a disaster. In *Proceedings of 5th International Joint Conference on Natural Language Processing*, pages 965–973. Asian Federation of Natural Language Processing, 2011.
- [Neumann *et al.* 97]
G. Neumann, R. Backofen, J. Baur, M. Becker, and C. B. un. An information extraction core system for real world german text processing. In *Proceedings of the 5th Conference on Applied Natural Language Processing, ANLP-97*, pages 209–216, 1997.
- [Neumann *et al.* 00]
G. Neumann, C. Braun, and J. Piskorski. A divide-and-conquer strategy for shallow parsing of german free texts. In *Proceedings of the 6th Conference on Applied Natural Language Processing, ANLP-00*, 2000.

- [Newbold *et al.* 10]
 N. Newbold, H. McLaughlin, and L. Gillam. Rank by readability: Document weighting for information retrieval. In H. Cunningham, A. Hanbury, and S. Rüger, editors, *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference)*. LNCS volume number: 6107, Lecture Notes in Computer Science, Vienna, Austria, May 2010. Springer.
- [Newell 90]
 A. Newell. *Unified Theories of Cognition*. Harvard University Press, Cambridge, MA, 1990.
- [Newell *et al.* 98]
 A. Newell, S. Langer, and M. Hickey. The role of natural language processing in alternative and augmentative communication. *Journal of Natural Language Engineering*, 4(1):1–17, 1998.
- [Newman *et al.* 08]
 A. Newman, Y.-F. Li, and J. Hunter. A Scale-Out RDF Molecule Store for Improved Co-Identification, Querying and Inferencing. In *The 4th International Workshop on Scalable Semantic Web knowledge Base Systems (SSWS) 2008*, Karlsruhe, Germany, 2008.
- [Ney & Aubert 96]
 H. Ney and X. Aubert. Dynamic programming search: from digit strings to large vocabulary word graphs. In C.-H. Lee, F. K. Soong, and K. K. Paliwal, editors, *Automatic speech and speaker recognition: Advanced topics*, chapter 16, pages 385–412. Kluwer Academic, Amsterdam, 1996.
- [Ney & Ortmanns 00]
 H. Ney and S. Ortmanns. Progress in dynamic programming search for lvcsr. *Proc. of the IEEE*, 88:1224–1240, August 2000.
- [Ney 84]
 H. Ney. The use of a one-stage dynamic programming algorithm for connected word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32:263–271, 1984.
- [Ney 97a]
 H. Ney. Corpus-based statistical methods in speech and language processing. In G. Bloothoof and S. Young, editors, *Corpus-Based Methods in Speech and Language*, pages 1–26. Kluwer Academic Publishers, Dordrecht, 1997.
- [Ney 97b]
 H. Ney. Language models. In D. Gibbon, R. Moore, and R. Winski, editors, *Handbook for Spoken Language Systems*, pages 237–287. de Gruyter, Berlin, 1997.
- [Ney 99a]
 H. Ney. Statistical modelling – from speech recognition to text translation. In G. Chollet, M. D. Benedetto, A. Esposito, and M. Marinaro, editors, *Computational Models of Speech Pattern Processing*, pages 190–215. Springer, Berlin, Int. School on Neural Nets “Eduardo R. Caianiello”, Salerno, Italien, October 1999.
- [Ney 99b]
 H. Ney. Stochastic modelling: From pattern recognition to speech recognition and language translation. In W. F. et al., editor, *Mustererkennung*, pages 325–337. Springer, Berlin, DAGM-Jahrestagung (Deutsche Arbeitsgemeinschaft für Mustererkennung, Bonn, September 1999).
- [Ney 99c]
 H. Ney. The use of the maximum likelihood criterion in language modelling. In K. Ponting, editor, *Computational Models of Speech Pattern Processing*, pages 259–279. Springer, Berlin, NATO Advanced Studies Institute, St. Helier, Jersey, 1997, July 1999.
- [Ney 00]
 H. Ney. Stochastic modelling: From pattern classification to speech recognition and translation. In *Int. Conf. on Pattern Recognition, Barcelona*, volume 3, pages 25–32, Barcelona, September 2000.

- [Ney 01a]
H. Ney. The statistical approach to spoken language translation. In *Proc. IEEE Automatic Speech Recognition and Understanding Workshop*, page 8 pages, Madonna di Campiglio, Italy, December 2001. CD ROM, IEEE Catalog No. 01EX544.
- [Ney 01b]
H. Ney. Stochastic modelling: From pattern classification to language translation. In *Proc. Data-Driven Machine Translation Workshop, 39th Annual Meeting of the Ass. for Computational Linguistics (ACL)*, pages 33–37, Toulouse, July 2001.
- [Ney 03]
H. Ney. Maschinelle Sprachverarbeitung: Der statistische Ansatz in der Spracherkennung und Sprachübersetzung. *Informatik Spektrum*, 26(2):94–102, May 2003.
- [Ney et al. 92]
H. Ney, D. Mergel, A. Noll, and A. Paesler. Data-driven search organisation for continuous speech recognition. *IEEE Transactions on Speech Processing*, 40:272–281, 1992.
- [Ney et al. 94]
H. Ney, U. Essen, and R. Kneser. On structuring probabilistic dependences in stochastic language modelling. *Computer Speech and Language*, 8:1–38, 1994.
- [Ney et al. 97]
H. Ney, S. Martin, and F. Wessel. Statistical language modelling by leaving-one-out. In G. Bloothoof and S. Young, editors, *Corpus-Based Methods in Speech and Language*, pages 174–207. Kluwer Academic Publishers, Dordrecht, 1997.
- [Ney et al. 00]
H. Ney, S. Nießen, F. J. Och, C. Tillmann, H. Sawaf, and S. Vogel. Algorithms for Statistical Translation of Spoken Language. *IEEE Trans. on Speech and Audio Processing, Special Issue on Language Modeling and Dialogue Systems*, 8:24–36, January 2000.
- [Ng & Cardie 01]
V. Ng and C. Cardie. Improving Machine Learning Approaches to Coreference Resolution. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pages 104–111, Philadelphia, Pennsylvania, 2001.
- [Ng & Lee 96]
H. T. Ng and H. B. Lee. Integrating multiple knowledge sources to disambiguate word sense: An exemplar-based approach. In *Proceedings of ACL96*, 1996.
- [Ng & Zelle 97]
H. Ng and J. Zelle. Corpus-based approaches to semantic interpretation in NLP. *AI Magazine*, 18(4):45–64, 1997.
- [Ngai & Yarowsky 00]
G. Ngai and D. Yarowsky. Rule Writing or Annotation: Cost-efficient Resource Usage for Base Noun Phrase Chunking. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics*, pages 117–125, Hongkong, 2000.
- [Ngai et al. 02]
G. Ngai, M. Carpuat, and P. Fung. Identifying concepts across languages: A first step towards a corpus-based approach to automatic ontology alignment. In *Proc. of COLING 2002*, Taipei, Taiwan, 2002.
- [Nguyen & Cao 08]
H. Nguyen and T. Cao. Named entity disambiguation: A hybrid statistical and rule-based incremental approach. In J. Domingue and C. Anutariya, editors, *The Semantic Web*, volume 5367 of *Lecture Notes in Computer Science*, pages 420–433. Springer Berlin / Heidelberg, 2008.

- [Nguyen *et al.* 12a]
D. T. Nguyen, N. P. Nguyen, and M. T. Thai. Sources of misinformation in Online Social Networks: Who to suspect? In *MILITARY COMMUNICATIONS CONFERENCE, 2012 - MILCOM 2012*, pages 1–6. IEEE, October 2012.
- [Nguyen *et al.* 12b]
N. P. Nguyen, G. Yan, M. T. Thai, and S. Eidenbenz. Containment of misinformation spread in online social networks. In *Proceedings of the 3rd Annual ACM Web Science Conference, WebSci '12*, pages 213–222, New York, NY, USA, 2012. ACM.
- [NHS Information Authority 99a]
NHS Information Authority. Coding and classification: Clinical terms (the READ codes). <http://www.coding.nhsia.nhs.uk>, 1999.
- [NHS Information Authority 99b]
NHS Information Authority. Global standard for healthcare terminology. Unpublished report, 1999. Ref: 1999-IA-80.
- [Nicholson & Macgregor 02]
D. Nicholson and G. Macgregor. Learning Lessons Holistically in the Glasgow Digital Library. *D-Lib Magazine*, 8(7/8), 2002.
- [Nielsen 94]
J. Nielsen. Enhancing the Explanatory Power of Usability Heuristics. In *CHI '94: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 152–158, New York, NY, USA, 1994. ACM.
- [Nielsen 00]
J. Nielsen. *Designing Web Usability: The Practice of Simplicity*. New Riders Publishing, 2000.
- [Niesler & Woodland 96]
T. Niesler and P. Woodland. A variable-length category-based n-gram language model. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 164–167, Atlanta GA, 1996.
- [Nikolov *et al.* 95]
N. Nikolov, C. Mellish, and G. Ritchie. Sentence Generation from Conceptual Graphs. In *Proceedings of 3rd Int. Conf. on Conceptual Structures (ICCS'95)*, number 954 in LNAI, Santa Cruz, CA, 1995. Springer-Verlag.
- [Niles & Silverman 90]
L. T. Niles and H. F. Silverman. Combining hidden Markov models and neural network classifiers. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 417–420, Albuquerque, 1990.
- [Niles 91]
L. T. Niles. Timit phoneme recognition using an HMM-derived recurrent neural network. In *Proceedings European Conference on Speech Communication and Technology*, pages 559–562, Genova, Italy, 1991.
- [Niles *et al.* 90]
L. T. Niles, H. F. Silverman, and M. A. Bush. Neural networks, maximum mutual information training and maximum likelihood training. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 493–496, Albuquerque, 1990.
- [Nilsson 71a]
N. J. Nilsson. *Problem Solving Methods of Artificial Intelligence*. McGraw-Hill, New York, 1971.

- [Nilsson 71b]
N. Nilsson. *Problem-solving Methods in Artificial Intelligence*. McGraw-Hill, New York, 1971.
- [Nilsson 80]
N. Nilsson. *Principles of Artificial Intelligence*. Tioga Publishing Company, Palo Alto, CA, 1980.
- [Nirenburg 94]
S. Nirenburg. The workstation substrate of the pangloss project. In *Proceedings of the Conference on the Future Generation of Natural Language Processing systems (FGNLP-2)*, 1994.
- [NIST 02a]
NIST. Nist evaluation. 2002. <http://www.nist.gov/speech/tests/mt/doc/ngram-study.pdf>.
- [NIST 02b]
NIST. Nist scoring. 2002. <http://www.nist.gov/speech/tests/mt/mt2001/resource/>.
- [NLM]
NLM. Unified Medical Language System (UMLS). Technical report, National Library of Medicine, <http://www.nlm.nih.gov/research/umls/umlsmain.html>.
- [Norcliffe & Slater 91]
A. Norcliffe and G. Slater. *Mathematics of Software Construction*. Ellis Horwood, Chichester, 1991.
- [Norman 98]
F. Norman. Organizing medical networked information (omni). *Medical Informatics*, 23(1):43–51, 1998.
- [Novak *et al.* 04]
B. Novak, M. Grobelnik, and D. Mladenić. Dealing With Unlabelled Data. Technical report, SEKT project deliverable D1.2.1, 2004. <http://www.sekt-project.org/rd/deliverables/wp01/sekt-d-1-2-1-Dealing%20with%20Unlabelled%20Data.pdf>.
- [Novotney & Callison-Burch 10a]
S. Novotney and C. Callison-Burch. Cheap, fast and good enough: Automatic speech recognition with non-expert transcription. In *Proceedings of HLT-NAACL*, pages 207–215, 2010.
- [Novotney & Callison-Burch 10b]
S. Novotney and C. Callison-Burch. Shared task: crowdsourced accessibility elicitation of wikipedia articles. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, pages 41–44, 2010.
- [Noy & Klein 03]
N. F. Noy and M. Klein. Ontology evolution: not the same as schema evolution. *Knowledge and Information Systems*, 5, 2003.
- [Noy & McGuinness 01]
N. F. Noy and D. L. McGuinness. Ontology development 101: A guide to creating your first ontology. Technical Report KSL-01-05, Stanford Knowledge Systems Laboratory, March 2001.
- [Noy *et al.* 01]
N. Noy, M. Sintek, S. Decker, M. Crubézy, R. Ferguson, and M. Musen. Creating Semantic Web Contents with Protégé-2000. *IEEE Intelligent Systems*, 16(2):60–71, 2001.
- [NTALIANIS *et al.* 13]
K. NTALIANIS, N. MASTORAKIS, A. DOULAMIS, and P. TOMARAS. Social media video collection summarization based on social graph user interactions. In Z. Chen and E. Lopez-Neri, editors, *Recent Advances in Knowledge Engineering and Systems Science*, pages 162–167. World Scientific and Engineering Academy and Society, 2013.

- [Nyberg *et al.* 98]
 E. Nyberg, T. Mitamura, and C. Kamprath. The KANT Translation System: From R&D to Large-Scale Deployment. In *LISA Newsletter, Vol 2:1, March*, 1998.
- [OAI 01]
 OAI. The Open Archives Initiative Protocol for Metadata Harvesting. Technical report, Open Archives Initiative, <http://www.openarchives.org>, 2001.
- [Oakes & Paice 98]
 M. P. Oakes and C. D. Paice. Term extraction for automatic abstracting. In *Proc. of 1st Workshop on Computational Terminology, Computerm '98*, pages 91–95, Montreal, Canada, 1998. COLING/ACL.
- [Oakes & Paice 99]
 M. P. Oakes and C. D. Paice. The Automatic Generation of Templates for Automatic Abstracting. In *21st BCS IRSG Colloquium on IR*, Glasgow, 1999.
- [Oakes & Paice 01a]
 M. P. Oakes and C. D. Paice. Term extraction for automatic abstracting. In D. Bourigault, C. Jacquemin, and M.-C. L'Homme, editors, *Recent Advances in Computational Terminology*, volume 2 of *Natural Language Processing*, chapter 17, pages 353–370. John Benjamins Publishing Company, 2001.
- [Oakes & Paice 01b]
 M. Oakes and C. Paice. Term extraction for automatic abstracting. In *Recent Advances in Computational Terminology*, pages 353–369. John Benjamins, 2001.
- [Oakes & Taylor 94]
 M. Oakes and M. Taylor. Morphological analysis in vocabulary selection for the ringdoc pharmacological database. In P. Barahona, M. Veloso, and J. Braynt, editors, *Proc. of 12th International Congress of the European Federation for Medical Informatics (MIE '94)*, pages 523–528, Lisbon, Portugal, 1994.
- [Oard 03]
 D. W. Oard. The surprise language exercises. *ACM Transactions on Asian Language Information Processing (TALIP)*, 2(2):79–84, 2003.
- [Oberlander *et al.* 98]
 J. Oberlander, M. O'Donnell, A. Knott, and C. Mellish. Conversation in the museum: Experiments in dynamic hypermedia with the intelligent labelling explorer. *New Review of Hypermedia and Multimedia*, 4:11 — 32, 1998.
- [Och & Ney 00a]
 F. J. Och and H. Ney. A Comparison of Alignment Models for Statistical Machine Translation. In *Int. Conf. on Computational Linguistics*, volume 2, pages 1086–1090, July 2000.
- [Och & Ney 00b]
 F. J. Och and H. Ney. Improved statistical alignment models. In *Proceedings of the Association for Computational Linguistics*, pages 440–447, Hongkong, China, October 2000.
- [Och & Ney 02a]
 F. Och and H. Ney. Discriminative Training and Maximum Entropy Models for Statistical Machine Translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 295–302, Philadelphia, 2002. <http://citeseer.nj.nec.com/541909.html>.
- [Och & Ney 02b]
 F. J. Och and H. Ney. A systematic comparison of various alignment models. *Computational Linguistics*, page in press, 2002.

- [Och & Weber 98]
 F. J. Och and H. Weber. Improving Statistical Natural Language Translation with Categories and Rules. In *Int. Conf. on Computational Linguistics*, volume 2, pages 985–989, August 1998.
- [Och 99]
 F. J. Och. An Efficient Method for Determining Bilingual Word Classes. In *Proceedings of the European Chapter of the Association for Computational Linguistics*, pages 71–76, June 1999.
- [Och 02]
 F. J. Och. Yasmet. 2002. <http://www-i6.Informatik.RWTH-Aachen.de/web/Software/index.html>.
- [Och et al. 99]
 F. J. Och, C. Tillmann, and H. Ney. Improved Alignment Models for Statistical Machine Translation. In *EMNLP'99*, pages 20–28, June 1999.
- [O'Connor et al. 10a]
 B. O'Connor, R. Balasubramanyan, B. Routledge, and N. A. Smith. From tweets to polls: Linking text sentiment to public opinion time series. In *Proceedings of the Fourth International Conference on Weblogs and Social Media*, 2010.
- [O'Connor et al. 10b]
 B. O'Connor, M. Krieger, and D. Ahn. Tweetmotif: Exploratory search and topic summarization for twitter. In *Proceedings of the Fourth International Conference on Weblogs and Social Media*, 2010.
- [O'Connor et al. 10c]
 B. O'Connor, M. Krieger, and D. Ahn. TweetMotif: Exploratory Search and Topic Summarization for Twitter. In *Proceedings of the Fourth AAAI Conference on Weblogs and Social Media (ICWSM)*, pages 384–385, 2010.
- [Ó. Corcho and A. Gómez-Pérez and R. González-Cabero and MC Suárez-Figueroa 04]
 Ó. Corcho and A. Gómez-Pérez and R. González-Cabero and MC Suárez-Figueroa. ODEval: a Tool for evaluating RDF(S), DAML+OIL, and OWL Concept Taxonomies. In *First IFIP Conference on Artificial Intelligence Applications and Innovations (AIAI2004)*, 2004.
- [Odell 95]
 J. J. Odell. *The Use of Context in Large Vocabulary Speech Recognition*. Unpublished PhD thesis, University of Cambridge, 1995.
- [Odell et al. 94]
 J. J. Odell, P. C. Woodland, and S. J. Young. A one pass decoder design for large vocabulary recognition. Princeton NJ, March 1994. ARPA.
- [O'Donnell et al. 00]
 M. O'Donnell, A. Knott, J. Oberlander, and C. Mellish. Optimising text quality in generation from relational databases. In *Proceedings of the International Natural Language Generation Conference (INLG'00)*, pages 133–140, 2000.
- [Oflazer 96]
 K. Oflazer. Error-tolerant tree matching. In *Proc. of 16th International Conference of Computational Linguistics*, volume 2, pages 860–864, Copenhagen, Denmark, 1996.
- [Ogden & Bernick 97]
 W. Ogden and P. Bernick. Using Natural Language Interfaces. In M. Helander, editor, *Handbook of Human-Computer Interaction*. Elsevier Science Publishers B.V. (North-Holland), 1997.
- [Ogden & Richards 23]
 C. K. Ogden and I. A. Richards. *The Meaning of Meaning*. Harcourt Brace Jovanovich, Orlando, Florida, 1923.

- [Ogden 99]
B. Ogden. TIPSTER annotation and the Corpus Encoding Standard. <http://crl.nmsu.edu/-Research/Projects/tipster/annotation>, 1999.
- [Ogleby 99]
C. L. Ogleby. From Rubble To Virtual Reality: Photogrammetry And The Virtual World Of Ancient Ayutthaya, Thailan. In *Technical Meeting of the Photogrammetric Society*, Plzen, Chzeck Republic, January 1999.
- [Ognyanov & Kiryakov 02]
D. Ognyanov and A. Kiryakov. Tracking Changes in RDF(S) Repositories. In *13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02)*, pages 373–378, Siguenza, Spain, 2002.
- [Ogren 06]
P. Ogren. Knowtator: A Protege Plug-In For Annotated Corpus Construction . In *HLT-NAACL - Demos*, 2006.
- [Ohtake *et al.* 13]
K. Ohtake, J. Goto, S. De Saeger, K. Torisawa, J. Mizuno, and K. Inui. Nict disaster information analysis system. In *The Companion Volume of the Proceedings of IJCNLP 2013: System Demonstrations*, pages 29–32. Asian Federation of Natural Language Processing, 2013.
- [Olson & Lee 97]
M. Olson and B. Lee. Object Databases for SGML Document Management. In *IEEE International Conference on Systems Sciences*, 1997.
- [Olsson 97]
F. Olsson. Tagging and morphological processing in the svensk system. Unpublished M.Sc. thesis, University of Uppsala, 1997. <http://http://stp.ling.uu.se/fredriko/exjobb.ps>.
- [Olsson *et al.* 98]
F. Olsson, B. Gambäck, and M. Eriksson. Reusing Swedish Language Processing Resources in SVENSK. In *Workshop on Minimising the Efforts for LR Acquisition*, Granada, Spain, 1998.
- [Olsson *et al.* 02]
F. Olsson, G. Eriksson, K. Franzén, L. Asker, and P. Lidén. Notions of Correctness when Evaluating Protein Name Taggers. In *Proceedings of COLING 2002*, Taipei, Taiwan, 2002.
- [Onyshkevych & Nirenburg 94]
B. Onyshkevych and N. Nirenburg. The lexicon in the sceme of kbmt things. Memoranda in Computer and Cognitive Science MCCS-94-277, CRL/NMSU, 1994.
- [Onyshkevych *et al.* 93]
B. Onyshkevych, M. Okurowski, and L. Carlson. Tasks, Domains, and Languages for Information Extraction. In *Proceedings of Tipster Text Program (Phase I)*. Morgan Kaufmann, California, 1993.
- [Onyshkevych *et al.* 96]
B. Onyshkevych, Boyan, and S. Nirenburg. Microkosmos. *Machine Translation 10:1-2 (Special Issue on building lexicons for MT)*, 1996.
- [Oppermann & Specht 99]
R. Oppermann and M. Specht. A nomadic information system for adaptive exhibition guidance. In *ICHIM*, pages 103–109, 1999.
- [Opsahl 13]
K. Opsahl. Everything We Know About NSA Spying: “Through a PRISM, Darkly”. In *Proceedings of the Chaos Communication Congress*, 2013.

- [OPT00]
Opta soccer web page. <http://www.optasoccer.com>, October 2000.
- [Oren 05]
E. Oren. SemperWiki: a semantic personal Wiki. In *Semantic Desktop (ISWC)*, November 2005.
- [Oren *et al.* 06a]
E. Oren, R. Delbru, K. Moeller, M. Voelkel, and S. Handschuh. Annotation and Navigation in Semantic Wikis. In *First Workshop on Semantic Wikis: From Wiki to Semantic (SemWiki2006)*, 2006.
- [Oren *et al.* 06b]
E. Oren, J. G. Breslin, and S. Decker. How semantics make better wikis. In *WWW '06: Proceedings of the 15th international conference on World Wide Web*, pages 1071–1072, New York, NY, USA, 2006. ACM Press.
- [Oren *et al.* 08]
E. Oren, R. Delbru, M. Catasta, R. Cyganiak, H. Stenzhorn, and G. Tummarello. Sindice.com: A document-oriented lookup index for open linked data. *International Journal of Metadata, Semantics and Ontologies*, 3(1), 2008.
- [Orwell 91]
G. Orwell. *1984*. Univers Publishing House, Bucharest, 1991.
- [Osborne & Dredze 14]
M. Osborne and M. Dredze. Facebook, Twitter and Google Plus for Breaking News: Is there a winner? In *Proceedings of the International Conference on Weblogs and Social Media*, 2014.
- [Osenova & Simov 04]
P. Osenova and K. Simov. BulTreeBank stylebook. Technical Report BTB-TR05, BulTreeBank Project, May 2004.
- [O’Shaughnessy 87]
D. O’Shaughnessy. *Speech Communication, Human and Machine*. Addison-Wesley, Reading, MA, 1987.
- [O’Sullivan *et al.* 04]
D. O’Sullivan, B. Smyth, D. Wilson, K. McDonald, and A. Smeaton. Interactive Television Personalization. In L. Ardissono, A. Kobsa, and M. Maybury, editors, *[Ardissono et al. 04]*, chapter 4, pages 73–91. Kluwer, 2004.
- [Ounis *et al.* 06]
I. Ounis, M. de Rijke, C. Macdonald, G. Mishne, and I. Soboroff. Overview of the TREC-2006 blog track. In *Proceedings of the Fifteenth Text REtrieval Conference (TREC 2006)*. NIST, 2006.
- [Ounis *et al.* 11]
I. Ounis, C. Macdonald, J. Lin, and I. Soboroff. Overview of the trec-2011 microblog track. In *Proceedings of the 20th Text REtrieval Conference (TREC 2011)*, 2011.
- [Ousterhout 94]
J. Ousterhout. *Tcl and the Tk Toolkit*. Addison-Wesley, Reading, MA, 1994.
- [Owoputi *et al.* 13]
O. Owoputi, B. O’Connor, C. Dyer, K. Gimpel, N. Schneider, and N. A. Smith. Improved part-of-speech tagging for online conversational text with word clusters. In *HLT-NAACL*, pages 380–390, 2013.

- [Pado 07]
S. Pado. *Cross-Lingual Annotation Projection Models for Role-Semantic Information*. Unpublished PhD thesis, German Research Center for Artificial Intelligence and Saarland University, 2007.
- [Paek *et al.* 10]
T. Paek, M. Gamon, S. Counts, D. M. Chickering, and A. Dhesi. Predicting the importance of newsfeed posts and social network friends. In M. Fox and D. Poole, editors, *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010, Atlanta, Georgia, USA, July 11-15, 2010*. AAAI Press, 2010.
- [Paggio 98]
P. Paggio. Validating the TEMAA LE evaluation methodology: a case study on Danish spelling checkers. *Journal of Natural Language Engineering*, 4(3):211–228, 1998.
- [Paice & Jones 93a]
C. D. Paice and P. A. Jones. The identification of important concepts in highly structured technical papers. In *Proc. of ACM-SIGIR '93*, pages 69–77, Pittsburgh, PA, 1993.
- [Paice & Jones 93b]
C. D. Paice and P. A. Jones. The Identification of Important Concepts in Highly Structured Technical Papers. In R. Korfhage, E. Rasmussen, and P. Willett, editors, *Proc. of the 16th ACM-SIGIR Conference*, pages 69–78, 1993.
- [Paice & Oakes 99]
C. D. Paice and M. P. Oakes. A Concept-Based Method for Automatic Abstracting. Technical Report Research Report 27, Library and Information Commission, 1999.
- [Paice *et al.* 94]
C. D. Paice, W. J. Black, F. C. Johnson, and A. Neal. Automatic Abstracting. Technical Report R&D Report 6166, British Library, 1994.
- [Paiva 98]
D. Paiva. A Survey of Applied Natural Language Generation Systems. Technical Report ITRI-98-03, Information Technology Research Institute, Brighton, 1998.
- [Pak & Paroubek 10a]
A. Pak and P. Paroubek. Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC)*, 2010.
- [Pak & Paroubek 10b]
A. Pak and P. Paroubek. Twitter Based System: Using Twitter for Disambiguating Sentiment Ambiguous Adjectives. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pages 436–439, 2010.
- [Pallett *et al.* 94]
D. S. Pallett, J. G. Fiscus, W. M. Fisher, J. S. Garofolo, B. Lund, and M. Przybocki. 1993 benchmark tests for the ARPA Spoken Language Program. In *Proceedings ARPA Workshop on Human Language Technology*, pages 15–40. ARPA, 1994.
- [Pallett *et al.* 95]
D. Pallett, J. G. Fiscus, W. M. Fisher, J. Garofolo, B. A. Lund, A. Martin, and M. A. Przybocki. 1994 benchmark tests for the arpa spoken language program. In *Proceedings of ARPA Workshop on Spoken Language Technology*, pages 5–36, 1995.
- [Palma *et al.* 08]
R. Palma, P. Haase, Y. Wang, and M. D’Aquin. Propagation models and strategies. Technical Report D1.3.1, NeOn Project Deliverable, 2008.

- [Palma *et al.* 09]
R. Palma, P. Haase, and Q. Ji. Change management to support collaborative workflows. Technical Report D1.3.2, NeOn Project Deliverable, 2009.
- [Palmer & Day 97]
D. Palmer and D. Day. A statistical profile of the named entity task. In *In Proceedings of the 5th ACL Conference for Applied Natural Language Processing (ANLP'97)*, pages 190–193, Washington, USA, 1997.
- [Palmer *et al.* 99]
D. Palmer, J. Burger, and M. Ostendorf. Information extraction from broadcast news speech data. In *Proceedings of the DARPA Broadcast News and Understanding Workshop*, 1999.
- [Palmer *et al.* 05]
M. Palmer, D. Gildea, and P. Kingsbury. The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–106, 2005.
- [Pan *et al.* 04]
H. Pan, L. Zuo, V. Choudhary, Z. Zhang, S. H. Leow, F. T. Chong, Y. Huang, V. W. S. Ong, B. Mohanty, S. L. Tan, S. P. T. Krishnan, and V. B. Bajic. Dragon tf association miner: a system for exploring transcription factor associations through text-mining. *Nucleic Acids Research*, 32(suppl 2):W230–W234, 2004.
- [Pan *et al.* 06]
F. Pan, R. Mulkar, and J. Hobbs. An annotated corpus of typical durations of events. In *Proceedings of LREC 2006*, Genoa, May 2006. ELRA.
- [Pan *et al.* 07]
J. Pan, L. Lancieri, D. Maynard, F. Gandon, R. Cuel, and A. Leger. Success stories and best practices. Technical Report D1.4.2v2, KnowledgeWeb Deliverable, 2007.
- [Pang & Lee 04]
B. Pang and L. Lee. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the Association for Computational Linguistics*, pages 271–278, 2004.
- [Pang & Lee 05]
B. Pang and L. Lee. Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pages 115–124, 2005.
- [Pang & Lee 08]
B. Pang and L. Lee. Opinion mining and sentiment analysis. *Information Retrieval*, 2(1), 2008.
- [Pang *et al.* 02]
B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up? Sentiment Classification Using Machine Learning Techniques. In *Proceedings of the 2002 Conference on EMNLP*, pages 79–86, 2002.
- [Pankaj & Ghemawat 90]
Pankaj and Ghemawat. The snowball effect. *International Journal of Industrial Organization*, 8(3):335 – 351, 1990.
- [Pantel & Pennacchioni 06]
P. Pantel and M. Pennacchioni. Espresso: Leveraging generic patterns for automatically harvesting semantic relations. In *Proceedings of Conference on Computational Linguistics / Association for Computational Linguistics (COLING/ACL-06)*, pages 113–120, Sydney, Australia, 2006.

- [Pantel & Pennacchiotti 08]
P. Pantel and M. Pennacchiotti. Automatically harvesting and ontologizing semantic relations. In *Proceeding of the 2008 conference on Ontology Learning and Population: Bridging the Gap between Text and Knowledge*, pages 171–195. IOS Press, 2008.
- [Pantel & Ravichandran 04]
P. Pantel and D. Ravichandran. Automatically labeling semantic classes. In *Proceedings of HLT/NAACL-04*, pages 321–328, Boston, MA, 2004.
- [Pantos & Bekiari 95]
P. Pantos and C. Bekiari. POLEMON: A Project to Computerize the Monument Records at the Greek Ministry of Culture. In *Proceedings of the Conference on Archaeological Heritage: Inventory and Documentations Standards in Europe*, Oxford, UK, September 1995.
- [Paolo *et al.* 02]
A. Paolo, S. Montemagni, and V. Pirrelli. Controlled bootstrapping of lexico-semantic classes as a bridge between paradigmatic and syntagmatic knowledge: Methodology and evaluation. In *LREC 2002*, 2002.
- [Pape *et al.* 01]
D. Pape, J. Anstey, B. Carter, J. Leigh, M. Roussou, and T. Portlock. Virtual Heritage at iGrid 2000. In *Proceedings of INET 2001*, Stockholm, Sweden, June 2001.
- [Papineni *et al.* 97]
K. A. Papineni, S. Roukos, and R. T. Ward. Feature-based language understanding. In *European Conference on Speech Communication and Technology (EUROSPEECH)*, pages 1435–1438, Rhodes, Greece, September 1997.
- [Papineni *et al.* 98]
K. A. Papineni, S. Roukos, and R. T. Ward. Maximum likelihood and discriminative training of direct translation models. In *Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 189–192, Seattle, Washington, USA, May 1998.
- [Parent & Eskenazi 10]
G. Parent and M. Eskenazi. Clustering dictionary definitions using amazon mechanical turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, pages 21–29, 2010.
- [Parent & Eskenazi 11]
G. Parent and M. Eskenazi. Speaking to the crowd: Looking at past achievements in using crowd-sourcing for speech and predicting future challenges. In *In Proceedings of INTERSPEECH 2011*, pages 3037–3040, 2011.
- [Paris & Linden 96]
C. Paris and K. V. Linden. DRAFTER: An Interactive Support Tool for Writing Multilingual Instructions. *IEEE Computer, Special Issue on Interactive NLP*, July 1996.
- [Paris 88]
C. L. Paris. Tailoring object descriptions to the user’s level of expertise. *Computational Linguistics*, 14 (3):64–78, September 1988. Special Issue on User Modelling.
- [Paris 91]
C. L. Paris. Generation and explanation: Building an explanation facility for the explainable expert systems framework. In C. L. Paris, W. R. Swartout, and W. C. Mann, editors, *Natural language generation in artificial intelligence and computational linguistics*. Kluwer Academic Publishers, July 1991. Presented at the Fourth International Workshop on Natural Language Generation. Santa Catalina Island, CA, July, 1988.

- [Park *et al.* 05]
A. Park, T. Hazen, and J. Glass. Automatic processing of audio lectures for information retrieval: Vocabulary selection and language modeling. In *Proc., ICASSP 05*, Philadelphia, 2005.
- [Park *et al.* 13]
J. Park, V. Barash, C. Fink, and M. Cha. Emoticon Style: Interpreting Differences in Emoticons Across Cultures. In *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, pages 466–475. AAAI Press, 2013.
- [Parrott 01]
W. G. Parrott. *Emotions in social psychology: Essential readings*. Psychology Press, 2001.
- [Paskaleva *et al.* 02]
E. Paskaleva, G. Angelova, M. Yankova, K. Bontcheva, H. Cunningham, and Y. Wilks. Slavonic named entities in gate. Technical Report CS-02-01, University of Sheffield, 2002.
- [Passant & Laublet 08]
A. Passant and P. Laublet. Meaning of a tag: A collaborative approach to bridge the gap between tagging and linked data. In *Proceedings of the Linked Data on the Web Workshop (LDOW), Beijing, China*, 2008.
- [Passant *et al.* 10]
A. Passant, J. G. Breslin, and S. Decker. Rethinking microblogging: open, distributed, semantic. In *Proceedings of the 10th International Conference on Web Engineering*, pages 263–277, 2010.
- [Pasonneau *et al.* 96]
B. Pasonneau, K. Kukich, J. Robin, V. Hatzivassiloglou, L. Lefkowitz, and H. Jin. Generating summaries of work flow diagrams. In *Proc. of the International Conference on NLP and Industrial Applications (NLP+IA'96)*, Moncton, Canada, 1996.
- [Pastra 00]
K. Pastra. Basic semantic element extraction: The rule-writing experience. Unpublished M.Sc. thesis, Department of Language Engineering, UMIST, 2000.
- [Pastra *et al.* 01]
K. Pastra, H. Saggion, and Y. Wilks. Socis: Scene of crime information system. Technical Report CS-19-01, Department of Computer Science, University of Sheffield, 2001.
- [Pastra *et al.* 02a]
K. Pastra, D. Maynard, H. Cunningham, O. Hamza, and Y. Wilks. How feasible is the reuse of grammars for named entity recognition? In *Proceedings of the 3rd Language Resources and Evaluation Conference*, 2002. <http://gate.ac.uk/sale/lrec2002/reusability.ps>.
- [Pastra *et al.* 02b]
K. Pastra, D. Maynard, H. Cunningham, O. Hamza, and Y. Wilks. How feasible is the reuse of grammars for Named Entity Recognition? In *Proceedings of 3rd Language Resources and Evaluation Conference*, 2002.
- [Patel *et al.* 03]
C. Patel, K. Supekar, and Y. Lee. OntoGenie: Extracting Ontology Instances from WWW. In *Human Language Technology for the Semantic Web and Web Services, ISWC'03*, Sanibel Island, Florida, 2003.
- [Patel-Schneider 85]
P. Patel-Schneider. A decidable first-order logic for knowledge representation. *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, 1:455–458, 1985.

- [Paterson & Wegman 78]
M. Paterson and M. Wegman. Linear unification. *Journal of Computer System Science*, 16(2):158–167, 1978.
- [Patrick & Cunningham 03]
J. Patrick and H. Cunningham, editors. *Workshop on the Software Engineering and Architecture of Language Technology Systems (SEALTS)*. HLT-NAACL, Association for Computational Linguistics, 2003.
- [Patrick & Whalen 92]
A. S. Patrick and T. E. Whalen. Field Testing a Natural Language Information System: Usage Characteristics and Users’ Comments. *Interacting with Computers*, 4(2):218–230, 1992.
- [Paul & Baker 92]
D. B. Paul and J. M. Baker. The design for the Wall Street Journal-based CSR corpus. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 899–902, Banff, 1992.
- [Paul 92]
D. B. Paul. An efficient A* stack decoder algorithm for continuous speech recognition with a stochastic language model. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 1, pages 25–28, San Francisco, 1992.
- [Paul et al. 91]
D. B. Paul, J. K. Baker, and J. M. Baker. On the interaction between true source, training and testing language models. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 569–572, Toronto, 1991.
- [Paul et al. 10]
M. J. Paul, C. Zhai, and R. Girju. Summarizing contrastive viewpoints in opinionated text. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 66–76, 2010.
- [Paul et al. 11]
S. A. Paul, L. Hong, and E. H. Chi. What is a question? crowdsourcing tweet categorization. In *CHI’2011 Workshop on Crowdsourcing and Human Computation*, 2011.
- [Paulson 91]
L. Paulson. *ML for the Working Programmer*. Cambridge University Press, Cambridge, 1991.
- [Pavlovic 08]
D. Pavlovic. On quantum statistics in data analysis. In *Proceedings of the Second Quantum Interaction Symposium (QI-2008)*. College Publications, 2008.
- [Pazienza 97]
M. Pazienza, editor. *Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology*. Springer-Verlag, Berlin Heidelberg, 1997.
- [Pazienza 03]
M. T. Pazienza, editor. *Information Extraction in the Web Era*. Springer-Verlag, 2003.
- [Pazienza et al. 05]
M. T. Pazienza, A. Stellato, L. Henriksen, P. Paggio, and F. Zanzotto. Ontology mapping to support multilingual ontology-based question answering. In *4th International Semantic Web Conference (ISWC-2005)*, 2005.
- [Peñas & Gonzalo 01]
V. F. Peñas, A. and J. Gonzalo. Corpus-based terminology extraction applied to information access. In *Proceedings of Corpus Linguistics 2001*, Lancaster, UK, 2001.

- [Pearson 98]
J. Pearson. *Terms in Context*. John Benjamins, Amsterdam/Philadelphia, 1998.
- [Pease & Murray 03]
A. Pease and W. Murray. An english to logic traslator for ontology-based knowledge representation languages. In *In Proceedings of the 2003 IEEE International Conference on Natural Language Processing and Knowledge Engineering*, pages 777–783, Beijing, China, 2003.
- [Pedersen & Bruce 97]
T. Pedersen and R. Bruce. Distinguishing word senses in untagged text. In *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*, Providence, RI, August 1997.
- [Pennacchiotti & Popescu 11]
M. Pennacchiotti and A. Popescu. A Machine Learning Approach to Twitter User Classification. In *Proceedings of ICWSM 2011*, pages 281–288, 2011.
- [Pennebaker *et al.* 07]
J. D. Pennebaker, C. K. Chung, M. Ireland, G. A., and R. J. Booth. The LIWC2007 Application. Technical report, 2007. <http://www.liwc.net/liwcdescription.php>.
- [Penrose 90]
R. Penrose. *The Emperor’s New Mind*. Vintage, London, 1990.
- [Pereira & Shieber 84]
F. Pereira and S. Shieber. The semantics of grammar formalisms seen as computer languages. *Proceedings of the Tenth International Conference on Computational Linguistics*, 1984.
- [Pereira & Shieber 87]
F. Pereira and S. Shieber. *Prolog and Natural-Language Analysis*. Number 10 in CLSI Lecture Notes. Stanford University, Stanford, CA, 1987.
- [Perna & Spector 04]
J. Perna and A. Spector. Introduction to the Special Issue on Unstructured Information Management. *IBM Systems Journal*, 43(3), 2004.
- [Peshkin & Pfeffer 03]
L. Peshkin and A. Pfeffer. Bayesian Information Extraction Network . In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI-03)*, 2003.
- [Petasis *et al.* 99]
G. Petasis, G. Paliouras, V. Karkaletsis, C. Spyropoulos, and I. Androutsopoulos. Resolving part-of-speech ambiguity in the greek language using learning techniques. In *Proc. of the ECCAI Advanced Course on Artificial Intelligence (ACAI)*, Chania, Greece, 1999.
- [Petasis *et al.* 02]
G. Petasis, V. Karkaletsis, G. Paliouras, I. Androutsopoulos, and C. Spyropoulos. Ellogon: A new text engineering platform. In *Proceedings of 3rd Language Resources and Evaluation Conference (LREC’2002)*, Gran Canaria, Spain, 2002.
- [Peters & Kilgarriff 00]
W. Peters and A. Kilgarriff. Discovering Semantic Regularity in Lexical Resources. *International Journal for Lexicography*, 13(4), 2000.
- [Peters & Wilks 01]
W. Peters and Y. Wilks. Distribution-oriented extension of wordnet’s ontological framework. In *RANLP 2001*, Tzigov Chark, Bulgaria, 2001.

- [Peters 02]
W. Peters. Self-enriching properties of wordnet: Relationships between word senses. In *LREC 2002 Third International Conference on Language Resources and Evaluation*, Gran Canaria, 2002.
- [Peters 09]
W. Peters. Text-based legal ontology enrichment. *Proceedings of LOAIT*, pages 55–66, 2009.
- [Peters *et al.* 98a]
W. Peters, H. Cunningham, C. McCauley, K. Bontcheva, and Y. Wilks. Uniform Language Resource Access and Distribution. In *Workshop on Distributing and Accessing Lexical Resources at Conference on Language Resources Evaluation*, Granada, Spain, 1998.
- [Peters *et al.* 98b]
W. Peters, P. Vossen, P. Diez-Orzas, and G. Adriaens. Cross-linguistic Alignment of Wordnets with an Inter-Lingual-Index. *Computers and the Humanities*, 32(2-3):221–251, 1998.
- [Peters *et al.* 02]
W. Peters, L. Guthrie, and Y. Wilks. Computational psycholinguistics: The case of regular polysemy. In K. Petrova, A. Eftimova, and R. Shopov, editors, *Litora Psycholinguistica*, pages 328–338. SEMA RSH, 2002.
- [Peters *et al.* 05]
W. Peters, N. Aswani, K. Bontcheva, and H. Cunningham. Quantitative Evaluation Tools and Corpora v1. Technical report, SEKT project deliverable D2.5.1, 2005.
- [Petrov *et al.* 11]
S. Petrov, D. Das, and R. McDonald. A universal part-of-speech tagset. *arXiv preprint arXiv:1104.2086*, 2011.
- [Petrović *et al.* 10]
S. Petrović, M. Osborne, and V. Lavrenko. Streaming first story detection with application to Twitter. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 181–189, 2010.
- [Petrovic *et al.* 13]
S. Petrovic, M. Osborne, and V. Lavrenko. I Wish I Didn't Say That! Analyzing and Predicting Deleted Messages in Twitter. *arXiv preprint arXiv:1305.3107*, 2013.
- [Pfitzner *et al.* 12]
R. Pfitzner, A. Garas, and F. Schweitzer. Emotional divergence influences information spreading in twitter. *ICWSM*, 12:2–5, 2012.
- [Phan *et al.* 06]
X.-H. Phan, L.-M. Nguyen, and S. Horiguchi. Personal name resolution crossover documents by a semantics-based approach. *IEICE Trans. Inf. & Syst.*, Feb 2006, 2006.
- [Phelan *et al.* 09]
O. Phelan, K. McCarthy, and B. Smyth. Using Twitter to recommend real-time topical news. In *Proceedings of the 2009 ACM Conference on Recommender Systems*, pages 385–388, 2009.
- [Phillips 96]
J. L. Phillips. *How to Think about Statistics*. W. H. Freeman and Company, New York, 1996.
- [Piao *et al.* 07]
S. S. Piao, S. Ananiadou, Y. Tsuruoka, Y. Sasaki, and J. McNaught. Mining opinion polarity relations of citations. In *Proceedings of the 7th International Workshop on Computational Semantics*, 2007.
- [Picht & Draskau 85]
H. Picht and J. Draskau. *Terminology: an introduction*. University of Surrey, Guildford, UK, 1985.

- [Pickering *et al.* 02]
M. Pickering, D. Heesch, R. O’Callaghan, S. Rüger, and D. Bull. Video retrieval using global features in keyframes. In *Proceedings of TREC 2002*, Gaithersburg, MD, 2002.
- [Pieraccini & Levin 91]
R. Pieraccini and E. Levin. Stochastic Representation of Semantic Structure for Speech Understanding. In *European Conference on Speech Communication and Technology (EUROSPEECH)*, volume 2, pages 383–386, Genova, Italy, September 1991.
- [Pietrosanti & Graziadio 97]
E. Pietrosanti and B. Graziadio. Artificial Intelligence and Legal Text Management: Tools and Techniques for Intelligent Document Processing and Retrieval. In *Natural Language Processing: Extracting Information for Business Needs*, pages 277–291, London, March 1997. Unicom Seminars Ltd.
- [Pillman *et al.* 11]
W. Pillman, S. Schade, and P. Smits. *Innovations in sharing environmental observations and information, Proceedings of the 25th EnviroInfo Conference*. Shaker-Verlag, 2011.
- [Pirolli 09]
P. Pirolli. Powers of 10: Modeling complex information-seeking systems at multiple scales. *IEEE Computer*, 42(3):33–40, 2009.
- [Piskorski & Neumann 00]
J. Piskorski and G. Neumann. An intelligent text extraction and navigation system. In *Proceedings of the 6th Conference on Recherche d’Information Assistée par Ordinateur, RIAO-2000*, 2000.
- [Piskorski *et al.* 02]
J. Piskorski, W. Drożdżyński, O. Scherf, and F. Xu. A Flexible XML-based Regular Compiler for Creation and Conversion of Linguistic Resources. In *Proceedings of the LREC 2002 Third International Conference on Language Resources and Evaluation*, pages 1124–1129, Las Palmas, Canary Islands - Spain, 2002.
- [Plank *et al.* 14]
B. Plank, D. Hovy, R. McDonald, and A. Søgaard. Adapting taggers to Twitter with not-so-distant supervision. In J. Tsujii and J. Hajic, editors, *Proceedings of COLING: Technical Papers*, pages 1783–1792, Dublin, Ireland, 2014. Dublin City University and Association for Computational Linguistics.
- [Plumbaum *et al.* 11]
T. Plumbaum, S. Wu, E. W. De Luca, and S. Albayrak. User Modeling for the Social Semantic Web. In *2nd Workshop on Semantic Personalized Information Management: Retrieval and Recommendation, in conjunction with ISWC 2011*, 2011.
- [Poesio *et al.* 12]
M. Poesio, N. Diewald, M. Stührenberg, J. Chamberlain, D. Jettka, D. Goecke, and U. Kruschwitz. Markup infrastructure for the anaphoric bank: Supporting web collaboration. In A. Mehler, K.-U. Kühnberger, H. Lobin, H. Lungen, A. Storrer, and A. Witt, editors, *Modeling, Learning, and Processing of Text Technological Data Structures*, volume 370 of *Studies in Computational Intelligence*, pages 175–195. Springer Berlin / Heidelberg, 2012.
- [Poesio *et al.* 13]
M. Poesio, U. Kruschwitz, J. Chamberlain, L. Robaldo, and L. Ducceschi. Phrase Detectives: Utilizing Collective Intelligence for Internet-Scale Language Resource Creation. *Transactions on Interactive Intelligent Systems*, 3, 2013.
- [Poirier 99]
H. Poirier. The XeLDA Framework (presentation at Baslow workshop on Distributing and Accessing

- Linguistic Resources, Sheffield, 1999). <http://www.dcs.shef.ac.uk/~hamish/dalr/baslow/xelda.pdf>, 1999.
- [Pol04]
Semantic Technology, March 2004.
- [Polajnar *et al.* 05]
 T. Polajnar, V. Tablan, and H. Cunningham. User-friendly ontology authoring using a controlled language. Technical Report CS Report No. CS-05-10, University of Sheffield, Sheffield, UK, 2005.
- [Pollard & Sag 87]
 C. Pollard and I. Sag. *Information-Based Syntax and Semantics*, volume 1: Fundamentals of *Lecture Notes No.13*. CSLI, Stanford, CA, 1987.
- [Pollard & Sag 94]
 C. Pollard and I. Sag. *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics. University of Chicago Press, Chicago, 1994.
- [Pollet *et al.* 11]
 T. V. Pollet, S. G. Roberts, and R. I. Dunbar. Use of social network sites and instant messaging does not lead to increased offline social network size, or to emotionally closer relationships with offline network members. *Cyberpsychology, Behaviour, and Social Networking*, 14(4):253–258, 2011.
- [Pons-Porrata *et al.* 03]
 A. Pons-Porrata, R. Berlanga-Llavori, and J. Ruiz-Shulcloper. Building a hierarchy of events and topics for newspaper digital libraries. In F. Sebastiani, editor, *Proceedings of 25th European Conference on IR Research (ECIR'03)*, volume 2633, pages 588–596, Pisa, Italy, 2003. Springer-Verlag.
- [Popescu & Etzioni 05]
 A. Popescu and O. Etzioni. Extracting product features and opinions from reviews. In *Proc. of the Conference on Empirical Methods for Natural Language Processing (EMNLP)*, pages 339–346, Vancouver, Canada, 2005.
- [Popescu *et al.* 03a]
 A. Popescu, O. Etzioni, and H. Kautz. Towards a theory of natural language interfaces to databases. In *Proceedings of the International Conference on Intelligent User Interfaces (IUI'03)*, 2003.
- [Popescu *et al.* 03b]
 A.-M. Popescu, O. Etzioni, and H. Kautz. Towards a Theory of Natural Language Interfaces to Databases. In *IUI '03: Proceedings of the 8th international conference on Intelligent user interfaces*, pages 149–157, New York, NY, USA, 2003. ACM.
- [Popescu *et al.* 04]
 A. Popescu, A. Yates, and O. Etzioni. Class Extraction from the World Wide Web. In *Proceedings of AAAI 2004 Workshop on Adaptive Text Extraction and Mining (ATEM'04)*, 2004. <http://www.cs.washington.edu/research/nowita11/papers/WS104PopescuA.pdf>.
- [Popov *et al.* 03a]
 B. Popov, A. Kiryakov, D. Ognyanoff, D. Manov, A. Kirilov, and M. Goranov. Towards Semantic Web Information Extraction. In *Human Language Technologies Workshop at the 2nd International Semantic Web Conference (ISWC2003)*, Florida, USA, 2003.
- [Popov *et al.* 03b]
 B. Popov, A. Kiryakov, A. Kirilov, D. Manov, D. Ognyanoff, and M. Goranov. KIM – Semantic Annotation Platform. In *2nd International Semantic Web Conference (ISWC2003)*, pages 484–499, Berlin, 2003. Springer.

- [Popov *et al.* 04a]
 B. Popov, A. Kiryakov, A. Kirilov, D. Manov, D. Ognyanoff, and M. Goranov. KIM – A Semantic Platform for Information Extraction and Retrieval. *Natural Language Engineering*, 10:375–392, 2004.
- [Popov *et al.* 04b]
 B. Popov, A. Kiryakov, D. Ognyanoff, D. Manov, and A. Kirilov. KIM – A Semantic Platform for Information Extraction and Retrieval. *Natural Language Engineering*, 10:375–392, 2004.
- [Poritz 82]
 A. B. Poritz. Linear predictive hidden Markov models and the speech signal. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1291–1294, Paris, 1982.
- [Poritz 88]
 A. B. Poritz. Hidden Markov models: a guided tour. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 7–13, New York, 1988.
- [Porter 80]
 M. Porter. An algorithm for suffix stripping. *Program*, 14(3):130–137, 1980.
- [Porter 01]
 M. Porter. Snowball: A language for stemming algorithms, 2001.
- [Porzel & Malaka 04]
 R. Porzel and R. Malaka. A task-based approach for ontology evaluation. In P. Buitelaar, S. Handschuh, and B. Magnini, editors, *Proceedings of ECAI 2004 Workshop on Ontology Learning and Population*, Valencia, Spain, August 2004.
- [Potamianos *et al.* 00]
 A. Potamianos, E. Ammicht, and H.-K. J. Kuo. Dialogue Management in the Bell Labs Communicator System. In *Int. Conf. on Spoken Language Processing (ICSLP)*, volume 2, pages 603–606, Beijing, China, October 2000.
- [Potter 01]
 S. Potter. A survey of knowledge acquisition from natural language. AKT project report Task 1.1.2, 2001. [http://www.aiai.ed.ac.uk/project/akt/work/stephenp/TMA of KAfromNL.pdf](http://www.aiai.ed.ac.uk/project/akt/work/stephenp/TMA%20of%20KAfromNL.pdf).
- [Poulos 96]
 P. Poulos. Deriving XI models from WordNet. Unpublished M.Sc. thesis, Department of Computer Science, University of Sheffield, 1996.
- [Pound *et al.* 10]
 J. Pound, P. Mika, and H. Zaragoza. Ad-hoc object retrieval in the web of data. In *Proceedings of the 19th International Conference on World Wide Web*, pages 771–780. ACM, 2010.
- [Power *et al.* 98]
 R. Power, D. Scott, and R. Evans. What You See Is What You Meant: direct knowledge editing with natural language feedback. In *Proceedings of the 13th Biennial European Conference on Artificial Intelligence (ECAI'98)*, Brighton, UK, August 1998.
- [Power *et al.* 13]
 R. Power, B. Robinson, and D. Ratcliffe. Finding Fires with Twitter. In *Proceedings of the Australasian Language Technology Association Workshop*, pages 80–89, 2013.
- [Powsner *et al.* 98]
 S. M. Powsner, J. C. Wyatt, and P. Wright. Opportunities for and challenges of computerisation. *Lancet*, 352(9140):1617–1622, Nov 1998.

- [PPS06]
Fourth Workshop on Principles and Practice of Semantic Web Reasoning, Budva, Montenegro, June 2006.
- [Preece *et al.* 94]
 J. Preece, Y. Rogers, H. Sharp, D. Benyon, S. Holland, and T. Carey. *Human-Computer Interaction*. Addison-Wesley, Wokingham, UK, 1994.
- [Preotiuc-Pietro *et al.* 12]
 D. Preotiuc-Pietro, S. Samangooei, T. Cohn, N. Gibbins, and M. Niranjana. Trendminer: An architecture for real time analysis of social media text. In *Proceedings of the workshop on Real-Time Analysis and Mining of Social Streams*, 2012.
- [Preoțiuc-Pietro *et al.* 15]
 D. Preoțiuc-Pietro, V. Lampos, and N. Aletras. An analysis of the user occupational class through Twitter content. In *Proceedings of ACL*. ACL, 2015.
- [Press *et al.* 95]
 W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C – the Art of Scientific Computing*. Cambridge University Press, U.K., 1995.
- [Pressman 94]
 R. Pressman. *Software Engineering, a Practitioner’s Approach (European Edition)*. McGraw Hill, New York, 1994.
- [Presutti *et al.* 08]
 V. Presutti, A. Gangemi, D. Stefano, G. Aguado, M. Suárez-Figueroa, E. Montiel-Ponsoda, and M. Poveda. A library of Ontology Design Patterns: reusable solutions for collaborative design of networked ontologies. Technical report, NeOn project deliverable D2.5.1, 2008.
- [Prieto-Diaz & Freeman 87]
 R. Prieto-Diaz and P. Freeman. Status Report: Software Reusability. *IEEE Software*, Vol. 4 No. 1, 1987.
- [Prieto-Diaz & Freeman 93]
 R. Prieto-Diaz and P. Freeman. Status Report: Software Reusability. *IEEE Software*, Vol. 10, No. 3, 1993.
- [Procter 78a]
 P. Procter. *Longman Dictionary of Contemporary English*. Longman Group, Essex, 1978.
- [Procter 78b]
 P. Procter. *Longman Dictionary of Contemporary English*. Longman Group, 1978.
- [Procter *et al.* 13a]
 R. Procter, J. Crump, S. Karstedt, A. Voss, and M. Cantijoch. Reading the riots: What were the police doing on twitter? *Policing and society*, 23(4):413–436, 2013.
- [Procter *et al.* 13b]
 R. Procter, J. Crump, S. Karstedt, A. Voss, and M. Cantijoch. Reading the riots: What were the police doing on twitter? *Policing and society*, 23(4):413–436, 2013.
- [Procter *et al.* 13c]
 R. Procter, F. Vis, and A. Voss. Reading the riots on twitter: methodological innovation for the analysis of big data. *International journal of social research methodology*, 16(3):197–214, 2013.
- [Procter 79]
 P. Procter, editor. *The Longman Dictionary of Contemporary English*. Longman, Harlow, 1979.

- [Prokofyev *et al.* 15]
R. Prokofyev, A. Tonon, M. Luggen, L. Vouilloz, D. E. Difallah, and P. Cudré-Mauroux. Sanaphor: Ontology-based coreference resolution. In *The Semantic Web-ISWC 2015*, pages 458–473. Springer, 2015.
- [Prud’hommeaux & Seaborne 08]
E. Prud’hommeaux and A. Seaborne. SPARQL Query Language for RDF. W3C recommendation – 15 january 2008, W3C, <http://www.w3.org/TR/rdf-sparql-query/>, 2008.
- [Przybocki *et al.* 99]
M. Przybocki, J. Fiscus, J. Garofolo, and D. Pallett. 1998 HUB-4 information extraction evaluation. In *Proceedings of the DARPA Broadcast News Workshop*, pages 13–18, Herndon, VA, February 1999.
- [Pulman 96]
S. Pulman. Controlled Language for Knowledge Representation. In *CLAW96: Proceedings of the First International Workshop on Controlled Language Applications*, pages 233–242, Leuven, Belgium, 1996.
- [Pustejovsky *et al.* 97]
J. Pustejovsky, B. Boguraev, M. Verhagen, and P. Buitelaar. Semantic tagging and typed hyper-linking. In *AAAI Symposium on NLP and the WWW*, Stanford, 1997.
- [Pustejovsky *et al.* 03]
J. Pustejovsky, P. Hanks, R. Sauri, A. See, R. Gaizauskas, A. Setzer, D. Radev, B. Sundheim, D. Day, L. Ferro, et al. The Timebank Corpus. In *Corpus Linguistics*, pages 647–656, 2003.
- [Pustejovsky *et al.* 04]
J. Pustejovsky, B. Ingria, R. Sauri, J. Castano, J. Littman, and R. Gaizauskas. The Specification Language TimeML. In *The Language of Time: A Reader*, pages 545–557. Oxford University Press, 2004.
- [Pustejovsky *et al.* 10]
J. Pustejovsky, K. Lee, H. Bunt, and L. Romary. ISO-TimeML: An international standard for semantic annotation. In *LREC*, 2010.
- [P.W.Lord *et al.* 03]
P.W.Lord, R. Stevens, A. Brass, and C.A.Goble. Investigating semantic similarity measures across the Gene Ontology: the relationship between sequence and annotation. *Bioinformatics*, 19(10):1275–83, 2003.
- [Qazvinian *et al.* 11]
V. Qazvinian, E. Rosengren, D. R. Radev, and Q. Mei. Rumor has it: Identifying misinformation in microblogs. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP ’11*, pages 1589–1599, 2011.
- [Qu *et al.* 06]
Y. Qu, W. Hu, and G. Cheng. Constructing virtual documents for ontology matching. In *Proceedings of WWW2006*, pages 23–31, 2006.
- [Qu *et al.* 08]
L. Qu, C. Müller, and I. Gurevych. Using tag semantic network for keyphrase extraction in blogs. In *Proceedings of the 17th Conference on Information and Knowledge Management*, pages 1381–1382, 2008.
- [Quinlan 86]
J. R. Quinlan. Induction of decision trees. *Mach. Learn.*, 1(1):81–106, March 1986.

- [Quinlan 93]
J. R. Quinlan. *C4.5 Programs for Machine Learning*. Morgan Kaufmann, California, 1993.
- [Quinn & Bederson 11]
A. J. Quinn and B. B. Bederson. Human computation: a survey and taxonomy of a growing field. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems, CHI '11*, pages 1403–1412, 2011.
- [R. Florian 04]
A. I. H. J. N. K. X. L. N. N. S. R. R. Florian, H. Hassan. A statistical model for multilingual entity detection and tracking. In *Proc. of HLT/NAACL 2004*, Boston, MA, 2004.
- [Rabiner 89]
L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77:257–286, 1989.
- [Rabiner *et al.* 89]
L. R. Rabiner, J. G. Wilpon, and F. K. Soong. High performance connected digit recognition using hidden Markov models. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37:1214–1225, 1989.
- [Radev & McKeown 98]
D. R. Radev and K. R. McKeown. Generating natural language summaries from multiple on-line sources. *Computational Linguistics*, 24(3):469–500, September 1998.
- [Radev *et al.* 00]
D. R. Radev, H. Jing, and M. Budzikowska. Centroid-based summarization of multiple documents: sentence extraction, utility-based evaluation, and user studies. In *ANLP/NAACL Workshop on Summarization*, Seattle, WA, April 2000.
- [Radev *et al.* 04a]
D. Radev, T. Allison, S. Blair-Goldensohn, J. Blitzer, A. Çelebi, S. Dimitrov, E. Drabek, A. Hakim, W. Lam, D. Liu, J. Otterbacher, H. Qi, H. Saggion, S. Teufel, M. Topper, A. Winkel, and Z. Zhang. MEAD — A platform for multidocument multilingual text summarization. In *Conference on Language Resources and Evaluation (LREC)*, Lisbon, Portugal, 2004.
- [Radev *et al.* 04b]
D. R. Radev, H. Jing, M. Styś, and D. Tam. Centroid-based summarization of multiple documents. *Information Processing and Management*, 40(6):919–938, November 2004.
- [Radford *et al.* 95]
I. Radford, V. M. Arranz, S. Ananiadou, and J. Tsujii. Dynamic context matching for knowledge acquisition from small corpora. In S. Bolasco, L. Lebart, and A. Salem, editors, *Analisi Statistica dei Dati Testuali*, volume 1, pages 321–239, Rome, Italy, 1995. CISU.
- [Radford *et al.* 97]
I. Radford, S. Ananiadou, M. V. Arranz, and J. Tsujii. Term disambiguation by adding structural constraints to lexically based context matching techniques. In *Proc. of Natural Language Processing Pacific Rim Symposium (NLPRS)*, Phuket, Thailand, 1997.
- [Rafelsberger & Scharl 09]
W. Rafelsberger and A. Scharl. Games with a purpose for social networking platforms. In *Proceedings of the 20th ACM conference on Hypertext and hypermedia, HT '09*, pages 193–198, 2009.
- [Rafols *et al.* 10]
I. Rafols, A. L. Porter, and L. Leydesdorff. Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for information Science and Technology*, 61(9):1871–1887, 2010.

- [Raghunathan *et al.* 10]
K. Raghunathan, H. Lee, S. Rangarajan, N. Chambers, M. Surdeanu, D. Jurafsky, and C. Manning. A multi-pass sieve for coreference resolution. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 492–501. Association for Computational Linguistics, 2010.
- [Rajman *et al.* 03]
M. Rajman, M. Vesely, and P. Andrews. Report of WG1 - state of the art, evaluation and recommendations regarding document processing and visualization techniques. Technical Report D3.1, IST project IST-2001-37574, Network of Excellence in Text Mining and its Applications in Statistics (NEMIS), 2003.
- [Ramage 99]
M. Ramage. *The Learning Way: Evaluating Cooperative Systems*. Unpublished PhD thesis, University of Lancaster, 1999.
- [Ramage *et al.* 09]
D. Ramage, D. Hall, R. Nallapati, and C. D. Manning. Labeled LDA: a supervised topic model for credit attribution in multi-labeled corpora. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 248–256, 2009.
- [Ramage *et al.* 10]
D. Ramage, S. Dumais, and D. Liebling. Characterizing microblogs with topic models. In *Proceedings of the Fourth International Conference on Weblogs and Social Media (ICWSM)*, 2010.
- [Ramakrishnan *et al.* 10]
C. Ramakrishnan, W. A. Baumgartner, J. A. Blake, G. A. P. C. Burns, K. Bretonnel Cohen, H. Drabkin, J. Eppig, E. Hovy, C. N. Hsu, L. E. Hunter, T. Ingulfsen, H. R. Onda, S. Pokkunuri, E. Riloff, C. Roeder, and K. Verspoor. Building the scientific knowledge mine (SciKnowMine): a community-driven framework for text mining tools in direct service to biocuration. In R. Witte, H. Cunningham, J. Patrick, E. Beisswanger, E. Buyko, U. Hahn, K. Verspoor, and A. R. Coden, editors, *New Challenges for NLP Frameworks (NLPFrameworks 2010)*, LREC 2010, pages 9–14, Valletta, Malta, May 2010. ELRA.
- [Ramanath *et al.* 13]
R. Ramanath, M. Choudhury, K. Bali, and R. S. Roy. Crowd prefers the middle path: A new iaa metric for crowdsourcing reveals turker biases in query segmentation. In *Proceedings of the annual conference of the Association for Computational Linguistics*, volume 1, pages 1713–1722, 2013.
- [Ramsay 86]
A. Ramsay. Knowing that and knowing what. Unpublished paper, cognitive studies programme, University of Sussex, 1986.
- [Ramsay 88]
A. Ramsay. *Formal Methods in Artificial Intelligence*. Cambridge University Press, Cambridge, 1988.
- [Ramsay 00]
A. Ramsay. Research in the Department of Language Engineering. <http://www.ccl.umist.ac.uk/research/booklet.html>, UMIST Department of Language Engineering, 2000.
- [Ramshaw & Marcus 95]
L. Ramshaw and M. Marcus. Text Chunking Using Transformation-Based Learning. In *Proceedings of the Third ACL Workshop on Very Large Corpora*, 1995.

- [Rao *et al.* 92]
 R. Rao, H. J. S.K. Card, J. Mackinlay, and G. Robertson. The Information Grid: a Framework for Information Retrieval and Retrieval-Centered Applications. In *Proceedings of the fifth annual ACM symposium on User interface software and technology (UIST '92)*, pages 23–32, Monterey, CA, 1992.
- [Rao *et al.* 11]
 D. Rao, P. McNamee, and M. Dredze. Entity linking: Finding extracted entities in a knowledge base. In *Multi-source, Multi-lingual Information Extraction and Summarization*. 2011.
- [Rao *et al.* 13]
 D. Rao, P. McNamee, and M. Dredze. Entity linking: Finding extracted entities in a knowledge base. In *Multi-source, Multi-lingual Inf. Extraction and Summarization*. Springer, 2013.
- [Raskin & Pan 05]
 R. G. Raskin and M. J. Pan. Knowledge representation in the semantic web for earth and environmental terminology (sweet). *Computers and Geosciences*, 31(9):1119–1125, 2005.
- [Rasmussen & Willett 87]
 E. Rasmussen and P. Willett. Non-hierarchical document clustering using the icl distribution array processor. In *SIGIR '87: Proceedings of the 10th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 132–139, New York, NY, USA, 1987. ACM Press.
- [Rastier *et al.* 94]
 F. Rastier, M. Cavazza, and A. Abeille. *Semantique pour l'analyse: de la linguistique a l'informatique*. Masson, Paris, 1994.
- [Rath & Pepper 99]
 H. H. Rath and S. Pepper. Topic maps: Introduction and allegro. In *Proceedings of Markuo Technologies'99*, Philadelphia, USA, 1999.
- [Ratinov & Roth 09]
 L. Ratinov and D. Roth. Design challenges and misconceptions in named entity recognition. In *Proceedings of the Thirteenth Conference on Computational Natural Language Learning*, pages 147–155. Association for Computational Linguistics, 2009.
- [Ratkiewicz *et al.* 11]
 J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, A. Flammini, and F. Menczer. Detecting and tracking political abuse in social media. In *5th International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2011.
- [Ratnaparkhi 96a]
 A. Ratnaparkhi. A maximum entropy part-of-speech tagger. In *Proc. of the Empirical Methods in Natural Language Processing Conference*, University of Pennsylvania, 1996.
- [Ratnaparkhi 96b]
 A. Ratnaparkhi. A maximum entropy model for part-of-speech tagging. In E. Brill and K. Church, editors, *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 133–142. Association for Computational Linguistics, University of Pennsylvania, Philadelphia, PA, May 1996.
- [Ratnaparkhi 96c]
 A. Ratnaparkhi. A maximum entropy model for part-of-speech tagging. In E. Brill and K. Church, editors, *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 133–142. Association for Computational Linguistics, University of Pennsylvania, Philadelphia, PA, May 1996.

- [Ratnaparkhi 98]
 A. Ratnaparkhi. *Maximum Entropy Models for Natural Language Ambiguity Resolution*. Unpublished PhD thesis, Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA, 1998.
- [Rauber & Aschenbrenner 01]
 A. Rauber and A. Aschenbrenner. Part of our culture is born digital — On efforts to preserve it for future generations. *TRANS. On-line Journal for Cultural Studies (Internet-Zeitschrift für Kulturwissenschaften)*, 10, July 2001.
- [Rauber & Muller-Kogler 01]
 A. Rauber and A. Muller-Kogler. Integrating automatic genre analysis into digital libraries. In *Proceedings of the First ACM-IEEE Joint Conf on Digital Libraries*, Roanoke, VA, 2001.
- [Raux *et al.* 11]
 S. Raux, N. Grünwald, and C. Prieur. Describing the web in less than 140 characters. In L. A. Adamic, R. A. Baeza-Yates, and S. Counts, editors, *Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, July 17-21, 2011*. The AAAI Press, 2011.
- [Ravikant & Rifkin 10]
 N. Ravikant and A. Rifkin. Why Twitter Is Massively Undervalued Compared To Facebook. *TechCrunch*, 2010. <http://techcrunch.com/2010/10/16/why-twitter-is-massively-undervalued-compared-to-facebook/>.
- [Ravin & Kazi 99]
 Y. Ravin and Z. Kazi. Is Hillary Rodham Clinton the president? disambiguating names across documents. In *Proceedings of the ACL 1999 Workshop on Coreference and its Applications*, Jun 1999.
- [Raymond 99]
 E. Raymond. The cathedral and the bazaar. *Knowledge, Technology & Policy*, 12(3):23–49, September 1999.
- [Rayner & Wyard 95]
 M. Rayner and P. Wyard. Robust parsing of n-best speech hypothesis lists using a general grammar based language model. In *Proceedings European Conference on Speech Communication and Technology*, pages 1793–1796, madrid, 1995.
- [Rayner *et al.* 96]
 M. Rayner, D. Carter, and P. Bouillon. Adapting the core language engine to french and spanish. *NLP+IA*, 1996.
- [Rayson & Wilson 96]
 P. Rayson and A. Wilson. The ACAMRIT semantic tagging system: progress report. In W. Evett and I. Rose, editors, *Proc. of AISB-96 Workshop on Language Engineering for Document Analysis and Recognition*, pages 13–20, Nottingham Trent University, England, 1996.
- [Reape & Mellish 99]
 M. Reape and C. Mellish. Just what *is* aggregation anyway? In *Proceedings of the European Workshop on Natural Language Generation (EWNLG'99)*, pages 20 – 29, Toulouse, France, May 1999.
- [Recasens *et al.* 13]
 M. Recasens, M.-C. de Marneffe, and C. Potts. The life and death of discourse entities: Identifying singleton mentions. In *HLT-NAACL*, pages 627–633, 2013.

- [Reckman *et al.* 13]
H. Reckman, C. Baird, J. Crawford, R. Crowell, L. Micciulla, S. Sethi, and F. Veress. teragram: Rule-based detection of sentiment phrases using sas sentiment analysis. In *Proceedings of SemEval 2013 International Workshop on Semantic Evaluation*, Atlanta, Georgia, 2013.
- [Recupero & Cambria 15]
D. R. Recupero and E. Cambria. Eswc’14 challenge on concept-level sentiment analysis. In *Semantic Web Evaluation Challenge*, pages 3–20. Springer, 2015.
- [Recupero *et al.* 15]
D. R. Recupero, M. Dragoni, and V. Presutti. Eswc 15 challenge on concept-level sentiment analysis. In *Semantic Web Evaluation Challenge*, pages 211–222. Springer, 2015.
- [Reeve & Han 05]
L. H. Reeve and H. Han. Survey of semantic annotation platforms. In Haddad *et al.* [Haddad *et al.* 05], pages 1634–1638.
- [Reichman 85]
R. Reichman. *Getting Computers to Talk Like You and Me*. MIT Press, Cambridge, Mass., 1985.
- [Reidsma *et al.* 03a]
D. Reidsma, J. Kuper, T. Declerck, H. Saggion, and H. Cunningham. Cross-Document Annotation for Multimedia Retrieval. In *Proceedings 3rd Workshop on NLP and XML (NLPXML-2003)*, pages 41–48, Budapest, 2003.
- [Reidsma *et al.* 03b]
D. Reidsma, J. Kuper, T. Declerck, H. Saggion, and H. Cunningham. Cross-Document Ontology-Based Information Extraction for Multimedia Retrieval. In *Proceedings of ICCS’03*, pages 41–48, Desden, 2003.
- [Reidsma *et al.* 05]
D. Reidsma, N. Jovanovic, and D. Hofs. Designing annotation tools based on properties of annotation problems. In *Measuring Behavior 2005, 5th International Conference on Methods and Techniques in Behavioral Research*, Wageningen, The Netherlands, 2005.
- [Reinert 86]
M. Reinert. Un logiciel d’analyse lexicale (Alceste). *Cahiers de l’Analyse des Donnees*, 4, 1986.
- [Reinhart 83]
T. Reinhart. *Anaphora and Semantic Interpretation*. Croom Helm, London, 1983.
- [Reiter & Dale 99]
E. Reiter and R. Dale. Building Natural Language Generation Systems. *Journal of Natural Language Engineering*, Vol. 3 Part 1, 1999.
- [Reiter & Dale 00a]
E. Reiter and R. Dale. *Building Natural Language Generation Systems*. Cambridge University Press, Cambridge, 2000.
- [Reiter & Dale 00b]
E. Reiter and R. Dale. *Building Natural Language Generation Systems*. Cambridge University Press, Cambridge, England, 2000.
- [Reiter 80]
R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13, 1980.
- [Reiter 84]
R. Reiter. Towards a logical reconstruction of relational database theory. In M. Brodie, J. Mylopoulos, and J. Schmidt, editors, *On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases, and Programming Languages*. Springer-Verlag, New York, 1984.

- [Reiter 94a]
E. Reiter. Has a Consensus NL Generation Architecture Appeared, and is it Psycholinguistically Plausible? In *Proceedings of the Seventh International Workshop on Natural Language Generation (INLGW-1994)*, 1994. <http://xxx.lanl.gov/abs/CS.c1/9411032>.
- [Reiter 94b]
E. Reiter. Has a Consensus NL Generation Architecture Appeared, and is it Psycholinguistically Plausible? In *Proceedings of 7th Int. Workshop on NL Generation (INLG-94)*, pages 163–170, Kennebunkport, Maine, USA, 1994.
- [Reiter 95]
E. Reiter. NLG vs. Templates. In *Proceedings of the 5th European Workshop on Natural Language Generation (EWNLG-95)*, pages 95 – 106, Leiden, The Netherlands, 1995.
- [Reiter 99]
E. Reiter. Are Reference Architectures Standardisation Tools or Descriptive Aids? In *Proceedings of the AISB'99 Workshop on Reference Architectures and Data Standards for NLP*, Edinburgh, April 1999. The Society for the Study of Artificial Intelligence and Simulation of Behaviour.
- [Reiter 00]
E. Reiter. Pipelines and size constraints. *Computational Linguistics*, 26:251–259, 2000.
- [Reiter *et al.* 95]
E. Reiter, C. Mellish, and J. Levine. Automatic generation of technical documentation. *Journal of Applied Artificial Intelligence*, 9(3):259–287, 1995.
- [Reiter *et al.* 03a]
E. Reiter, R. Robertson, and L. Osman. Lessons from a Failure: Generating Tailored Smoking Cessation Letters. *Artificial Intelligence*, 144:41–58, 2003.
- [Reiter *et al.* 03b]
E. Reiter, S. Sripada, and S. Williams. Acquiring and Using Limited User Models in NLG. In *Proceedings of the 2003 European Natural Language Generation Workshop (EWNLG'03)*, pages 87 – 94. Association for Computational Linguistics, 2003.
- [Reithinger 91]
N. Reithinger. POPEL—a Parallel and Incremental Natural Language Generation System. In C. Paris, W. Swartout, and W. Mann, editors, *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, Dordrecht, Netherlands, 1991. Kluwer.
- [Remus *et al.* 10]
R. Remus, U. Quasthoff, and G. Heyer. SentiWS - A Publicly Available German-language Resource for Sentiment Analysis. In N. C. C. Chair), K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, and D. Tapias, editors, *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, Valletta, Malta, may 2010. European Language Resources Association (ELRA).
- [Ren *et al.* 13]
Z. Ren, S. Liang, E. Meij, and M. de Rijke. Personalized time-aware tweets summarization. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '13*, pages 513–522, New York, NY, USA, 2013. ACM.
- [Renals & Dalby 89]
S. Renals and J. Dalby. Analysis of a neural network model for speech recognition. In *Proceedings European Conference on Speech Communication and Technology*, volume 1, pages 333–336, 1989.
- [Renals & Hochberg 94]
S. Renals and M. Hochberg. Using gamma filters to model temporal dependencies in speech. pages 1491–1494, Yokohama, Japan, 1994.

- [Renals & Hochberg 95]
S. Renals and M. Hochberg. Efficient search using posterior phone probability estimates. volume 1, pages 596–599, Detroit, 1995.
- [Renals & MacKay 93]
S. Renals and D. MacKay. Bayesian regularisation methods in a hybrid MLP–HMM system. In *Proceedings European Conference on Speech Communication and Technology*, pages 1719–1722, Berlin, 1993.
- [Renals & Morgan 92]
S. Renals and N. Morgan. Connectionist probability estimation in HMM speech recognition. Technical Report TR-92-081, International Computer Science Institute, Berkeley CA, USA, 1992.
- [Renals & Rohwer 89]
S. Renals and R. Rohwer. Learning phoneme recognition using neural networks. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 413–416, Glasgow, 1989.
- [Renals 87]
S. Renals. A connectionist approach to speech recognition. MSc thesis, University of Edinburgh, 1987.
- [Renals 90]
S. Renals. *Speech and Neural Network Dynamics*. Unpublished PhD thesis, University of Edinburgh, 1990.
- [Renals 96]
S. Renals. Phone deactivation pruning in large vocabulary continuous speech recognition. *IEEE Signal Processing Letters*, 3:4–6, 1996.
- [Renals *et al.* 91a]
S. Renals, D. McKelvie, and F. McInnes. A comparative study of continuous speech recognition using neural networks and hidden Markov models. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 369–372, Toronto, 1991.
- [Renals *et al.* 91b]
S. Renals, N. Morgan, and H. Bourlard. Probability estimation by feed-forward networks in continuous speech recognition. In *Proceedings IEEE Workshop on Neural Networks for Signal Processing*, pages 309–318, Princeton NJ, 1991.
- [Renals *et al.* 91c]
S. Renals, N. Morgan, and H. Bourlard. Probability estimation by feed-forward networks in continuous speech recognition. Technical Report TR-91-030, International Computer Science Institute, Berkeley CA, USA, 1991.
- [Renals *et al.* 92]
S. Renals, N. Morgan, H. Bourlard, H. Franco, and M. Cohen. Connectionist optimisation of tied mixture hidden Markov models. In J. E. Moody, S. J. Hanson, and R. P. Lippmann, editors, *Advances in Neural Information Processing Systems*, volume 4, pages 167–174. Morgan Kaufmann, California, San Mateo CA, 1992.
- [Renals *et al.* 94]
S. Renals, N. Morgan, H. Bourlard, M. Cohen, and H. Franco. Connectionist probability estimators in HMM speech recognition. *IEEE Transactions on Signal Processing*, 2:161–175, 1994.
- [Renals *et al.* 99]
S. Renals, Y. Gotoh, R. Gaizauskas, and M. Stevenson. Baseline IE-NE Experiments Using The SPRACH/LASIE System. In *Proceedings of the DARPA Broadcast News Workshop*, Herndon, VA, 1999.

- [Renals *et al.* 00]
S. Renals, D. Abberley, D. Kirby, and T. Robinson. Indexing and retrieval of broadcast news. *Speech Communication*, 32(1-2), 2000.
- [Renn 99]
J. Renn. Challenges of the Information Revolution for the Max Planck Society. Lecture given at the conference of the Max Planck Society “The Transformation of Science, Research between Printed Information and the Challenges of Electronic Networks”, June 1999. Schloß Elmau.
- [Resnick *et al.* 94]
P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. Grouplens: An open architecture for collaborative filtering of netnews. In *Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*, pages 175 – 186, Chapel Hill, North Carolina, United States, 1994. ACM.
- [Resnik 93]
P. Resnik. *Selection and Information: A Class-Based Approach to Lexical Relationships*. Unpublished PhD thesis, Institute for Research in Cognitive Science, University of Pennsylvania, 1993.
- [Resnik 95a]
P. Resnik. Disambiguating noun groupings with respect to WordNet senses. In *Proc. of 3rd Workshop on Very Large Corpora*. MIT, 1995.
- [Resnik 95b]
P. Resnik. Using information content to evaluate semantic similarity in a taxonomy. In *Proc. of 14th International Joint Conference on Artificial Intelligence*, pages 448–453, Montreal, Canada, 1995.
- [Resnik 95c]
P. Resnik. Using information content to evaluate semantic similarity in a taxonomy. In *Proc. of 14th International Joint Conference on Artificial Intelligence*, pages 448–453, Montreal, Canada, 1995.
- [Resnik 97]
P. Resnik. Selectional preference and sense disambiguation. In *Proc. of ACL-SIGLEX Workshop "Tagging Text with Lexical Semantics: Why, What and How?"*, Washington, D.C., 1997.
- [Resnik 99]
P. Resnik. Semantic Similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language. *Journal of Artificial Intelligence Research*, pages 95–13, 1999.
- [Rey 79]
A. Rey. *La terminologie: noms et notions*. Que sais-je? PUF, 1979.
- [Reyes *et al.* 13]
A. Reyes, P. Rosso, and T. Veale. A multidimensional approach for detecting irony in Twitter. *Language Resources and Evaluation*, pages 1–30, 2013.
- [Richardson *et al.* 95]
F. Richardson, M. Ostendorf, and J. R. Rohlicek. Lattice-based search strategies for large vocabulary speech recognition. pages 576–579, Detroit, 1995.
- [Rick 83]
E. Rick. Users are individuals: individualising user models. *International Journal of Man-Machine Studies*, 18:199–214, 1983.
- [Rifkin & Klautu 04]
R. Rifkin and A. Klautu. In Defense of One-Vs-All Classification. *Journal of Machine Learning Research*, 5:101–141, 2004.

- [Rigau *et al.* 97]
 G. Rigau, J. Atserias, and E. Agirre. Combining unsupervised lexical knowledge methods for word sense disambiguation. In *Proc. of ACL/EACL*, pages 48–55, Madrid, Spain, 1997.
- [Rikitoku 07]
 M. Rikitoku. F-term classification Experiments at NTCIR-6 for Justsystems. In *Proceedings of the Sixth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access*, pages 420–427, Tokyo, Japan, 2007.
- [Riloff & Jones 99]
 E. Riloff and R. Jones. Learning Dictionaries for Information Extraction by Multi-Level Bootstrapping. In *Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI'99)*, 1999.
- [Riloff & Lehnert 92]
 E. Riloff and W. Lehnert. Classifying texts using relevancy signatures. In *Proc. of AAAI*, 1992.
- [Riloff & Lehnert 93]
 E. Riloff and W. Lehnert. Automated dictionary construction for information extraction from text. In *Proceedings of Ninth IEEE Conference on Artificial Intelligence for Applications*, pages 93–99, 1993.
- [Riloff & Lehnert 94]
 E. Riloff and W. Lehnert. Information extraction as a basis for high-precision text classification. *ACM Transactions on Information Systems*, pages 296–233, 1994.
- [Riloff & Shoen 95]
 E. Riloff and J. Shoen. Automatically acquiring conceptual patterns without an annotated corpus. In *Proceedings of the Third Workshop on Very Large Corpora*, 1995.
- [Riloff 93a]
 E. Riloff. Automatically constructing a dictionary for information extraction tasks. *Proceedings of the Eleventh Annual Conference on Artificial Intelligence*, pages 811–816, 1993.
- [Riloff 93b]
 E. Riloff. Automatically constructing a dictionary for information extraction tasks. In *Proceedings of Eleventh National Conference on Artificial Intelligence*, 1993.
- [Riloff 96a]
 E. Riloff. Automatically generating extraction patterns from untagged text. *Proceedings of the Thirteenth Annual Conference on Artificial Intelligence*, pages 1044–1049, 1996.
- [Riloff 96b]
 E. Riloff. An empirical study of automated dictionary construction for information extraction in three domains. *AI Journal*, 85, 1996.
- [Riloff *et al.* 02]
 E. Riloff, C. Schafer, and D. Yarowsky. Inducing information extraction systems for new languages via cross-language projection. In *Proceedings of the 19th international conference on Computational linguistics*, pages 1–7, Morristown, NJ, USA, 2002. Association for Computational Linguistics.
- [Riloff *et al.* 03]
 E. Riloff, J. Wiebe, and T. Wilson. Learning Subjective Nouns Using Extraction Pattern Bootstrapping. In *Proceedings of the Seventh Conference on Natural Language Learning (CoNLL-03)*, 2003.

- [Riloff *et al.* 06]
 E. Riloff, S. Patwardhan, and J. Wiebe. Feature Subsumption for Opinion Analysis. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP-06)*, 2006.
- [Rindflesch & Aronson 93]
 T. Rindflesch and A. Aronson. Semantic processing in information retrieval. In C. Safran, editor, *Proc. of 17th Annual Symposium on Computer Applications in Medical Care*, pages 611–615, 1993.
- [Rindflesch & Aronson 94]
 T. Rindflesch and A. Aronson. Ambiguity resolution while mapping free text to the umls metathesaurus. In J. Ozbolt, editor, *Proc. of 18th Annual Symposium on Computer Applications in Medical Care*, pages 240–244, 1994.
- [Risse *et al.* 11]
 T. Risse, S. Dietze, D. Maynard, and N. Tahmasebi. Using Events for Content Appraisal and Selection in Web Archives. In *Proceedings of DeRiVE 2011: Workshop in conjunction with the 10th International Semantic Web Conference 2011*, volume 779, pages 98–107, Bonn, Germany, October 2011. Marieke van Erp, Willem Robert van Hage, Laura Hollink, Anthony Jameson, Raphaël Troncy.
- [Ritchie 94a]
 G. Ritchie. Definite clause grammars with automatic structure binding. Unpublished software manual available in Department of Artificial Intelligence, 1994.
- [Ritchie 94b]
 G. Ritchie. Rules for translating syntax trees. Unpublished software manual available in Department of Artificial Intelligence, 1994.
- [Ritter *et al.* 11a]
 A. Ritter, S. Clark, Mausam, and O. Etzioni. Named entity recognition in tweets: An experimental study. In *Proc. of Empirical Methods for Natural Language Processing (EMNLP)*, Edinburgh, UK, 2011.
- [Ritter *et al.* 11b]
 A. Ritter, S. Clark, Mausam, and O. Etzioni. Named entity recognition in tweets: An experimental study. In *Proc. EMNLP*, 2011.
- [Rivest 92]
 R. Rivest. The MD5 message-digest algorithm. RFC 1321, Internet Engineering Task Force, April 1992.
- [Rizzo & Troncy 11]
 G. Rizzo and R. Troncy. NERD: A framework for evaluating named entity recognition tools in the Web of data. In *ISWC 2011, 10th International Semantic Web Conference*, Bonn, Germany, 2011.
- [Rizzo *et al.* 12]
 G. Rizzo, R. Troncy, S. Hellmann, and M. Bruemmer. NERD meets NIF: Lifting NLP extraction results to the Linked Data cloud. In *5th Workshop on Linked Data on the Web (LDoW)*, 2012.
- [Roberts *et al.* 08]
 A. Roberts, R. Gaizauskas, M. Hepple, and Y. Guo. Combining terminology resources and statistical methods for entity recognition: an evaluation. *Proceedings of the Conference on Language Resources and Evaluation (LRE'08)*, 2008.
- [Robertson *et al.* 95]
 S. E. Robertson, S. Walker, S. Jones, M. M. Hancock-Beaulieu, and M. Gatford. Okapi at trec-3. *NIST SPECIAL PUBLICATION SP*, pages 109–109, 1995.

- [Robin & McKeown 96]
 J. Robin and K. McKeown. Empirically designing and evaluating a new revision-based model for summary generation. *Artificial Intelligence*, 85(1-2), 1996.
- [Robinson & Fallside 87]
 A. J. Robinson and F. Fallside. Static and dynamic error propagation networks with application to speech coding. In D. Z. Anderson, editor, *Advances in Neural Information Processing Systems*, pages 632–641. American Institute of Physics, New York, 1987.
- [Robinson & Fallside 89]
 A. J. Robinson and F. Fallside. A dynamic connectionist model for phoneme recognition. In L. Personnaz and G. Dreyfus, editors, *Neural Networks: From Models to Applications*, pages 541–550. I. D. S. E. T., Paris, 1989.
- [Robinson & Fallside 90]
 A. J. Robinson and F. Fallside. Phoneme recognition from the TIMIT database using recurrent error propagation networks. Technical Report CUED/F-INFENG/TR.42, Cambridge University Engineering Department, 1990.
- [Robinson 94]
 A. J. Robinson. The application of recurrent nets to phone probability estimation. *IEEE Transactions on Neural Networks*, 5:298–305, 1994.
- [Robinson *et al.* 88]
 A. J. Robinson, M. Niranjan, and F. Fallside. Generalising the nodes of the error propagation network. Technical Report CUED/F-INFENG/TR.25, Cambridge University Engineering Department, 1988.
- [Robinson *et al.* 93]
 A. J. Robinson, L. Almeida, J.-M. Boite, H. Bourlard, F. Fallside, M. Hochberg, D. Kershaw, P. Kohn, Y. Konig, N. Morgan, J. P. Neto, S. Renals, M. Saerens, and C. Wooters. A neural network based, speaker independent, large vocabulary, continuous speech recognition system: the WERNICKE project. In *Proceedings European Conference on Speech Communication and Technology*, pages 1941–1944, Berlin, 1993.
- [Robinson *et al.* 95a]
 T. Robinson, J. Fransen, D. Pye, J. Foote, and S. Renals. WSJCAM0: A British English speech corpus for large vocabulary continuous speech recognition. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, 1995.
- [Robinson *et al.* 95b]
 T. Robinson, J. Fransen, D. Pye, J. Foote, and S. Renals. WSJCAM0: A British English speech corpus for large vocabulary continuous speech recognition. volume 1, pages 81–84, Detroit, 1995.
- [Robinson *et al.* 96]
 T. Robinson, M. Hochberg, and S. Renals. The use of recurrent networks in continuous speech recognition. In C. H. Lee, K. K. Paliwal, and F. K. Soong, editors, *Automatic Speech and Speaker Recognition – Advanced Topics*, chapter 10, pages 233–258. Kluwer Academic Publishers, Amsterdam, 1996.
- [Robinson-García *et al.* 14]
 N. Robinson-García, D. Torres-Salinas, Z. Zahedi, and R. Costas. New data, new possibilities: exploring the insides of altmetric.com. *arXiv preprint arXiv:1408.0135*, 2014.
- [Rocchini *et al.* 01]
 C. Rocchini, P. Cignoni, C. Montani, P. Pingi, and R. Scopigno. A low cost 3D scanner based on structured light. In A. Chalmers and T.-M. Rhyne, editors, *EG 2001 Proceedings*, volume 20(3), pages 299–308. Blackwell Publishing, 2001.

- [Rocha *et al.* 94]
R. Rocha, B. Rocha, and S. Huff. Automated translation between medical vocabularies using a frame-based interlingua. In *Proc. of SCAMC '94*, pages 690–694, 1994.
- [Rocha *et al.* 04]
C. Rocha, D. Schwabe, and M. P. Aragao. A hybrid approach for searching in the semantic web. In *Proceedings of the World Wide Web Conference*, 2004.
- [Rocha *et al.* 15]
L. Rocha, F. Mourão, T. Silveira, R. Chaves, G. Sá, F. Teixeira, R. Vieira, and R. Ferreira. Saci: Sentiment analysis by collective inspection on social media content. *Web Semantics: Science, Services and Agents on the World Wide Web*, 2015.
- [Roche & Schabes 95]
E. Roche and Y. Schabes. Deterministic Part-of-Speech Tagging with Finite-State Transducers. *Computational Linguistics*, 21(2):227–254, June 1995.
- [Roche & Schabes 97]
E. Roche and Y. Schabes, editors. *Finite-State Language Processing*. MIT Press, Cambridge, MA, 1997.
- [Rodgers *et al.* 97]
P. Rodgers, R. Gaizauskas, K. Humphreys, and H. Cunningham. Visual Execution and Data Visualisation in Natural Language Processing. In *IEEE Visual Language*, Capri, Italy, 1997.
- [Rodríguez & Araujo 02]
M. G. Rodríguez and C. P. S. Araujo, editors. Canary Islands, Spain, May 2002.
- [Rodriguez-Esteban 09]
R. Rodriguez-Esteban. Biomedical text mining and its applications. *PLoS Comput Biol*, 5(12):e1000597, 12 2009.
- [Roeder *et al.* 10]
C. Roeder, C. Jonquet, N. H. Shah, W. A. Baumgartner, K. Verspoor, and L. Hunter. A uima wrapper for the ncbo annotator. *Bioinformatics*, 26(14):1800–1801, 2010.
- [Romero *et al.* 11]
D. M. Romero, W. Galuba, S. Asur, and B. A. Huberman. Influence and passivity in social media. In *Machine learning and knowledge discovery in databases*, pages 18–33. Springer, 2011.
- [Rondeau 87]
G. Rondeau. Introduction to terminology. In J. C. Sager, editor, *Readings in Terminology*. CCL, UMIST, 1987. (English version of Chapters 1-5, edited by J.C. Sager 1980).
- [Roque *et al.* 11]
F. S. Roque, P. B. Jensen, H. Schmock, M. Dalgaard, M. Andreatta, T. Hansen, K. Søbey, S. Bredkjær, A. Juul, T. Werge, L. J. Jensen, and S. Brunak. Using electronic patient records to discover disease correlations and stratify patient cohorts. *PLoS Comput Biol*, 7(8):e1002141, 08 2011.
- [Rosch & Lloyd 78]
E. Rosch and B. Lloyd. *Cognition and categorization*. Lawrence Erlbaum Associates, Hillsdale, N.J., 1978.
- [Rosch 78]
E. Rosch. Principles of categorization. In E. Rosch and B. B. Lloyd, editors, *Cognition and Categorization*. Erlbaum, Hillsdale, NJ, 1978.

- [Roscheisen & Winograd 97]
M. Roscheisen and T. Winograd. A Network-Centric Design for Relationshipbased Security and Access Control. *Journal of Computer Security. Special Issue on Security in the World-Wide Web.*, 1997.
- [Rose *et al.* 02]
T. Rose, M. Stevenson, and M. Whitehead. The Reuters Corpus Volume 1 - from Yesterday's News to Tomorrow's Language Resources. In *LREC*, volume 2, pages 827–832, 2002.
- [Rosenblatt 58]
F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408, 1958.
- [Rosenbloom *et al.* 11]
S. T. Rosenbloom, J. C. Denny, H. Xu, N. Lorenzi, W. W. Stead, and K. B. Johnson. Data from clinical notes: a perspective on the tension between structure and flexible documentation. *J Am Med Inform Assoc*, 18(2):181–186, 2011.
- [Rosenfeld 95]
R. Rosenfeld. The CMU statistical language modelling toolkit and its use in the 1994 ARPA CSR evaluation. In *Proceedings of the ARPA Spoken Language Technology Workshop*, Austin TX, 1995.
- [Rösner & Kunze 02]
D. Rösner and M. Kunze. An XML-based document suite. In *Proceedings of the 19th International Conference on Computational Linguistics (COLING'02)*, Taipei, Taiwan, 2002.
- [Ross 82]
K. Ross. An improved left-corner parsing algorithm. COLING, 1982.
- [Ross *et al.* 98]
S. Ross, M. Moss, and L. Richmond. Planning and Designing a Programme of Digital Preservation Studies. In *Proceedings of the Electronic Access: Archives In The New Millennium conference*, pages 119–10. Public Record Office, June 1998.
- [Rossi & Magnani 12]
L. Rossi and M. Magnani. Conversation practices and network structure in twitter. In *ICWSM*, 2012.
- [Roth & Klakow 10]
B. Roth and D. Klakow. Combining wikipedia-based concept models for cross-language retrieval. In H. Cunningham, A. Hanbury, and S. Rüger, editors, *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference)*. LNCS volume number: 6107, Lecture Notes in Computer Science, Vienna, Austria, May 2010. Springer.
- [Roth & van denBosch (eds.) 02]
D. Roth and A. van den Bosch (eds.). *Proceedings of CoNLL-2002*. 2002.
- [Roth & Yih 01]
D. Roth and W. T. Yih. Relational Learning via Propositional Algorithms: An Information Extraction Case Study. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1257–1263, 2001.
- [Rotolo *et al.* 14]
D. Rotolo, I. Rafols, M. M. Hopkins, and L. Leydesdorff. Scientometric mapping as a strategic intelligence tool for the governance of emerging technologies. 2014.

- [Rousu *et al.* 05]
 J. Rousu, C. Saunders, S. Szedmak, and J. Shawe-Taylor. Learning Hierarchical Multi-Category Text Classification Models. In *Proceedings of 22nd International Conference on Machine Learning (ICML 2005)*, Bonn, Germany, 2005.
- [Rousu *et al.* 06]
 J. Rousu, C. Saunders, S. Szedmak, and J. Shawe-Taylor. Learning Hierarchical Multi-Category Text Classification Models. *Journal of Machine Learning Research*, 7:1601–1626, 2006.
- [Rout & Bontcheva 15]
 D. Rout and K. Bontcheva. A human-annotated dataset for evaluating tweet ranking algorithms. In *Proceedings of the 26th ACM Conference on Hypertext and Social Media*, 2015.
- [Rout *et al.* 13a]
 D. Rout, K. Bontcheva, and M. Hepple. Reliably evaluating summaries of twitter timelines. In *Proceedings of the AAAI Symposium on Analyzing Microtext*, 2013.
- [Rout *et al.* 13b]
 D. Rout, D. Preotiuc-Pietro, K. Bontcheva, and T. Cohn. Where’s @wally? a classification approach to geolocating users based on their social ties. In *Proceedings of the 24th ACM Conference on Hypertext and Social Media*, 2013.
- [Rowe & Alani 14]
 M. Rowe and H. Alani. Mining and comparing engagement dynamics across multiple social media platforms. In *Proceedings of the 2014 ACM conference on Web science*, pages 229–238. ACM, 2014.
- [Rowe & Stankovic 09]
 M. Rowe and M. Stankovic. Aligning Tweets with Events : Automation via Semantics. *Semantic Web*, 1, 2009.
- [Rowe *et al.* 11]
 M. Rowe, S. Anceletou, and H. Alani. Predicting discussions on the social semantic web. In *Proceedings of the 8th Extended Semantic Web Conference on the Semantic Web, ESWC’11*, pages 405–420. Springer-Verlag, 2011.
- [Rowe *et al.* 13]
 M. Rowe, M. Stankovic, A. Dadzie, B. Nunes, and A. Cano. Making sense of microposts (#msm2013): Big things come in small packages. In *Proceedings of the WWW Conference - Workshops*, 2013.
- [Rowe *et al.* 15]
 M. Rowe, M. Stankovic, and A.-S. Dadzie. #Microposts2015 – 5th Workshop on ‘Making Sense of Microposts’: Big things come in small packages. In *Proceedings of the 24th International Conference on World Wide Web Companion*, pages 1551–1552. International World Wide Web Conferences Steering Committee, 2015.
- [Royauté & Jacquemin 93]
 J. Royauté and C. Jacquemin. Indexation automatique et recherche des noms composés sous leurs différentes variations. In *Actes, Colloque Informatique et Langue Naturelle (ILIN)*, IRIN, Nantes, France, 1993.
- [Rubinoff 92]
 R. Rubinoff. Integrating text planning and linguistic choice. In *Aspects of Automated Natural Language Generation* [Dale *et al.* 92], pages 45–56.
- [Ruecker 08]
 B. Ruecker. Building an open source business process simulation tool with jboss jbpmp. Unpublished M.Sc. thesis, Stuttgart University of applied science, 2008.

- [Ruge 91]
G. Ruge. Experiments on linguistically-based term associations. In *Proc. of RIAO '91*, pages 528–545, Barcelona, 1991.
- [Russell & Norvig 95]
S. Russell and P. Norvig. *Artificial Intelligence - A modern Approach*. Prentice Hall International Editions, Upper Saddle River, NJ, 1995.
- [Rzhetsky *et al.* 09]
A. Rzhetsky, M. Seringhaus, and M. B. Gerstein. Getting started in text mining: Part two. *PLoS Comput Biol*, 5(7):e1000411, 07 2009.
- [Sabou & Pan 05]
M. Sabou and J. Pan. Towards Improving Web Service Repositories through Semantic Web Techniques. In *Workshop on Semantic Web Enabled Software Engineering (SWESE)*, Galway, Ireland, 2005.
- [Sabou 04]
M. Sabou. From Software APIs to Web Service Ontologies: a Semi-Automatic Extraction Method. In *Proceedings of the 3rd International Semantic Web Conference (ISWC'04)*, 2004.
- [Sabou 05]
M. Sabou. Visual Support for Ontology Learning: an Experience Report. In *Proceedings of the 9th International Conference on Information Visualisation (IV05)*, 2005.
- [Sabou 06]
M. Sabou. *Building Web Service Ontologies*. Unpublished PhD thesis, Vrije Universiteit, 2006.
- [Sabou *et al.* 05]
M. Sabou, C. Wroe, C. Goble, and G. Mishne. Learning domain ontologies for web service descriptions: an experiment in bioinformatics. In *WWW '05: Proceedings of the 14th international conference on World Wide Web*, pages 190–198, New York, NY, USA, 2005. ACM.
- [Sabou *et al.* 12]
M. Sabou, K. Bontcheva, and A. Scharl. Crowdsourcing research opportunities: Lessons from natural language processing. In *12th International Conference on Knowledge Management and Knowledge Technologies (I-KNOW)*, pages 17–24, 2012.
- [Sabou *et al.* 13]
M. Sabou, K. Bontcheva, A. Scharl, and M. Föls. Games with a Purpose or Mechanised Labour? A Comparative Study. In *Proc. of the 13th International Conference on Knowledge Management and Knowledge Technologies (i-KNOW)*. ACM, 2013.
- [Sabou *et al.* 14]
M. Sabou, K. Bontcheva, L. Derczynski, and A. Scharl. Corpus annotation through crowdsourcing: Towards best practice guidelines. In *Proceedings of the 9th international conference on language resources and evaluation (LREC'14)*, pages 859–866, 2014.
- [Sadek & De Mori 98]
D. Sadek and R. De Mori. *Spoken Dialogues with Computers*. Academic Press Limited, 24-28 Oval Road, London NW1 7DX, UK, 1998.
- [Sadilek *et al.* 12]
A. Sadilek, H. Kautz, and V. Silenzio. Modeling spread of disease from social interactions. In *International AAAI Conference on Weblogs and Social Media (ICWSM)*, pages 322–329. AAAI, 2012.

- [Saeys *et al.* 07]
 Y. Saeys, I. Inza, and P. Larrañaga. A review of feature selection techniques in Bioinformatics. *bioinformatics*, 23(19):2507–2517, 2007.
- [Sager 72]
 N. Sager. Syntactic formatting of scientific information. In *Proc. of 1972 Fall Joint Computer Conference*, volume 41 of *AFIPS Conf. Proc.*, pages 791–800, Montvale, NJ, 1972.
- [Sager 78a]
 J. Sager. Commentary. In *Table Ronde sur les Problèmes du Découragement du Terme*, pages 39–52. Service des Publications, Direction des Française, Montreal, Canada, 1978.
- [Sager 78b]
 N. Sager. Natural Language Information Formatting: The Automatic Conversion of Texts to a Structured Data Base. *Advances in Computers*, 17:89–162, 1978.
- [Sager 81]
 N. Sager. *Natural Language Information Processing*. Addison-Wesley, Reading, MA, 1981.
- [Sager 90]
 J. Sager. *A Practical Course in Terminology Processing*. John Benjamins, 1990.
- [Sager 94]
 J. Sager. *Language Engineering and Translation: consequences of automation*. John Benjamins, 1994.
- [Sager 97]
 J. C. Sager. Term formation. In S. Wright and G. Budin, editors, *Handbook of Terminology Management*, volume 1: Basic Aspects of Terminology Management, chapter 1.2.1, pages 25–41. John Benjamins, Amsterdam, 1997.
- [Sagerer 90]
 G. Sagerer. *Automatisches Verstehen gesprochener Sprache*. BI Wissenschaftsverlag, Bibliographisches Institut & F.A. Brockhaus AG, Mannheim, 1990.
- [Saggion & Funk 09]
 H. Saggion and A. Funk. Extracting opinions and facts for business intelligence. *RNTI Journal*, E(17):119–146, November 2009.
- [Saggion & Gaizauskas 04a]
 H. Saggion and R. Gaizauskas. Mining on-line sources for definition knowledge. In *Proceedings of the 17th FLAIRS 2004*, Miami Beach, Florida, USA, May 17-19 2004. AAAI.
- [Saggion & Gaizauskas 04b]
 H. Saggion and R. Gaizauskas. Multi-document summarization by cluster/profile relevance and redundancy removal. In *Proceedings of the Document Understanding Conference 2004*. NIST, 2004.
- [Saggion & Gaizauskas 05]
 H. Saggion and R. Gaizauskas. Experiments on statistical and pattern-based biographical summarization. In *Proceedings of EPIA 2005*, pages 611–621, 2005.
- [Saggion & Lapalme 00a]
 H. Saggion and G. Lapalme. Concept Identification and Presentation in the Context of Technical Text Summarization. In *Proceedings of the Workshop on Automatic Summarization. ANLP-NAACL2000*, Seattle, WA, USA, 30 April 2000. Association for Computational Linguistics.

- [Saggion & Lapalme 00b]
H. Saggion and G. Lapalme. Summary Generation and Evaluation in SumUM. In *Advances in Artificial Intelligence. International Joint Conference: 7th Ibero-American Conference on Artificial Intelligence and 15th Brazilian Symposium on Artificial Intelligence. IBERAMIA-SBIA 2000.*, volume 1952 of *Lecture Notes in Artificial Intelligence*, pages 329–38, Berlin, Germany, 2000. Springer-Verlag.
- [Saggion & Lapalme 02]
H. Saggion and G. Lapalme. Generating Indicative-Informative Summaries with SumUM. *Computational Linguistics*, 2002.
- [Saggion 04]
H. Saggion. Identifying definitions in text collections for question answering. *lrec*. In *Proceedings of Language Resources and Evaluation Conference*. ELDA, 2004.
- [Saggion 06]
H. Saggion. Multilingual Multidocument Summarization Tools and Evaluation. In *Proceedings of LREC 2006*, 2006.
- [Saggion 07]
H. Saggion. Shef: Semantic tagging and summarization techniques applied to cross-document coreference. In *Proceedings of SemEval 2007, Association for Computational Linguistics*, pages 292–295, June 2007.
- [Saggion 08]
H. Saggion. Summa: A robust and adaptable summarization tool. *Traitement Automatique des Langues*, 49(2), 2008.
- [Saggion et al. 01]
H. Saggion, H. Cunningham, D. Maynard, and Y. Wilks. Using the GATE architecture for named entity recognition in the football domain. In *2nd International Workshop on Spanish Language Processing and Language technologies (SPLPT2)*, pages 141–147, Jaen, Spain, 2001.
- [Saggion et al. 02a]
H. Saggion, H. Cunningham, Y. Wilks, T. Declerck, and P. Wittenburg. English Information Extraction from Football Reports in the Context of the MUMIS Project. *ELSNews*, 11(2), 2002.
- [Saggion et al. 02b]
H. Saggion, H. Cunningham, K. Bontcheva, D. Maynard, C. Ursu, O. Hamza, and Y. Wilks. Access to Multimedia Information through Multisource and Multilanguage Information Extraction. In *Proceedings of the 7th Workshop on Applications of Natural Language to Information Systems (NLDB 2002)*, Stockholm, Sweden, 2002.
- [Saggion et al. 02c]
H. Saggion, H. Cunningham, D. Maynard, K. Bontcheva, O. Hamza, C. Ursu, and Y. Wilks. Extracting Information for Automatic Indexing of Multimedia Material. In *3rd International Conference on Language Resources and Evaluation (LREC 2002)*, pages 669–676, Las Palmas, Gran Canaria, Spain, 2002.
- [Saggion et al. 02d]
H. Saggion, H. Cunningham, D. Maynard, K. Bontcheva, O. Hamza, C. Ursu, and Y. Wilks. Extracting Information for Information Indexing of Multimedia Material. In *Proceedings of 3rd Language Resources and Evaluation Conference (LREC'2002)*, 2002. http://gate.ac.uk/sale/lrec2002/mumis_lrec2002.ps.
- [Saggion et al. 02e]
H. Saggion, D. Radev, S. Teufel, L. Wai, and S. Strassel. Developing Infrastructure for the Evaluation of Single and Multi-document Summarization Systems in a Cross-lingual Environment. In *3rd*

International Conference on Language Resources and Evaluation (LREC 2002), pages 747–754, Las Palmas, Gran Canaria, Spain, 2002.

[Saggion *et al.* 03a]

H. Saggion, K. Bontcheva, and H. Cunningham. Robust Generic and Query-based Summarisation. In *Proceedings of the European Chapter of Computational Linguistics (EACL), Research Notes and Demos*, 2003.

[Saggion *et al.* 03b]

H. Saggion, H. Cunningham, K. Bontcheva, D. Maynard, O. Hamza, and Y. Wilks. Multimedia Indexing through Multisource and Multilingual Information Extraction; the MUMIS project. *Data and Knowledge Engineering*, 48:247–264, 2003.

[Saggion *et al.* 03c]

H. Saggion, J. Kuper, H. Cunningham, T. Declerck, P. Wittenburg, M. Puts, F. DeJong, and Y. Wilks. Event-coreference across Multiple, Multi-lingual Sources in the Mumis Project. In *Proceedings of the European Chapter of Computational Linguistics (EACL), Research Notes and Demos*, 2003.

[Saggion *et al.* 03d]

H. Saggion, J. Kuper, T. Declerck, D. Reidsma, and H. Cunningham. Intelligent Multimedia Indexing and Retrieval through Multi-source Information Extraction and Merging. In *IJCAI 2003*, Acapulco, Mexico, 2003.

[Saggion *et al.* 07]

H. Saggion, A. Funk, D. Maynard, and K. Bontcheva. Ontology-based information extraction for business applications. In *Proceedings of the 6th International Semantic Web Conference (ISWC 2007)*, Busan, Korea, November 2007.

[Sahlgren & Karlgren 08]

M. Sahlgren and J. Karlgren. Buzz monitoring in word space. In *European Conference on Intelligence and Security Informatics (EuroISI 2008)*, Esbjerg, Denmark, 2008.

[Sahlgren 05]

M. Sahlgren. An introduction to random indexing. In *Proc. of the Methods and Applications of Semantic Indexing Workshop*, Copenhagen, Denmark, 2005.

[Sahlgren *et al.* 08]

M. Sahlgren, A. Holst, and P. Kanerva. Permutations as a means to encode order in word space. In *Proceedings of the 30th Annual Meeting of the Cognitive Science Society (CogSci'08)*, Washington D.C., USA, 2008.

[SAIC 98]

SAIC. Proceedings of the Seventh Message Understanding Conference (MUC-7). http://www.itl.nist.gov/iaui/894.02/related_projects/muc/index.html, 1998.

[Saif *et al.* 12]

H. Saif, Y. He, and H. Alani. Alleviating data sparsity for Twitter sentiment analysis. In *Proceedings of the #MSM2012 Workshop, CEUR*, volume 838, 2012.

[Saif *et al.* 13]

H. Saif, M. Fernandez, Y. He, and H. Alani. Evaluation datasets for twitter sentiment analysis: a survey and a new dataset, the sts-gold. 2013.

[Sairamesh *et al.* 96]

J. Sairamesh, S. Kapidakis, and C. Nikolaou. Architectures for QoS based Retrieval in Digital Libraries. In *Proceedings of the Workshop on Networked Information Retrieval, SIGIR'96.*, 1996.

- [Saito & Nagata 03]
K. Saito and M. Nagata. Multi-Language Entity Recognition System based on HMM. In *ACL Workshop on Multilingual and Mixed-language Named Entity Recognition: Combining Statistical and Symbolic Models*, Sapporo, Japan, 2003.
- [Sakaki *et al.* 10a]
T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World Wide Web (WWW)*, pages 851–860. ACM, 2010.
- [Sakaki *et al.* 10b]
T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web, WWW '10*, pages 851–860, New York, NY, USA, 2010. ACM.
- [Salton & McGill 83]
G. Salton and M. J. McGill. *Introduction to Modern Information Retrieval*. McGraw-Hill, 1983.
- [Salton 70]
G. Salton. Automatic Processing of Foreign Language Documents. *Journal of the American Society for Information Sciences*, 21:187–194, 1970.
- [Salton 88]
G. Salton. *Automatic Text Processing*. Addison-Wesley Publishing Company, 1988.
- [Salton 89]
G. Salton. *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*. Addison-Wesley, Reading MA, 1989.
- [Salton *et al.* 75]
G. Salton, A. Wong, and C.-S. Yang. A vector space model for automatic indexing. *Communications of the ACM*, 18(11):613–620, 1975.
- [Sammon 69]
J. Sammon, Jr. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers*, 18:401–409, 1969.
- [Sampson 97]
G. Sampson. *Educating Eve*. Cassell, London, 1997.
- [Sampson 98]
G. Sampson. Review of: Greenbaum (ed.), *Comparing English Worldwide*. *Journal of Natural Language Engineering*, 4(4):363–365, 1998.
- [Sanderson 94]
M. Sanderson. Word sense disambiguation and information retrieval. In *Proc. of the 17th ACM/SIGIR Conference*, pages 142–150, 1994.
- [Sapcota 05]
B. Sapcota. Web Service Discovery in Distributed and Heterogeneous Environments. In *International Conference on Web Intelligence (WI'05)*, Compiegne, France, 2005.
- [Sapsford & Jupp 96]
R. Sapsford and V. Jupp, editors. *Data Collection and Analysis*. Open University, 1996.
- [Sapsford 96]
R. Sapsford. Extracting and presenting statistics. In Sapsford and Jupp [Sapsford & Jupp 96], chapter 8.

- [Sarawagi 08]
S. Sarawagi. Information Extraction. *Foundations and Trends in Databases*, 1(3):261–377, 2008.
- [Sarwar *et al.* 01]
B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th International Conference on World Wide Web*, pages 285–295. ACM, 2001.
- [Sassano 02]
M. Sassano. An Empirical Study of Active Learning with Support Vector Machines for Japanese Word Segmentation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2002.
- [Sassone 87]
P. Sassone. Cost-benefit methodology for office systems. *ACM Transactions on Office Information Systems*, 5(3):273–289, 1987.
- [Saurí *et al.* 05]
R. Saurí, R. Knippen, M. Verhagen, and J. Pustejovsky. Evita: a robust event recognizer for qa systems. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 700–707. Association for Computational Linguistics, 2005.
- [Savage 12]
N. Savage. Gaining wisdom from crowds. *Commun. ACM*, 55(3):13–15, March 2012.
- [Sayeed 13]
A. Sayeed. An opinion about opinions about opinions: subjectivity and the aggregate reader. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 691–696, Atlanta, Georgia, June 2013. Association for Computational Linguistics.
- [Sayyadi *et al.* 09]
H. Sayyadi, M. Hurst, and A. Maykov. Event Detection and Tracking in Social Streams. In *Proceedings of the Third International ICWSM Conference*, pages 311–314, 2009.
- [Sazedj & Pinto 05]
P. Sazedj and H. S. Pinto. Time to evaluate: Targeting annotation tools. In *Proceedings of ISWC 2005 SemAnnot Workshop*, Galway, Ireland, 2005.
- [Sazedj & Pinto 07]
P. Sazedj and H. S. Pinto. Mining the Web Through Verbs: A Case Study. In *Proceedings of the European Semantic Web Conference (ESWC'07), LNCS 4519*, pages 488–502, 2007.
- [S.Carberry *et al.* 98]
S.Carberry, K.Samuel, and K.Vijay-Shanker. Dialogue act tagging with transformation-based learning. In *Proceedings of COLING-ACL*, volume 2, pages 1150–1156, Montreal, Canada, 1998.
- [Scerri *et al.* 12]
S. Scerri, K. Cortis, I. Rivera, and S. Handschuh. Knowledge Discovery in Distributed Social Web Sharing Activities. In *Proceedings of the #MSM2012 Workshop, CEUR*, volume 838, 2012.
- [Schäfer 03]
U. Schäfer. WHAT: An XSLT-based Infrastructure for the Integration of Natural Language Processing Components. In *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 9–16, 2003.
- [Schaffert 06]
S. Schaffert. IkeWiki - A Semantic Wiki for Collaborative Knowledge Management. In *Semantic Technologies in Collaborative Applications (STICA'06)*, 2006.

- [S.Chakrabarti *et al.* 98]
 S.Chakrabarti, B.Dom, D.Gibson, J.Kleinberg, P. Raghavan, and S. Rajagopalan. Automatic resource compilation by analyzing hyperlink structure and associated text. In *Seventh International Web Conference (WWW 98)*, Brisbane, Australia, 1998.
- [Schank & Abelson 77]
 R. Schank and R. Abelson. *Scripts, Plans, Goals, and Understanding*. Lawrence Erlbaum Associates, Hillsdale, N.J., 1977.
- [Schank & Colby 73]
 R. Schank and M. C. Colby. *Computer Models of Thought and Language*. W.H. Freeman, San Francisco, 1973.
- [Schank & Riesbeck 81]
 R. Schank and C. Riesbeck. *Inside Computer Understanding Five Programs Plus Miniatures*. Lawrence Erlbaum Associates, Hillsdale, N.J., 1981.
- [Schank & Tesler 69]
 R. C. Schank and L. Tesler. A conceptual dependency parser for natural language. In *Proceedings of the 1969 conference on Computational linguistics*, pages 1–3. Association for Computational Linguistics, 1969.
- [Schank 75]
 R. Schank. *Conceptual Information Processing*. North-Holland, Amsterdam, 1975.
- [Schantl *et al.* 13]
 J. Schantl, R. Kaiser, C. Wagner, and M. Strohmaier. The utility of social and topical factors in anticipating repliers in twitter conversations. In *Proceedings of the 5th Annual ACM Web Science Conference, WebSci '13*, pages 376–385, New York, NY, USA, 2013. ACM.
- [Scharl & Weichselbraun 08]
 A. Scharl and A. Weichselbraun. An automated approach to investigating the online media coverage of US presidential elections. *Journal of Information Technology and Politics*, 5(1):121–132, 2008.
- [Scharl *et al.* 12]
 A. Scharl, M. Sabou, S. Gindl, W. Rafelsberger, and A. Weichselbraun. Leveraging the wisdom of the crowds for the acquisition of multilingual language resources. 2012.
- [Scheffer & Wrobel 01]
 T. Scheffer and S. Wrobel. Active Learning of Partially Hidden Markov Models. In *ECML-2001/PKDD-2001 Workshop on Active Learning, Database Sampling, Experimental Design: Views on Instance Selection*, 2001.
- [Scheffer *et al.* 02]
 T. Scheffer, S. Wrobel, B. Popov, D. Ognyanoff, C. Wiech, and S. Hoche. Learning hidden markov models for information extraction actively from partially labeled text. *Künstliche Intelligenz*, 2, 2002.
- [Schein *et al.* 02]
 A. I. Schein, A. Popescul, L. H. Ungar, and D. M. Pennock. Methods and metrics for cold-start recommendations. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '02*, pages 253–260, New York, NY, USA, 2002. ACM.
- [Schellner 02]
 I. Schellner. Japanese file index classification and f-terms. *World Patent Information*, 24:197–201, 2002.

- [Schentz *et al.* 11]
H. Schentz, J. Peterseil, B. Magagna, and M. Mirttil. Semantics in ecosystems research and monitoring. In W. Pillman, S. Schade, and P. Smits, editors, *Proceedings of the 25th International EnviroInfo Conference*, 2011.
- [Scherf *et al.* 05]
M. Scherf, A. Epple, and T. Werner. The next generation of literature analysis: Integration of genomic analysis into text mining. *Briefings in Bioinformatics*, 6(3):287–297, 2005.
- [Schick *et al.* 90]
A. Schick, L. Gordon, and S. Haka. Information overload: A temporal approach. *Accounting Organizations and Society*, 15:199–220, 1990.
- [Schilder *et al.* 08]
F. Schilder, R. Kondadadi, J. L. Leidner, , and J. G. Conrad. Thomson reuters at tac 2008: Aggressive filtering with fastsum for update and opinion summarization. In *Proceedings of the First Text Analysis Conference (TAC 2008)*, pages 396–405, 2008.
- [Schmid 94]
H. Schmid. Probabilistic part-of-speech tagging using decision trees. In *Proceedings of the international conference on new methods in language processing*, volume 12, pages 44–49. Citeseer, 1994.
- [Schneider & Rodd 01]
F. Schneider and M. Rodd, editors. *International Review of UK Research in Computer Science*. EPSRC, IEE and BCS, 2001.
- [Schneider 15]
N. Schneider. What i’ve learned about annotating informal text (and why you shouldn’t take my word for it). In *Proceedings of The 9th Linguistic Annotation Workshop*, pages 152–157. ACL, 2015.
- [Schohn & Cohn 00]
G. Schohn and D. Cohn. Less is More: Active Learning with Support Vector Machines. In *Proceedings of the Seventeenth International Conference on Machine Learning (ICML-00)*, 2000.
- [Schröder *et al.* 11]
M. Schröder, P. Baggia, F. Burkhardt, C. Pelachaud, C. Peter, and E. Zovato. EmotionML—an upcoming standard for representing emotions and related states. In *Affective Computing and Intelligent Interaction*, pages 316–325. Springer, 2011.
- [Schroeter & Hunter 07]
R. Schroeter and J. Hunter. Annotating Relationships between Multiple Mixed-media digital objects by extending Annotea. In *Proceedings of the European Semantic Web Conference (ESWC’2007)*, 2007.
- [Schubert & Pelletier 82]
L. Schubert and F. Pelletier. From english to logic: Context-free computation of ‘conventional’ logical translation. *American Journal of Computational Linguistics*, 8(1), 1982.
- [Schuler 05]
K. K. Schuler. *VerbNet: A broad-coverage, comprehensive verb lexicon*. Unpublished PhD thesis, University of Pennsylvania, 2005.
- [Schulze 87]
E. Schulze. The term, its properties and nature and its delimitation from other types of lexemes. In J. Sager, editor, *Readings in Terminology*. CCL, UMIST, 1987. English version.

- [Schürr *et al.* 95]
 A. Schürr, A. Winter, and A. Zündorf. Visual Programming with Graph Rewriting Systems. In *Proceedings VL'95 11th International IEEE Symposium on Visual Languages, Darmstadt*. IEEE Computer Society Press, 1995.
- [Schutz & Buitelaar 05]
 A. Schutz and P. Buitelaar. Relext: A tool for relation extraction from text in ontology extension. *The Semantic Web-ISWC 2005*, pages 593–606, 2005.
- [Schütz 94]
 J. Schütz. Developing Lingware in ALEP. *ALEP User Group News, CEC Luxembourg*, 1(1), October 1994.
- [Schutz 08a]
 A. Schutz. Keyphrase extraction from single documents in the open domain exploiting linguistic and statistical methods. MSc thesis, National University of Ireland, Galway, 2008.
- [Schutz 08b]
 A. Schutz. XtraK4Me: Extraction of keyphrases for metadata creation. GATE plug-in software, Semantic Information Systems and Language Engineering Group (SmILE), DERI, Galway, Ireland, August 2008.
- [Schütz *et al.* 91]
 J. Schütz, G. Thurmair, and R. Cencioni. An Architecture Sketch of Eurotra-II. In *MT Summit III*, pages 3–11, Washington D.C., 1991.
- [Schütze 92]
 H. Schütze. Dimensions of meaning. In *Proceedings of Supercomputing '92*, pages 787–796, Minneapolis, MN, 1992.
- [Schütze 98]
 H. Schütze. Automatic Word Sense Discrimination. *Computational Linguistics*, 24:97–124, 1998.
- [Schwaber & Beedle 01]
 K. Schwaber and M. Beedle. *Agile Software Development with Scrum*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2001.
- [Schwartz & Hearst 03]
 A. S. Schwartz and M. A. Hearst. A simple algorithm for identifying abbreviation definitions in biomedical text. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, pages 451–462, 2003.
- [Schwartz & Thurmair 96a]
 C. Schwartz and G. Thurmair. WP 4.1 Task S3: Basic Linguistic Tools, Global Functional Specifications. LE project LE1-2238 AVENTINUS internal technical report, GMS, Germany, July 1996.
- [Schwartz & Thurmair 96b]
 C. Schwartz and G. Thurmair. WP 4.1 Task T3: Term Substitution, Detailed Functional Specifications. LE project LE1-2238 AVENTINUS internal technical report, GMS, Germany, July 1996.
- [Schwartz *et al.* 96]
 R. Schwartz, L. Nguyen, and J. Makhoul. Multiple-pass search strategies. In C.-H. Lee, F. K. Soong, and K. K. Paliwal, editors, *Automatic Speech and Speaker Recognition – Advanced Topics*, chapter 18, pages 429–456. Kluwer Academic, Amsterdam, 1996.
- [Schwertel 00]
 U. Schwertel. Controlling plural ambiguities in Attempto Controlled English. In *Proceedings of the 3rd International Workshop on Controlled Language Applications*, Seattle, Washington, 2000.

- [Schwitter & Tilbrook 04]
R. Schwitter and M. Tilbrook. Controlled natural language meets the semantic web, 2004.
- [Schwitter 07]
R. Schwitter. Controlled natural languages. Technical report, Centre for Language Technology, Macquarie University, June 2007.
- [Scott & Gaizauskas. 00]
S. Scott and R. Gaizauskas. The University of Sheffield TREC-9 Q&A System. In *In Proceedings of the 9th Text REtrieval Conference*, 2000.
- [Scott 96]
M. Scott. *WordSmith Tools Manual*. Oxford University Press, 1996.
- [Scott 00]
M. Scott. Wordsmith. <http://www.liv.ac.uk/~ms2928>, October 2000.
- [Scott *et al.* 12]
D. Scott, R. Barone, and R. Koeling. Corpus annotation as a scientific task. In N. C. C. Chair), K. Choukri, T. Declerck, M. U. Doğan, B. Maegaard, J. Mariani, J. Odijk, and S. Piperidis, editors, *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey, may 2012. European Language Resources Association (ELRA).
- [Scotti & Nuzzo 01]
A. Scotti and D. Nuzzo. Pinakes – a modeling environment for scientific heritage database applications. In *Proc. of Reconstructing science – Contributions to the enhancement of the European scientific heritage Workshop*, Ravenna, Italy, 2001.
- [Searls 02]
D. B. Searls. The language of genes. *Nature*, 420:211–217, November 2002.
- [Searls 03]
D. B. Searls. Trees of life and of language. *Nature*, 426:391–392, November 2003.
- [Segel & Heer 10]
E. Segel and J. Heer. Narrative visualization: Telling stories with data. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1139–1148, 2010.
- [SEI 99]
SEI. How Do You Define Software Architecture. <http://www.sei.cmu.edu/architecture/definitions.html> visited March 29th 1999, March 1999.
- [Sekine *et al.* 02]
S. Sekine, K. Sudo, and C. Nobata. Extended named entity hierarchy. In Rodríguez and Araujo [Rodríguez & Araujo 02], pages 1818–1824.
- [Seneff & Polifroni 96]
S. Seneff and J. Polifroni. A New Restaurant Guide Conversational System: Issues In Rapid Prototyping For Specialized Domains. In *Int. Conf. on Spoken Language Processing (ICSLP)*, volume 2, pages 665–668, Philadelphia, Pennsylvania, USA, October 1996.
- [Seneff 92a]
S. Seneff. Tina: A natural language system fo spoken language applications. *Computational Linguistics*, 18(1):61–86, 1992.
- [Seneff 92b]
S. Seneff. TINA: A Natural Language System for Spoken Language Applications. *Computational Linguistics*, 18(1):61–86, 1992.

- [Seneff *et al.* 95]
S. Seneff, M. McCandless, and V. Zue. Integrating natural language into the word graph search for simultaneous speech recognition and understanding. In *Proceedings European Conference on Speech Communication and Technology*, pages 1781–1784, Madrid, 1995.
- [Settles 05]
B. Settles. ABNER: An open source tool for automatically tagging genes, proteins, and other entity names in text. *Bioinformatics*, 21(14):3191–3192, 2005.
- [Settles 12]
B. Settles. *Active Learning*. Morgan and Claypool, 2012.
- [Shah *et al.* 02]
U. Shah, T. Finin, A. Joshi, R. S. Cost, and J. Mayfield. Information retrieval on the semantic web. In *Proceedings of the 11th ACM Conference on Information and Knowledge Management (CIKM'02)*, Washington, USA, Nov 2002.
- [Shamima Mithun 07]
V. H. Shamima Mithun, Leila Kosseim. Resolving quantifier and number restriction to question owl ontologies. In *Proceedings of The First International Workshop on Question Answering (QA2007)*, Xi'an, China, October 2007.
- [Shamma *et al.* 10]
D. Shamma, L. Kennedy, and E. Churchill. Tweetgeist: Can the Twitter timeline reveal the structure of broadcast events? In *Proceedings of CSCW 2010*, 2010.
- [Shani & Gunawardana 11]
G. Shani and A. Gunawardana. Evaluating recommendation systems. In *Recommender systems handbook*, pages 257–297. Springer, 2011.
- [Shao *et al.* 16]
C. Shao, G. L. Ciampaglia, A. Flammini, and F. Menczer. Hoaxy: A platform for tracking online misinformation. In *WWW'2016*, pages 745–750, 2016.
- [Shaon *et al.* 11]
A. Shaon, A. Wolf, S. Crompton, R. Boczek, W. Rogers, and M. Jackson. An open source linked data framework for publishing environmental data under the UK location strategy. In *Terra Cognita 2011: Foundations, Technologies and Applications of the Geospatial Web*, pages 62–74, 2011.
- [Shapiro 82]
S. Shapiro. Generalized Augmented Transition Network Grammars for Generation from Semantic Networks. *American Journal of Computational Linguistics*, 8(2):12 – 25, 1982.
- [Sharifi *et al.* 10]
B. Sharifi, M. A. Hutton, and J. Kalita. Summarizing Microblogs Automatically. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 685–688, Los Angeles, California, June 2010.
- [Sharkey 92]
N. Sharkey, editor. *Connectionist Natural Language Processing*. Kluwer, Dordrecht, Netherlands, 1992.
- [Sharpe 98]
T. Sharpe. Use of internet resources by pre-clinical and clinical medical students. In *Proc. of 2nd Annual BMIS Conference: "Making Medical Informatist Work"*, Manchester, UK, 1998. British Medical Informatics Society.

- [Shatkay 05]
H. Shatkay. Hairpins in bookstacks: Information retrieval from biomedical text. *Briefings in Bioinformatics*, 6(3):222–238, 2005.
- [Shaver *et al.* 87]
P. Shaver, J. Schwartz, D. Kirson, and C. O’connor. Emotion knowledge: further exploration of a prototype approach. *Journal of personality and social psychology*, 52(6):1061, 1987.
- [Shaw & Clements 97]
M. Shaw and P. Clements. A Field Guide to Boxology: Preliminary Classification of Architectural Styles for Software Systems. In *Proceedings COMPSAC97, 21st Int’l Computer Software and Applications Conference*, pages 6–13, 1997.
- [Shaw & Gaines 88]
M. L. G. Shaw and B. R. Gaines. KITTEN: knowledge initiation and transfer tools for experts and novices. In B. R. Gaines and J. H. Boose, editors, *Knowledge-based Systems: Vol 2: Knowledge Acquisition Tools for Expert Systems*. Academic Press, London, 1988.
- [Shaw & Garlan 96]
M. Shaw and D. Garlan. *Software Architecture*. Prentice Hall, New York, 1996.
- [Shaw 93]
M. Shaw. Software Architecture: Beyond Objects. In *OOPSLA 93*, Washington, DC, 1993.
- [Shawe-Taylor & Cristianini 99]
J. Shawe-Taylor and N. Cristianini. Margin distribution bounds on generalization. In *Proceedings of European Conference on Computational Learning Theory, EuroCOLT’99*, pages 263–273, 1999.
- [Shchekotykhin *et al.* 07]
K. M. Shchekotykhin, D. Jannach, G. Friedrich, and O. Kozeruk. AllRight: Automatic Ontology Instantiation from Tabular Web Documents. In *Proceedings of ISWC/ASWC*, pages 466–479, 2007.
- [Sheffield 00]
Sheffield. The MUMIS Football English Corpus. Technical report, University of Sheffield, 2000.
- [Shen *et al.* 12a]
W. Shen, J. Wang, P. Luo, and M. Wang. LINDEN: Linking named entities with knowledge base via semantic knowledge. In *Proceedings of the 21st Conference on World Wide Web*, pages 449–458, 2012.
- [Shen *et al.* 12b]
W. Shen, J. Wang, P. Luo, and M. Wang. LINDEN: Linking named entities with knowledge base via semantic knowledge. In *Proc. of WWW*, pages 449–458, 2012.
- [Shen *et al.* 13]
W. Shen, J. Wang, P. Luo, and M. Wang. Linking named entities in tweets with knowledge base via user interest modeling. In *Proc. of the 19th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, pages 68–76. ACM, 2013.
- [Sheng *et al.* 08]
V. S. Sheng, F. Provost, and P. G. Ipeirotis. Get another label? improving data quality and data mining using multiple, noisy labelers. In *Proceedings of the 14th International Conference on Knowledge Discovery and Data Mining*, pages 614–622, 2008.
- [Sheremetyeva 03]
S. Sheremetyeva. Natural Language Analysis of Patent Claims. *Proceedings of the workshop “Patent Corpus Processing” in conjunction with 41 st Annual Meeting of the Association for Computational Linguistics (ACL 2003)*, pages 7–12, 2003.

- [Sheremetyeva *et al.* 96]
S. Sheremetyeva, S. Nirenburg, and I. Nirenburg. Generating patent claims from interactive input. In INLG'96 [INLG'96 96], pages 51–60.
- [Shieber 92]
S. Shieber. *Constraint-Based Grammar Formalisms*. MIT Press, Cambridge, MA, 1992.
- [Shirky 03]
C. Shirky. A Group Is Its Own Worst Enemy. http://www.shirky.com/writings/group_enemy.html, 2003.
- [Shneiderman 96]
B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings of the IEEE Symposium on Visual Languages*, pages 336–343, 1996.
- [Shu 88]
N. Shu. *Visual Programming*. Van Nostrand Reinhold, New York, 1988.
- [Siberski 05]
W. Siberski. Semantic web framework requirements analysis. Technical Report D1.2.2, KnowledgeWeb Deliverable, 2005.
- [SIC00]
SICSTUS. Intelligent Systems Laboratory. Swedish Institute of Computer Science, PO Box 1263. SE-164 29 Kista, Sweden. *SICStus Prolog User's Manual*, May 2000.
- [Sidner 79]
C. Sidner. Toward a computation of intrasentential coreference. Technical Report TR-537, MIT. Artificial Intelligence Laboratory, 1979.
- [Sidner 81]
C. Sidner. Focusing for interpretation of pronouns. *American Journal of Computational Linguistics*, 7:217–231, 1981.
- [Sidner 83]
C. Sidner. Focusing in the comprehension of definite anaphora. In B. M and B. R.C, editors, *Computational Models of Discourse*. Cambridge (MA) : The MIT Press, 1983.
- [Siekmann & Wrightson 80]
J. Siekmann and G. Wrightson. Paramodulated connection graphs. *Acta Informatica*, 13, 1980.
- [Siekmann 84]
J. Siekmann. Universal unification. In R. Shostak, editor, *Proceedings of the 7th International Conference on Automated Deduction*. Springer-Verlag, New York, 1984.
- [Siekmann 86]
J. Siekmann. Unification theory. *Proceedings of the 7th European Conference on Artificial Intelligence*, 1986.
- [Siersdorfer *et al.* 10]
S. Siersdorfer, J. Hare, E. Minack, and F. Deng. Analyzing and predicting sentiment of images on the social web. In *ACM Multimedia 2010*, pages 715–718. ACM, October 2010.
- [Sigurd 91]
B. Sigurd. Referent Grammar in Text Generation. In C. Paris, W. Swartout, and W. Mann, editors, *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, Dordrecht, Netherlands, 1991. Kluwer.

- [Sigurdsson *et al.* 08]
K. Sigurdsson, M. Stack, and I. Ranitovic. Heritrix user manual. Software documentation, Internet Archive, 2008.
- [Silva & Rocha 03]
N. Silva and J. Rocha. An Ontology MAPPING FRAMework for the Semantic Web. In *Proc. of the 6th International Conference on Business Information Systems*, Colorado Springs, USA, 2003.
- [Simard *et al.* 92]
M. Simard, G. Foster, and P. Isabelle. Using cognates to align sentences in bilingual corpora. In *Proc. of the Fourth International Conference on Theoretical and Methodological Issues in Machine Translation*, pages 67–81, Montreal, Canada, 1992.
- [Simkins 92]
N. K. Simkins. ALEP User Guide. CEC Luxemburg, 1992.
- [Simkins 94]
N. K. Simkins. An Open Architecture for Language Engineering. In *First CEC Language Engineering Convention*, Paris, 1994.
- [Simon 83]
H. Simon. Why should machines learn ? In J. Carbonell, R. Michalski, and T. Mitchell, editors, *Machine Learning: An Artificial Intelligence Approach*. Tioga Pub. Co., Palo Alto, CA, 1983.
- [Simon 95]
H. Simon. Artificial intelligence and empirical science. *Artificial Intelligence*, 77(1), August 1995.
- [Simon 06]
S. Simon. Stats: Steve’s attempt to teach statistics. Technical report, Children’s Mercy Hospitals & Clinics, Kansas City, Missouri, November 2006.
- [Simov & Osenova 03]
K. Simov and P. Osenova. Practical annotation scheme for an HPSG treebank of Bulgarian. In *Proceedings of the 4th International Workshop on Linguistically Interpreteted Corpora (LINC-2003)*, Budapest, Hungary, 2003.
- [Simov *et al.* 02]
K. Simov, G. Popova, and P. Osenova. HPSG-based syntactic treebank of Bulgarian (BulTreeBank). In Wilson *et al.* [Wilson *et al.* 02], pages 135–142.
- [Simov *et al.* 04a]
K. Simov, P. Osenova, A. Simov, and M. Kouylekov. Design and implementation of the Bulgarian HPSG-based treebank. *Journal of Research on Language and Computation*, 2(4):495–522, December 2004.
- [Simov *et al.* 04b]
K. Simov, P. Osenova, and M. Slavcheva. BulTreeBank morphosyntactic tagset. Technical Report BTB-TR03, BulTreeBank Project, March 2004.
- [Simperl *et al.* 10]
E. Simperl, I. Thurlow, P. Warren, F. Dengler, J. Davies, M. Grobelnik, D. Mladenić, J. M. Gomez-Perez, and C. R. Moreno. Overcoming information overload in the enterprise: The active approach. *IEEE Internet Computing*, 14:39–46, 2010.
- [Sinclair *et al.* 87]
G. Sinclair, P. Hanks, G. Fox, R. Moon, and P. e. a. Stock. *The Collins Cobuild English Language Dictionary*. Collins, Glasgow, 1987.

- [Singer & Lippmann 92]
E. Singer and R. P. Lippmann. A speech recogniser using radial basis function neural networks in an HMM framework. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 629–632, San Francisco CA, 1992.
- [Singh *et al.* 02]
P. Singh, T. Lin, E. Mueller, G. Lim, T. Perkins, and W. Li Zhu. Open mind common sense: Knowledge acquisition from the general public. In R. Meersman and Z. Tari, editors, *On the Move to Meaningful Internet Systems 2002: CoopIS, DOA, and ODBASE*, volume 2519 of *Lecture Notes in Computer Science*, pages 1223–1237. Springer Berlin / Heidelberg, 2002.
- [Singhal 12]
A. Singhal. Introducing the knowledge graph: things, not strings. <http://googleblog.blogspot.it/2012/05/introducing-knowledge-graph-things-not.html>, May 2012.
- [Siorpaes & Hepp 08]
K. Siorpaes and M. Hepp. Games with a purpose for the semantic web. *Intelligent Systems, IEEE*, 23(3):50–60, may-june 2008.
- [Sitbon & Bruza 08]
L. Sitbon and P. Bruza. On the relevance of documents for semantic representation. In *Proceedings of the 13th Australasian Document Computing Symposium*, Hobart, Australia, December 2008.
- [Sitter & Daelemans 03]
A. D. Sitter and W. Daelemans. Information extraction via double classification. In *Proceedings of ECML/PRDD 2003 Workshop on Adaptive Text Extraction and Mining (ATEM 2003)*, Cavtat-Dubrovnik, Croatia, 2003.
- [Sivashanmugam *et al.* 03]
K. Sivashanmugam, A. Sheth, J. Miller, K. Verma, R. Aggarwal, and P. Rajasekaran. Metadata and semantics for web services and processes. In W. Bennis, P. Dadam, S. Kirn, and R. Unland, editors, *Datenbanken und Informationssysteme (Databases and Information Systems): Festschrift zum 60. Geburtstag von Gunter Schlageter*, pages 245–271. Hagen, 2003.
- [Sizov *et al.* 03]
S. Sizov, J. Graupmann, and M. Theobald. From focused crawling to expert information: an application framework for web exploration and portal generation. In *Proceedings of the 29th International Conference on Very Large Data Bases*, pages 1105–1108, Berlin, 2003.
- [Skeels & Grudin 09]
M. M. Skeels and J. Grudin. When social networks cross boundaries: A case study of workplace use of Facebook and LinkedIn. In *Proceedings of the ACM 2009 international conference on Supporting group work*, GROUP '09, pages 95–104, New York, NY, USA, 2009. ACM.
- [Skuce 03a]
D. Skuce. A Controlled Language for Knowledge Formulation on the Semantic Web. <http://www.site.uottawa.ca:4321/factguru2.pdf>, 2003.
- [Skuce 03b]
D. Skuce. Knowledge formulation on the semantic web, 2003. Also available at <http://www.site.uottawa.ca:4321/factguru2.pdf>.
- [Slator *et al.* 86]
B. M. Slator, M. P. Anderson, and W. Conley. Pygmalion at the Interface. *Communications of the ACM*, 29(7):599–604, 1986.
- [Slocum 81]
J. Slocum. A practical comparison of parsing strategies. ACL, 1981.

- [Sloman 89]
A. Sloman. Must intelligent systems be scruffy? In *Evolving Knowledge in Natural Science and Artificial Intelligence*, Eds J.E.Tiles, G.T.McKee, G.C.Dean. Pitman, 1989.
- [Smadja 93]
F. Smadja. Retrieving collocations from text: Xtract. *Computational Linguistics*, 19(1):143–177, 1993.
- [Small & Rieger 82]
S. Small and C. Rieger. Parsing and comprehending with word experts (a theory and its realisation). In W. Lehnert and M. Ringle, editors, *Strategies for Natural Language Processing*. Lawrence Erlbaum Associates, Hillsdale, NJ, lawrence erlbaum associates edition, 1982.
- [Small *et al.* 88]
S. Small, G. Cottrell, and M. Tanenhaus, editors. *Lexical Ambiguity Resolution: Perspectives from Psycholinguistics, Neuropsychology and Artificial Intelligence*. Morgan Kaufmann, California, San Mateo, California, 1988.
- [Smeaton & Quigley 96]
A. Smeaton and I. Quigley. Experiments on using semantic distances between words in image caption retrieval. In *Proc. of 19th International Conference on Research and Development in Information Retrieval*, Zurich, Switzerland, 1996.
- [Smeaton *et al.* 04]
A. Smeaton, C. Gurrin, H. Lee, K. McDonald, N. Murphy, N. O’Connor, D. O’Sullivan, B. Smyth, and D. Wilson. The Físchlár-News-Stories System: Personalised Access to an Archive of TV News. In *Proceedings of RIAO 2004*, Avignon, France, 2004.
- [Smith & Osborne 06]
A. Smith and M. Osborne. Using gazetteers in discriminative information extraction. In *Proceedings of the Tenth Conference on Computational Natural Language Learning (CoNLL-X)*, pages 133–140. Association for Computational Linguistics, 2006.
- [Smith 82]
B. Smith. Linguistic and computational semantics. *Proceedings of the 20th Annual Meeting of the Association of Computational Linguistics*, 1982.
- [Smith *et al.* 02]
D. A. Smith, A. Mahoney, and G. Crane. Integrating harvesting into digital library content. In *Proceedings of the second ACM/IEEE-CS joint conference on Digital libraries*, pages 183–184. ACM Press, 2002.
- [Smolnik 03]
S. Smolnik. Knowledge discovery methods: An exploration of different approaches. In K. Freyberg, H.-J. Petsche, and B. Klein, editors, *Proceedings of the WM 2003 Workshop on Knowledge Management and Philosophy*, Luzern, Switzerland, Apr 2003.
- [Smullyan 68]
R. Smullyan. *First-Order Logic*. Springer-Verlag, Berlin, 1968.
- [Smyth *et al.* 95]
P. Smyth, U. Fayyad, M. Burl, P. Perona, and P. Baldi. Inferring ground truth from subjective labelling of venus images. *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS*, (7):1085–1092, 1995.
- [Snow *et al.* 08a]
R. Snow, B. O’Connor, D. Jurafsky, and A. Y. Ng. Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 254–263, 2008.

- [Snow *et al.* 08b]
R. Snow, B. O'Connor, D. Jurafsky, and A. Y. Ng. Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 254–263, 2008.
- [Soderland & Lehnert 94]
S. Soderland and W. Lehnert. Wrap-up: a trainable discourse model for information extraction. *Journal of Artificial Intelligence Research*, 2:131–158, 1994.
- [Soderland 97a]
S. Soderland. Learning text analysis rules for domain-specific natural language processing. *Ph.D. Thesis*, 1997.
- [Soderland 97b]
S. Soderland. Learning to extract text-based information from the world wide web. *Proceedings of Third International Conference on Knowledge Discovery and Data Mining (KDD-97)*, 1997.
- [Soderland 99a]
S. Soderland. Learning information extraction rules for semi-structured and free text. *Machine Learning*, 34(1):233–272, 1999.
- [Soderland 99b]
S. Soderland. Learning Information Extraction Rules for Semi-Structures and Free Text. *Machine Learning, Special Issue on Natural Language Learning*, 34(1-3), Feb. 1999.
- [Soderland *et al.* 95]
S. Soderland, D. Fisher, J. Aseltine, and W. Lehnert. CRYSTAL: Inducing a conceptual dictionary. In *Proc. of the 14th International Joint Conference on Artificial Intelligence*, 1995.
- [Soderland *et al.* 04]
S. Soderland, O. Etzioni, T. Shaked, and D. S. Weld. The Use of Web-based Statistics to Validate Information Extraction. In *Proceedings of AAAI 2004 Workshop on Adaptive Text Extraction and Mining (ATEM'04)*, 2004. <http://www.cs.washington.edu/research/knowitall/papers/SoderlandATEM04.pdf>.
- [Soderlang 99]
S. Soderlang. Learning information extraction rules for semi-structured and free texts. *Machine Learning*, 34:233–272, 1999.
- [Software 06]
L. Software. Getting started with bpm - white paper. Technical report, Lombardi, June 2006.
- [Sohn *et al.* 11]
S. Sohn, J.-P. A. Kocher, C. G. Chute, and G. K. Savova. Drug side effect extraction from clinical narratives of psychiatry and psychology patients. *Journal of the American Medical Informatics Association*, 18(Suppl 1):i144–i149, 2011.
- [Sokoll 06]
K. Sokoll. Optimizing drug development strategies. *Pharma and BioIngredients*, November 2006.
- [Solskinnsbakk & Gulla 11]
G. Solskinnsbakk and J. A. Gulla. Semantic annotation from social data. In *Proceedings of the Fourth International Workshop on Social Data on the Web Workshop*, 2011.
- [Somasundaran & Wiebe 10]
S. Somasundaran and J. Wiebe. Recognizing stances in ideological on-line debates. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 116–124. Association for Computational Linguistics, 2010.

- [Somers 92]
H. Somers. Example-based and corpus-based approaches to machine translation. In *Proc. of International Symposium on Natural Language Understanding and Artificial Intelligence*, pages 87–101, Lizukz, Japan, 1992.
- [Somers 96]
H. Somers, editor. *Terminology, LSP and Translation: Studies in Language Engineering in honour of Juan C. Sager*. John Benjamins, 1996.
- [Somers 97]
H. L. Somers. Machine Translation and Minority Languages. In *Translating and the Computer 19: Papers from the Aslib conference*, London, 1997.
- [Somers *et al.* 94]
H. Somers, I. McLean, and D. Jones. Experiments in multilingual example-based generation. In A. Monaghan, editor, *Proc. of 3rd Conference on the Cognitive Science of Natural Language Processing*, Dublin, Ireland, 1994.
- [Sominsky *et al.* 08]
I. Sominsky, A. Coden, and M. Tanenblatt. CFE – a system for testing, evaluation and machine learning of UIMA based applications. In *Language Resources and Evaluation Conference (LREC)*, 2008.
- [Sommerville 01]
I. Sommerville. *Software Engineering*. Addison-Wesley, Reading, MA, 2001.
- [Sonderland 99]
S. Sonderland. Learning Information Extraction Rules for Semi-structured and Free-text. *Machine Learning*, 34:1–44, 1999.
- [Song *et al.* 04]
Y. Song, E. Yi, E. Kim, and G. G. Lee. POSBIOTM-NER: A Machine Learning Approach for Bio-Named Entity Recognition. In *Workshop on a critical assessment of text mining methods in molecular biology*, Granada, Spain (http://www.pdg.cnb.uam.es/BioLINK/workshop_BioCreative_04/), 2004.
- [Song *et al.* 07]
X. Song, Y. Chi, K. Hino, and B. L. Tseng. Summarization system by identifying influential blogs. In *Proceedings of ICWSM*, pages 325–326, 2007.
- [Soon *et al.* 01]
W. M. Soon, H. T. Nh, and D. C. Y. Lim. Machine Learning Approach to Coreference Resolution of Noun Phrases. *Computational Linguistics*, 27(4):521–544, 2001.
- [Soong & Huang 91]
F. K. Soong and E.-F. Huang. A tree-trellis based fast search for finding the n-best sentence hypotheses in continuous speech recognition. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 1, pages 705–708, Toronto, 1991.
- [Sousa *et al.* 10]
D. Sousa, L. Sarmento, and E. Mendes Rodrigues. Characterization of the twitter @replies network: are user ties social or topical? In *Proceedings of the 2nd international workshop on Search and mining user-generated contents*, SMUC '10, pages 63–70, New York, NY, USA, 2010. ACM.
- [Souzis 05]
A. Souzis. Building a Semantic Wiki. *IEEE Intelligent Systems*, 20(5):87–91, 2005.
- [Sowa 83]
J. Sowa. *Conceptual Structures: Information Processing in Mind and Machine*. Systems Programming Series. Addison-Wesley, New York, 1983.

- [Sowa 84a]
J. Sowa. *Conceptual Structures: Information Processing in Mind and Machine*. Addison-Wesley, Reading, MA, 1984.
- [Sowa 84b]
J. Sowa. *Conceptual Structures : Information Processing in Mind and Machine*. Addison-Wesley, Reading, MA, 1984.
- [Sowa 91]
J. Sowa, editor. *Principles of Semantic Networks: Explorations in the Representation of Knowledge*. Morgan Kaufmann, California, San Mateo, CA, 1991.
- [Sowa 02]
J. Sowa. Architectures for intelligent systems. *IBM Systems Journal*, 41(3), 2002.
- [Sparck-Jones & Galliers 96]
K. Sparck-Jones and J. Galliers. *Evaluating Natural Language Processing Systems*. Springer, Berlin, 1996.
- [Sparck Jones & Tait 84]
K. Sparck Jones and J. Tait. Automatic search term variant generation. *Journal of Documentation*, 40(1):50–66, 1984.
- [Sparck-Jones & Wilks 85]
K. Sparck-Jones and Y. Wilks, editors. *Automatic Natural Language Parsing*. Ellis Horwood, 1985.
- [Sparck-Jones 68]
K. Sparck-Jones. *Synonymy and Semantic Classification*. Edinburgh UP, Edinburgh, 1968. Reprinted 1986.
- [Spärck-Jones 70]
K. Spärck-Jones. Some thoughts on classification for retrieval. *Journal of Documentation*, 26:89–101, 1970.
- [Sparck Jones 85]
K. Sparck Jones. Compound noun interpretation problems. In F. Fallside and W. Woods, editors, *Computer Speech Processing*, pages 363–381. Prentice Hall International, London, 1985.
- [Sparck Jones 93a]
K. Sparck Jones. Discourse Modelling for Automatic Summarising. Technical Report 290, University of Cambridge, Computer Laboratory, February 1993.
- [Sparck Jones 93b]
K. Sparck Jones. What Might Be in a Summary? In K. Knorz and Womser-Hacker, editors, *Information Retrieval 93: Von der Modellierung zur Anwendung*, 1993.
- [Sparck-Jones 94]
K. Sparck-Jones. Towards Better NLP System Evaluation. In *Proceedings of the Second ARPA Workshop on Human Language Technology*. Morgan Kaufmann, California, March 1994.
- [Spasic *et al.* 05]
I. Spasic, S. Ananiadou, J. McNaught, and A. Kumar. Text mining and ontologies in biomedicine: Making sense of raw text. *Briefings in Bioinformatics*, 6(3):239–251, 2005.
- [Spasić *et al.* 08]
I. Spasić, D. Schober, S.-A. Sansone, D. Rebholz-Schuhmann, D. B. Kell, and N. W. Paton. Facilitating the development of controlled vocabularies for metabolomics technologies with text mining. *BMC bioinformatics*, 9(5):S5, 2008.

- [Specia & Motta 07]
L. Specia and E. Motta. Integrating Folksonomies with the Semantic Web. In *ESWC*, Innsbruck, Austria, 2007.
- [Sperberg-McQueen & Burnard 94]
C. Sperberg-McQueen and L. Burnard. *Guidelines for Electronic Text Encoding and Interchange (TEI P3)*. ACH, ACL, ALLC, 1994. <http://etext.virginia.edu/TEI.html>.
- [Sperberg-McQueen & Burnard 95]
C. Sperberg-McQueen and L. Burnard. The design of the tei encoding scheme. *Computers and the Humanities*, 29:17–39, 1995. 10.1007/BF01830314.
- [Sperberg-McQueen & Burnard 02]
C. Sperberg-McQueen and L. Burnard, editors. *Guidelines for Electronic Text Encoding and Interchange (TEI P4)*. The TEI Consortium, 2002.
- [Spink *et al.* 01]
A. Spink, D. Wolfram, M. B. J. Jansen, and T. Saracevic. Searching the web: The public and their queries. *Journal of the American Society for Information Science and Technology*, 52(3):226–234, 2001.
- [Spivey 89]
J. Spivey. *The Z Notation: A Reference Manual*. Prentice Hall, New York, 1989.
- [Spivey 92]
J. Spivey. *The Z Notation: A Reference Manual*. Prentice Hall, New York, 2nd edition, 1992.
- [Spo01]
Sportec al minuto. <http://www.sportec.com>, May 2001.
- [Spreyer & Frank 08]
K. Spreyer and A. Frank. Projection-based acquisition of a temporal labeller. In *IJCNLP*, pages 489–496. ACL, 2008.
- [Sproat *et al.* 01]
R. Sproat, A. Black, S. Chen, S. Kumar, M. Ostendorf, and C. Richards. Normalization of non-standard words. *Computer Speech & Language*, 15(3):287–333, 2001.
- [Spyropoulos 99]
C. Spyropoulos. Final Report of the Greek Information Extraction (GIE) Project. Technical report, NKSR Demokritus, Athens, 1999.
- [Srihari *et al.* 00]
R. Srihari, C. Niu, and W. Li. A hybrid approach for named entity and sub-type tagging. In *Proceedings of the Sixth Applied Natural Language Processing Conference (ANLP'00)*, Seattle, Washington, 2000.
- [Srinivas 97]
B. Srinivas. *Complexity of Lexical Descriptions and its Relevance to Partial Parsing*. Unpublished PhD thesis, University of Pennsylvania, US, 1997.
- [Srinivasan & Libbus 04]
P. Srinivasan and B. Libbus. Mining medline for implicit links between dietary substances and diseases. *Bioinformatics*, 20(suppl 1):i290–i296, 2004.
- [Staab & Studer 04]
S. Staab and R. Studer, editors. *Handbook on Ontologies*. International Handbooks on Information Systems. Springer, 2004.

- [Staab 95]
S. Staab. GLR-parsing of word lattices using a beam search method. In *Proceedings European Conference on Speech Communication and Technology*, pages 1777–1780, Madrid, 1995.
- [Stairmand & Black 96]
M. A. Stairmand and W. J. Black. Conceptual and contextual indexing of documents using WordNet-derived lexical chains. In *Proc. of 18th BCS-IRSG Annual Colloquium on Information Retrieval Research*, 1996.
- [Stajner *et al.* 09]
T. Stajner, D. Rusu, B. Fortuna, D. Mladenović, and M. Grobelnik. Enrycher – service oriented text enrichment. *ktijssi*, 2009.
- [Starbird 17]
K. Starbird. Examining the alternative media ecosystem through the production of alternative narratives of mass shooting events on twitter. In *ICWSM*, pages 230–239. AAAI Press, 2017.
- [Stede & Huang 12]
M. Stede and C.-R. Huang. Inter-operability and reusability: the science of annotation. *Language Resources and Evaluation*, 46:91–94, 2012. 10.1007/s10579-011-9164-x.
- [Steele 84]
G. Steele. *Common Lisp: The Language*. Digital Equipment Press, 1984.
- [Stefanini & Deamzeau 95]
M. Stefanini and Y. Deamzeau. TALISMAN: a multi-agent system for Natural Language Processing. In *Proceedings of IEEE Conference on Advances in Artificial Intelligence, 12th Brazilian Symposium on AI*, 1995.
- [Steinberger *et al.* 11a]
J. Steinberger, M. Kabadjov, R. Steinberger, H. Tanev, M. Turchi, and V. Zavarella. Jrc’s participation at tac 2011: Guided and multilingual summarization tasks. In *Proceedings of the Text Analysis Conference (TAC) 2011*, 2011.
- [Steinberger *et al.* 11b]
R. Steinberger, B. Pouliquen, M. Kabadjov, J. Belyaeva, and E. van der Goot. JRC-Names: A freely available, highly multilingual named entity resource. In *Proceedings of the 8th International Conference in Recent Advances in Natural Language Processing*, pages 104–110, 2011.
- [Steinbiss *et al.* 93]
V. Steinbiss, H. Ney, R. Haeb-Umbach, B.-H. Tran, U. Essen, R. Kneser, M. Oerder, H.-G. Meier, X. Aubert, C. Dugast, and D. Geller. The Philips Research System for Large-Vocabulary Continuous-Speech Recognition. In *Proc. of Eurospeech ’93*, pages 2125–2128, Berlin, Germany, 1993.
- [Steinman & Carver 96]
S. Steinman and K. Carver. *Visual Programming With Prograph CPX*. Manning Publication, 1996.
- [Steinmetz *et al.* 09]
N. Steinmetz, A. Funk, and M. Maleshkova. Web service crawling and annotation (tutorial). In *Future Internet Symposium (FIS 2009)*, September 2009.
- [Stenmark 01]
D. Stenmark. The relationship between information and knowledge. In *Proceedings of 24th Information Systems Research Seminar in Scandinavia (IRIS)*, Norway, Aug 2001.
- [Stephens 07]
S. Stephens. Effective data integration is of increasing importance within the life sciences, February 2007.

- [Stern 97]
R. Stern. Specification of the 1996 Hub 4 Broadcast News Evaluation. In *Proc. 1997 DARPA Speech Recognition Workshop*, Chantilly, Virginia, 1997.
- [Stevenson & Ciravegna 03]
M. Stevenson and F. Ciravegna. Information Extraction as a Semantic Web Technology: Requirements and Promises. In *Proceedings of the Adaptive Text Extraction and Mining Workshop*, 2003.
- [Stevenson & Wilks 01]
M. Stevenson and Y. Wilks. The interaction of knowledge sources in word sense disambiguation. *Computational Linguistics*, 2001.
- [Stevenson 02]
M. Stevenson. Combining disambiguation techniques to enrich an ontology. In *Proceedings of ECAI-02 workshop "Machine Learning and Natural Language Processing for Ontology Engineering"*, Lyon, France, 2002.
- [Stevenson et al. 98]
M. Stevenson, H. Cunningham, and Y. Wilks. Sense tagging and language engineering. In *Proceedings of the 13th European Conference on Artificial Intelligence (ECAI-98)*, pages 185–189, Brighton, U.K., 1998.
- [Stewart 14]
I. Stewart. Now We Stronger Than Ever: African-American Syntax in Twitter. In *Proceedings of the EACL Student Research Workshop*, pages 31–37, 2014.
- [Stewart et al. 09]
R. Stewart, M. Soremekun, G. Perera, M. Broadbent, F. Callard, M. Denis, M. Hotopf, G. Thornicroft, and S. Lovestone. The South London and Maudsley NHS foundation trust biomedical research centre (SLAM BRC) case register: development and descriptive data. *BMC Psychiatry*, 9:51–62, 2009.
- [Stoilos et al. 10]
G. Stoilos, B. C. Grau, and I. Horrocks. How incomplete is your semantic web reasoner? In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI)*, 2010.
- [Stojanovic 05]
N. Stojanovic. On the Role of a User’s Knowledge Gap in an Information Retrieval Process. In *Proceedings of the Third International Conference on Knowledge Capture*, October 2005.
- [Stojanovic et al. 02]
L. Stojanovic, A. Maedche, B. Motik, and N. Stojanovic. User-driven ontology evolution management. In *13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02)*, pages 285–300, Siguenza, Spain, 2002.
- [Stoker 06]
G. Stoker. Explaining political disenchantment: finding pathways to democratic renewal. *The Political Quarterly*, 77(2):184–194, 2006.
- [Stolcke & Omohundro 93]
A. Stolcke and S. Omohundro. Hidden Markov model induction by Bayesian model merging. In S. J. Hanson, J. D. Cowan, and C. L. Giles, editors, *Advances in Neural Information Processing Systems*, volume 5. Morgan Kaufmann, California, 1993.
- [Stollberg et al. 04]
M. Stollberg, A. Zhdanova, and D. Fensel. h-TechSight – A Next Generation Knowledge Management Platform. *Journal of Information and Knowledge Management*, 3 (1):1–22, 2004.

- [Stone *et al.* 66]
P. J. Stone, D. C. Dunphy, and M. S. Smith. The general inquirer: A computer approach to content analysis. 1966.
- [Stone *et al.* 08]
B. Stone, S. Dennis, and P. J. Kwantes. A systematic comparison of semantic models on human similarity rating data: The effectiveness of subspacing. In *Proceedings of 30th Annual Conference of the Cognitive Science Society (CogSci08)*, pages 1813–1819, Washington, DC, USA, 23–26, July 2008.
- [Strohmaier *et al.* 10]
M. Strohmaier, C. Koerner, and R. Kern. Why do users tag? detecting users’ motivation for tagging in social tagging systems. 2010.
- [Strötgen & Gertz 10]
J. Strötgen and M. Gertz. HeidelTime: High quality rule-based extraction and normalization of temporal expressions. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pages 321–324. Association for Computational Linguistics, 2010.
- [Strötgen *et al.* 13]
J. Strötgen, J. Zell, and M. Gertz. HeidelTime: Tuning English and Developing Spanish Resources for TempEval-3. In *Proceedings of the 7th International Workshop on Semantic Evaluation (SemEval 2013)*, pages 15–19. ACL, 2013.
- [Strzalkowski & Vauthey 91a]
T. Strzalkowski and B. Vauthey. Fast text processing for information retrieval. In *4th DARPA Speech and Natural Language Workshop*. Pacific Grove, CA, February 1991.
- [Strzalkowski & Vauthey 91b]
T. Strzalkowski and B. Vauthey. Natural language processing in automated information retrieval. PROTEUS Project Memorandum 42, Department of Computer Science, New York University, October 1991.
- [Stuehrenberg *et al.* 07]
M. Stuehrenberg, D. Goecke, N. Diewald, A. Mehler, and I. Cramer. Web-based annotation of anaphoric relations and lexical chains. In *Proceedings of the Linguistic Annotation Workshop*, pages 140–147, June 2007.
- [Su & Gulla 03]
X. Su and J. A. Gulla. Semantic enrichment for ontology mapping. In *Proceedings of the 9th International Conference on Applications of Natural Language to Information Systems*, Manchester, UK, 2003.
- [Suchanek *et al.* 07]
F. M. Suchanek, G. Kasneci, and G. Weikum. Yago: a core of semantic knowledge. In *Proceedings of the 16th international conference on World Wide Web*, pages 697–706. ACM, 2007.
- [Suchanek *et al.* 09]
F. M. Suchanek, M. Sozio, and G. Weikum. Sofie: a self-organizing framework for information extraction. In *Proceedings of the 18th International Conference on World Wide Web*, pages 631–640, 2009.
- [Sugiyama *et al.* 81]
K. Sugiyama, S. Tagawa, and M. Toda. Methods for visual understanding of hierarchical system structures. *IEEE Transactions on Systems, Man & Cybernetics*, 11(2):109–125, 1981.
- [Sugumaran & Voss 12]
Sugumaran and Voss. Real-time spatio-temporal analysis of West Nile Virus using Twitter Data. In *Proceedings of Computing for Geospatial Research and Applications*, 2012.

- [Suh *et al.* 10]
 B. Suh, L. Hong, P. Pirollo, and E. H. Chi. Want to be retweeted? large scale analytics on factors impacting retweet in twitter network. In *Proceedings of the 2010 IEEE Second International Conference on Social Computing, SOCIALCOM '10*, pages 177–184, Washington, DC, USA, 2010. IEEE Computer Society.
- [Sumita & Iida 91]
 E. Sumita and H. Iida. Experiments and prospects of example-based machine translation. In *Proc. of 29th Annual Meeting of the Association for Computational Linguistics*, pages 185–192, Berkeley, California, 1991.
- [Summers 93]
 D. Summers. *Longman Language Activator*. Longman Group, 1993.
- [Sun & Zhu 13]
 J. Sun and Y. Zhu. Microblogging personalized recommendation based on ego networks. In *Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) - Volume 01, WI-IAT '13*, pages 165–170, Washington, DC, USA, 2013. IEEE Computer Society.
- [Sun *et al.* 02]
 J. Sun, J. Gao, L. Zhang, M. Zhou, and C. Huang. Chinese named entity identification using class-based language model. In *Proceedings of 19th International Conference on Computational Linguistics (COLING-02)*, Taipei, Taiwan, 2002.
- [Sun *et al.* 03]
 J. Sun, M. Zhou, and J. Gao. A class-based language model approach to chinese named entity identification. *Computational Linguistics and Chinese Language Processing*, 8(2), 2003.
- [Sun *et al.* 13]
 W. Sun, A. Rumshisky, and O. Uzuner. Evaluating temporal relations in clinical text: 2012 i2b2 challenge. *Journal of the American Medical Informatics Association*, 20(5):806–813, 2013.
- [Sundheim 95a]
 B. Sundheim, editor. *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, Columbia, MD, 1995. ARPA, Morgan Kaufmann.
- [Sundheim 95b]
 B. M. Sundheim. Overview of results of the Proceedings of the Sixth Message Understanding Conference (MUC-6) evaluation. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, pages 13–31, Columbia, MD, November 1995.
- [Sundheim 98]
 B. Sundheim, editor. *Proceedings of the Seventh Message Understanding Conference (MUC-7)*. ARPA, Morgan Kaufmann, 1998.
- [Sure 03]
 Y. Sure. *Methodology, Tools and Case Studies for Ontology based Knowledge Management*. Unpublished PhD thesis, University of Karlsruhe, May 2003.
- [Sure *et al.* 02]
 Y. Sure, M. Erdmann, J. Angele, S. Staab, R. Studer, and D. Wenke. OntoEdit: Collaborative Ontology Development for the Semantic Web. In *1st International Semantic Web Conference (ISWC2002)*, Sardinia, 2002. Springer.
- [Sussna 93]
 M. Sussna. Word sense disambiguation for free-text indexing using a massive semantic network. In *Proc. of 2nd International Conference on Information and Knowledge Management*, Arlington, Virginia, 1993.

- [Swerts *et al.* 00]
M. Swerts, D. Litman, and J. Hirschberg. Corrections in Spoken Dialogue Systems. In *Int. Conf. on Spoken Language Processing (ICSLP)*, volume 2, pages 615–618, Beijing, China, October 2000.
- [Syeda-Mahmood *et al.* 05]
T. Syeda-Mahmood, G. Shah, R. Akkiraju, A. Ivan, and R. Goodwin. Searching service repositories by combining semantic and ontological matching. In *Proceedings of the IEEE International Conference on Web Services (ICWS 2005)*, pages 13–20, 2005.
- [Szomszor *et al.* 08]
M. Szomszor, H. Alani, I. Cantador, K. O’Hara, and N. Shadbolt. Semantic modelling of user interests based on cross-folksonomy analysis. In *Proceedings of the 7th International Conference on The Semantic Web (ISWC)*, pages 632–648. Springer-Verlag, 2008.
- [Tablan & Cunningham 08]
V. Tablan and H. Cunningham. Mixed-Mode Text Indexing in SAM: a Position Paper. Technical report, University of Sheffield, Sheffield, UK, 2008.
- [Tablan *et al.* 02a]
V. Tablan, C. Ursu, K. Bontcheva, H. Cunningham, D. Maynard, O. Hamza, T. McEnery, P. Baker, and M. Leisher. A Unicode-based Environment for Creation and Use of Language Resources. In *3rd Language Resources and Evaluation Conference*, Las Palmas, Canary Islands – Spain, 2002. ELRA. <http://gate.ac.uk/sale/ies103/ies103.pdf>.
- [Tablan *et al.* 02b]
V. Tablan, C. Ursu, K. Bontcheva, H. Cunningham, D. Maynard, O. Hamza, T. McEnery, P. Baker, and M. Leisher. A Unicode-based Environment for Creation and Use of Language Resources. In *Proceedings of the 3rd Language Resources and Evaluation Conference*, 2002.
- [Tablan *et al.* 03]
V. Tablan, K. Bontcheva, D. Maynard, and H. Cunningham. Ollie: on-line learning for information extraction. In *SEALTS ’03: Proceedings of the HLT-NAACL 2003 workshop on Software engineering and architecture of language technology systems*, volume 8, pages 17–24, Morristown, NJ, USA, 2003. Association for Computational Linguistics. <http://gate.ac.uk/sale/hlt03/ollie-sealts.pdf>.
- [Tablan *et al.* 06a]
V. Tablan, W. Peters, D. Maynard, H. Cunningham, and K. Bontcheva. Creating tools for morphological analysis of sumerian. In *5th Language Resources and Evaluation Conference (LREC)*, Genoa, Italy, May 2006. ELRA.
- [Tablan *et al.* 06b]
V. Tablan, T. Polajnar, H. Cunningham, and K. Bontcheva. User-friendly Ontology Authoring Using a Controlled Language. In *5th Language Resources and Evaluation Conference (LREC)*, Genoa, Italy, May 2006. ELRA.
- [Tablan *et al.* 08]
V. Tablan, D. Damljanovic, and K. Bontcheva. A Natural Language Query Interface to Structured Information. In *Proceedings of the 5th European Semantic Web Conference (ESWC, 1–5 June 2008)*, volume 5021 of *Lecture Notes in Computer Science*, pages 361–375, Tenerife, Spain, 1–5 June 2008. Springer-Verlag New York Inc.
- [Tablan *et al.* 11]
V. Tablan, I. Roberts, H. Cunningham, and K. Bontcheva. GATECloud.net: Cloud Infrastructure for Large-Scale, Open-Source Text Processing. In *UK e-Science All hands Meeting 2011*, York, UK, September 2011.

- [Tablan *et al.* 13a]
V. Tablan, I. Roberts, H. Cunningham, and K. Bontcheva. Gatecloud.net: a platform for large-scale, open-source text processing on the cloud. *Philosophical Transactions of the Royal Society A*, 371(1983), 2013.
- [Tablan *et al.* 13b]
V. Tablan, I. Roberts, H. Cunningham, and K. Bontcheva. GATECloud.net: a Platform for Large-Scale, Open-Source Text Processing on the Cloud. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, 371(1983):20120071, 2013.
- [Tablan *et al.* 15]
V. Tablan, K. Bontcheva, I. Roberts, and H. Cunningham. Mimir: an open-source semantic search framework for interactive information seeking and discovery. *Journal of Web Semantics*, 30:52–68, 2015.
- [Taboada *et al.* 11]
M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede. Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 1(September 2010):1–41, 2011.
- [Tait 82]
J. I. Tait. *Automatic Summarising of English Texts*. Unpublished PhD thesis, University of Cambridge, Computer Laboratory, December 1982.
- [Tait 08]
J. Tait. Information retrieval facility symposium in vienna. *SIGIR Forum*, 42:67–67, June 2008.
- [Takemoto *et al.* 96]
Y. Takemoto, T. Wakao, H. Yamada, R. Gaizauskas, and Y. Wilks. Description of the NEC/Sheffield System Used for MET Japanese. In *Proceedings of the TIPSTER Phase II Workshop*, 1996.
- [Tan & Kumar 00]
P. Tan and V. Kumar. Interestingness measures for association patterns: A perspective. Technical Report TR00-036, Department of Computer Science, University of Minnesota, 2000.
- [Tan *et al.* 10]
W. Tan, R. Madduri, A. Nenadic, S. Soiland-Reyes, D. Sulakhe, I. Foster, and C. Goble. Cagrid workflow toolkit: A taverna based workflow tool for cancer grid. *BMC Bioinformatics*, 11(1):542, 2010.
- [Tanabe & Wilbur 02]
L. Tanabe and W. J. Wilbur. Tagging Gene and Protein Names in Full Text Articles. In *Proceedings of the ACL-02 workshop on Natural Language Processing in the biomedical domain, 7–12 July 2002*, volume 3, pages 9–13, Philadelphia, PA, 2002. Association for Computational Linguistics.
- [Tannenbaum 89]
A. Tannenbaum. *Computer Networks, 2nd Edn.* Prentice-Hall, London, 1989.
- [Tao *et al.* 12]
K. Tao, F. Abel, C. Hauff, and G.-J. Houben. What makes a tweet relevant to a topic. In *Proceedings of the #MSM2012 Workshop, CEUR*, volume 838, 2012.
- [Tapanainen & Jarvinen 94]
P. Tapanainen and T. Jarvinen. Syntactic analysis of natural language using linguistic rules and corpus-based patterns. In *Proc. of 15th International Conference on Computational Linguistics (COLING)*, volume 1, pages 629–634, Kyoto, Japan, 1994.
- [Tarnlund 86]
S.-A. Tarnlund. Logic programming - from a logic point of view. *Proceedings of the 1986 Symposium on Logic Programming*, 1986.

- [Tashiro *et al.* 05]
 T. Tashiro, M. Rikitoku, and T. Nakagawa. Justsystem at NTCIR-5 Patent Classification. In *Proceedings of the Fifth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access*, Tokyo, Japan, 2005.
- [Taskar *et al.* 03]
 B. Taskar, C. Guestrin, and D. Koller. Max-Margin Markov Networks. In *Neural Information Processing Systems Conference (NIPS03)*, Vancouver, Canada, 2003.
- [Tattersall *et al.* 90]
 G. Tattersall, P. Linford, and R. Lingard. Neural Arrays for Speech Recognition. In *Speech and Language Processing*. Chapman and Hall, London, 1990.
- [Tazzoli *et al.* 04]
 R. Tazzoli, P. Castagna, and S. Campanini. Towards a Semantic Wiki Wiki Web. In *Poster Track, 3rd International Semantic Web Conference (ISWC2004)*, Hiorshima, Japan, 2004.
- [Team 11]
 T. F. D. Team. Anatomy of Facebook. *Facebook*, 2011. <https://www.facebook.com/notes/facebook-data-team/anatomy-of-facebook/10150388519243859>.
- [Teevan *et al.* 11]
 J. Teevan, D. Ramage, and M. R. Morris. #twittersearch: a comparison of microblog search and web search. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining (WSDM)*, pages 35–44, 2011.
- [Tello & Gómez-Pérez 04]
 A. L. Tello and A. Gómez-Pérez. Ontometric: A method to choose the appropriate ontology. *J. Database Manag.*, 15(2):1–18, 2004.
- [Tennant 81]
 H. Tennant. *Natural Language Processing*. Petrocelli Books, New York, 1981.
- [Tennant *et al.* 83a]
 H. Tennant, R. K.M., and T. C.W. Usable natural language interfaces through menu-based understanding. In *Proceedings of CHI 83*. CHI, 1983.
- [Tennant *et al.* 83b]
 H. Tennant, R. K.M., R. Saenz, T. C.W., and M. J.R. Menu-based natural language understanding. In *21st Annual Meeting of the Association for Computational linguistics proceedings*. ACL, 1983.
- [Terry *et al.* 88]
 M. Terry, S. Renals, R. Rohwer, and J. Harrington. A connectionist approach to speech recognition using peripheral auditory modelling. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 699–702, New York, 1988.
- [Tersmette *et al.* 88]
 K. Tersmette, A. Scott, G. Moore, N. Matheson, and R. Miller. Barrier Word Method for Detecting Molecular Biology Multiple Word Terms. In R. Greenes, editor, *Proc. of the 12th Annual Symposium on Computer Applications in Medical Care (SCAMC)*, pages 207–211, New York, 1988. IEEE.
- [Terziev *et al.* 05]
 I. Terziev, A. Kiryakov, and D. Manov. Base upper-level ontology (BULO) guidance. Deliverable D1.8.1, SEKT Consortium, July 2005.
- [Tessiore & v.Hahn 00]
 L. Tessiore and W. v. Hahn. *Functional validation of a machine translation system: Verbmobil*, pages 611–631. In Wahlster [Wahlster 00], July 2000.

- [Teufel & Moens 99]
S. Teufel and M. Moens. Argumentative classification of extracted sentences as a first step towards flexible abstracting. In I. Mani and M. Maybury, editors, *Advances in Automatic Text Summarization*, pages 155–171. The MIT Press, 1999.
- [Teufel 98]
S. Teufel. Meta-Discourse Markers and Problem-Structuring in Scientific Texts. In M. Stede, L. Wanner, and E. Hovy, editors, *Proceedings of the Workshop on Discourse Relations and Discourse Markers, COLING-ACL'98*, pages 43–49, 15th August 1998.
- [Thaddeus & Raja 06]
S. Thaddeus and S. K. Raja. A Semantic Web Tool for Knowledge-based Software Engineering. In *Workshop on Semantic Web Enabled Software Engineering (SWESE)*, Athens, G.A., USA, 2006.
- [Thaler *et al.* 11a]
S. Thaler, K. S. E. Simperl, and C. Hofer. A survey on games for knowledge acquisition. Technical Report Tech. Rep. STI TR 2011-05-01, Semantic Technology Institute, 2011.
- [Thaler *et al.* 11b]
S. Thaler, K. Siorpaes, C. Hofer, and E. Simperl. A Survey on Games for Knowledge Acquisition. Technical report, Semantic Technology Institute, Innsbruck, 2011.
- [The Object Management Group 92]
The Object Management Group. *The Common Object Request Broker: Architecture and Specification*. John Wiley, New York, 1992.
- [The Wellcome Trust Case Control Consortium 07]
The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 447:661–678, June 2007.
- [Thelwall *et al.* 10]
M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558, 2010.
- [Thelwall *et al.* 12]
M. Thelwall, K. Buckley, and G. Paltoglou. Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, 63(1):163–173, 2012.
- [Thelwall *et al.* 13]
M. Thelwall, S. Haustein, V. Larivière, and C. R. Sugimoto. Do altmetrics work? twitter and ten other social web services. *PloS one*, 8(5):e64841, 2013.
- [Thompson & McKelvie 96]
H. Thompson and D. McKelvie. A Software Architecture for Simple, Efficient SGML Applications. In *Proceedings of SGML Europe '96*, Munich, 1996.
- [Thompson & McKelvie 97]
H. Thompson and D. McKelvie. Hyperlink semantics for standoff markup of read-only documents. In *Proceedings of SGML Europe '97*, Barcelona, 1997.
- [Thompson 85]
H. Thompson. Natural language processing: a critical analysis of the structure of the field, with some implications for parsing. In K. Sparck-Jones and Y. Wilks, editors, *Automatic Natural Language Parsing*. Ellis Horwood, Chichester, 1985.
- [Thompson 95]
H. Thompson. MULTEXT Workpackage 2 Milestone B Deliverable Overview. LRE 62-050 Deliverable 2, 1995.

- [Thompson *et al.* 99]
 C. A. Thompson, M. E. Califf, and R. J. Mooney. Active learning for natural language parsing and information extraction. In *Proceedings of the International Conference on Machine Learning*, pages 406–414, 1999.
- [Thurmair 96a]
 G. Thurmair. AVENTINUS System Architecture. LE project LE1-2238 AVENTINUS internal technical report, GMS, Germany, 1996.
- [Thurmair 96b]
 G. Thurmair. Multilingual information processing in aventinus. Submitted 1996, 1996.
- [Thurmair 96c]
 G. Thurmair. TERRIOR: Terminology in AVENTINUS. LE project LE1-2238 AVENTINUS internal technical report, GMS, Germany, 1996.
- [Thurmair 96d]
 G. Thurmair. WP 4.1 Task S6: Text Handling, Detailed Functional Specifications. Technical Report WP41-S6-V1.1, Sail Labs GMBH., Munich, 1996. LE project LE1-2238 AVENTINUS internal report.
- [Thurmair 97]
 G. Thurmair. Information extraction for intelligence systems. In *Natural Language Processing: Extracting Information for Business Needs*, pages 135–149, London, March 1997. Unicom Seminars Ltd.
- [Tiedemann *et al.* 14]
 J. Tiedemann, Z. Agic, and J. Nivre. Treebank translation for cross-lingual parser induction. In *CoNLL*, pages 130–140, 2014.
- [Tillmann & Ney 02]
 C. Tillmann and H. Ney. Word re-ordering and a dp beam search algorithm for statistical machine translation. *Computational Linguistics*, page in press, 2002.
- [Tillmann *et al.* 00]
 C. Tillmann, S. Vogel, H. Ney, and H. Sawaf. Statistical translation of text and speech: First results with the rwth system. *Machine Translation*, 15:43–74, 2000.
- [TIPSTER 95]
 TIPSTER. The Generic Document Detection System. http://www.itl.nist.gov/div894/894.02/-related_projects/tipster/gen_ir.htm, 1995.
- [TIPSTER Architecture Committee 94]
 TIPSTER Architecture Committee. TIPSTER Text Phase II Architecture Concept. Technical report, New York University, 1994. <http://www.cs.nyu.edu/tipster>.
- [Tissot *et al.* 15a]
 H. Tissot, A. Roberts, L. Derczynski, G. Gorrell, and M. Didonet Del Fabro. Analysis of temporal expressions annotated in clinical notes. In *Proceedings of 11th Joint ACL-ISO Workshop on Interoperable Semantic Annotation*, pages 93–102, London, UK, 2015. ACL.
- [Tissot *et al.* 15b]
 H. Tissot, A. Roberts, G. Gorrell, L. Derczynski, and M. Didonet Del Fabro. UFPRSheffield: Contrasting Two Diverse Approaches to Time Expression Identification in Clinical TempEval. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, Denver, USA, 2015. ACL.
- [Titov & Mcdonald 08]
 I. Titov and R. Mcdonald. A joint model of text and aspect ratings for sentiment summarization. In *Proc. ACL-08: HLT*, pages 308–316, 2008.

- [Tjong Kim Sang & Buchholz 00a]
 E. Tjong Kim Sang and S. Buchholz. Introduction to the CoNLL-2000 shared task: Chunking. In *Proceedings of CoNLL-2000 and LLL-2000*, pages 127–132, Lisbon, Portugal, September 2000.
- [Tjong Kim Sang & Buchholz 00b]
 E. F. Tjong Kim Sang and S. Buchholz. Introduction to the CoNLL-2000 shared task: chunking. In *Proceedings of the 2nd workshop on Learning language in logic and the 4th conference on Computational natural language learning - Volume 7, ConLL '00*, pages 127–132, Stroudsburg, PA, USA, 2000. Association for Computational Linguistics.
- [Tjong Kim Sang & Meulder 03]
 E. F. Tjong Kim Sang and F. D. Meulder. Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition. In *Proceedings of CoNLL-2003*, pages 142–147. Edmonton, Canada, 2003.
- [Todorascu *et al.* 02]
 A. Todorascu, E. Kow, and L. Romary. Towards reusable nlp components. In *Proceedings of 3rd Language Resources and Evaluation Conference (LREC'2002)*, Gran Canaria, Spain, 2002.
- [Toffler 84]
 A. Toffler. *Future Shock*. Random House Publishing Group, 1984.
- [Tomanek 07]
 W. J. H. U. Tomanek, K. Efficient annotation with the jena annotation environment (jane). In *Proceedings of the Linguistic Annotation Workshop*, pages 9–16. Association for Computational Linguistics, June 2007.
- [Tombros *et al.* 98]
 A. Tombros, M. Sanderson, and P. Gray. Advantages of Query Biased Summaries in Information Retrieval. In *Intelligent Text Summarization. Papers from the 1998 AAAI Spring Symposium. Technical Report SS-98-06*, pages 34–43, Stanford (CA), USA, March 23-25 1998. The AAAI Press.
- [Tomita 86]
 M. Tomita. An efficient word lattice parsing algorithm for continuous speech recognition. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1569–1572, Tokyo, 1986.
- [Tomlinson 03]
 S. Tomlinson. Lexical and algorithmic stemming compared for 9 European languages with Hummingbird SearchServer(TM) at CLEF 2003. In C. Peters, editor, *Working notes for the CLEF 2003 Workshop*, pages 22–29, Trondheim, Norway, 2003.
- [Tong & Chang 01]
 S. Tong and E. Chang. Support vector machine active learning for image retrieval. In *Proc. of ACM Int. Conf. on Multimedia*, pages 107–118, 2001.
- [Tong & Koller 01]
 S. Tong and D. Koller. Support vector machine active learning with applications to text classification. *Journal of Machine Learning Research*, 2:45–66, 2001.
- [Tong 95]
 R. Tong. Interactive document retrieval using topic. In D. Harman, editor, *Overview of the Third retrieval Conference (TREC-3)*., 1995.
- [Torisawa & Yonezawa 94]
 K. Torisawa and A. Yonezawa. Lexicon acquisition in HPSG/JPSG-based grammars. Technical report, University of Tokyo, 1994.

- [Touretzky 86]
D. Touretzky. *The Mathematics of Inheritance Systems*. Morgan Kaufmann, California, Los Altos, Ca., 1986.
- [Toutanova & Manning 00]
K. Toutanova and C. D. Manning. Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In *Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora (EMNLP/VLC-2000)*, pages 63–70, Hong Kong, October 2000.
- [Toutanova et al. 03]
K. Toutanova, D. Klein, C. D. Manning, and Y. Singer. Feature-rich part-of-speech tagging with a cyclic dependency network. In *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, NAACL '03*, pages 173–180, 2003.
- [Tracz 95]
W. Tracz. Domain-Specific Software Architecture (DSSA) Frequently Asked Questions (FAQ). <http://www.oswego.com/dssa/faq/faq.html>, March 1995.
- [Tran et al. 11]
T. Tran, P. Mika, H. Wang, and M. Grobelnik. Semsearch'11: The 4th semantic search workshop. In *Proceedings of the 20th International Conference Companion on World Wide Web, WWW '11*, pages 315–316. ACM, 2011.
- [Trautteur]
G. Trautteur. *A Novel Interaction Between AI and Philosophy*.
- [Travis 11]
A. Travis. England riots: will harsher sentences act as a deterrent? *The Guardian*, 2011. <http://www.guardian.co.uk/uk/2011/aug/17/england-riots-harsher-sentences-deterrent>.
- [TRESTLE 00]
TRESTLE. The TRESTLE project. <http://www.dcs.shef.ac.uk/research/groups/nlp/trestle>, 2000.
- [Trujillo & Plowman 93]
A. Trujillo and D. Plowman. Automation of bilingual lexicon compilation. In Nirenburg, editor, *Progress in Machine Translation*. ISO Publishers, Washington, 1993.
- [Tsirliganis et al. 01]
N. Tsirliganis, G. Pavlidis, A. Tsompanopoulos, D. Papadopoulou, Z. Loukou, E. Politou, K. Stavroglou, and C. Chamzas. Integrated Documentation of Cultural Heritage through 3D imaging and multimedia database. In *Proceedings of VAST2001: Virtual Reality, Archaeology, and Cultural Heritage*, Glyfada, Athens, Greece, November 2001.
- [Tsirliganis et al. 02]
N. Tsirliganis, G. Pavlidis, A. Koutsoudis, E. Politou, A. Tsompanopoulos, K. Stavroglou, and C. Chamzas. New Ways in Digitization and Visualization of Cultural Objects. In *Proceedings of Proc. IEEE DSP 2002*, Santorini, Greece, July 2002.
- [Tsochantaridis et al. 04]
I. Tsochantaridis, T. Hofmann, T. Joachims, and Y. Altun. Support Vector Machine Learning for Interdependent and Structured Output Spaces. In *Proceedings of the 21 st International Conference on Machine Learning, Banff, Canada, 2004*.
- [Tsur et al. 10]
O. Tsur, D. Davidov, and A. Rappoport. Icwsn—a great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, pages 162–169, 2010.

- [Tsuruoka *et al.* 05]
 Y. Tsuruoka, Y. Tateishi, J.-D. Kim, T. Ohta, J. McNaught, S. Ananiadou, and J. Tsujii. Developing a robust part-of-speech tagger for biomedical text. In P. Bozanis and E. Houstis, editors, *Advances in Informatics: Proceedings of the 10th Panhellenic Conference on Informatics (PCI 2005), 11–13 November 2005*, volume 3746 of *Lecture Notes in Computer Science*, pages 382–392, Volas, Greece, 2005. Springer Berlin Heidelberg.
- [Tsutsumi 93]
 T. Tsutsumi. Natural language processing: the PLNLP approach. In K. Jenon, G. Heidhorn, and S. Richardson, editors, *Word Sense Disambiguation by Examples*, pages 263–272. Kluwer Academic Publishers, Dordrecht, 1993.
- [Tufekci 14]
 Z. Tufekci. Big questions for social media big data: Representativeness, validity and other methodological pitfalls. In *Proceedings of the 8th International AAAI Conference on Weblogs and Social Media*, 2014.
- [Tullis & Stetson 04]
 T. S. Tullis and J. N. Stetson. A Comparison of Questionnaires for Assessing Website Usability. In *Usability Professionals’ Association Conference*, Minneapolis, Minnesota, June 2004.
- [Tumasjan *et al.* 10]
 A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welp. Predicting elections with Twitter: What 140 characters reveal about political sentiment. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, pages 178–185, 2010.
- [Tummarello *et al.* 07]
 G. Tummarello, R. Delbru, and E. Oren. Sindice.com: Weaving the Open Linked Data. In *Proceedings of the 6th International Semantic Web Conference*, Busan, Korea, 2007.
- [Tur *et al.* 03]
 G. Tur, R. Schapire, and D. Hakkani-Tur. Active Learning for Spoken Language Understanding. In *Proceedings of 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 276–279, 2003.
- [Turian *et al.* 09]
 J. Turian, L. Ratinov, Y. Bengio, and D. Roth. A preliminary evaluation of word representations for named-entity recognition. In *NIPS Workshop on Grammar Induction, Representation of Language and Language Learning*, pages 1–8, 2009.
- [Turian *et al.* 10]
 J. Turian, L. Ratinov, and Y. Bengio. Word representations: a simple and general method for semi-supervised learning. In *Proceedings of the 48th annual meeting of the association for computational linguistics*, pages 384–394. Association for Computational Linguistics, 2010.
- [Turmo *et al.* 06]
 J. Turmo, A. Ageno, and N. Català. Adaptive information extraction. *ACM Computing Surveys (CSUR)*, 38(2):4, 2006.
- [Turner & Charniak 05]
 J. Turner and E. Charniak. Supervised and unsupervised learning for sentence compression. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, ACL ’05, pages 290–297, Stroudsburg, PA, USA, 2005. Association for Computational Linguistics.
- [Turner 84]
 R. Turner. *Logics for AI*. Ellis-Horwood, Chichester, 1984.

- [Turney 97]
P. Turney. Extraction of keyphrases from text: Evaluation of four algorithms. http://ai.iit.nrc.ca/II_public/extractor/reports/index.html, 1997.
- [Turney 02]
P. D. Turney. Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL '02)*, pages 417–424, Morristown, NJ, USA, July 2002.
- [Ueberla & Gransden 96]
J. Ueberla and I. Gransden. Clustered language models with context-dependent states. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 4, Philadelphia PA, 1996.
- [Ueberla 96]
J. Ueberla. An extended clustering algorithm for statistical language models. *IEEE Transactions on Signal Processing*, 4:313–316, 1996.
- [Ullman 82]
J. Ullman. *Principles of Database Systems*. Computer Science Press, 1982.
- [UML]
<http://www.nlm.nih.gov/pubs/factsheets/umls.html>.
- [Unicode Consortium 96]
Unicode Consortium. *The Unicode Standard, Version 2.0*. Addison-Wesley, Reading, MA, 1996.
- [University of Essex 99]
University of Essex. Description of the W3-Corpora web-site. <http://c1www.essex.ac.uk/w3c/>, 1999.
- [Uramoto 94]
N. Uramoto. A best-match algorithm for broad-coverage example-based disambiguation. In *Proc. of 15th International Conference on Computational Linguistics*, volume 2, pages 717–721, Kyoto, Japan, 1994.
- [Urbain & Frieder 10]
J. Urbain and O. Frieder. Exploring contextual models for chemical patent search. In H. Cunningham, A. Hanbury, and S. Rüger, editors, *Advances in Multidisciplinary Retrieval (the 1st Information Retrieval Facility Conference)*. LNCS volume number: 6107, Lecture Notes in Computer Science, Vienna, Austria, May 2010. Springer.
- [Uren *et al.* 06]
V. Uren, P. Cimiano, J. Iria, S. Handschuh, M. Vargas-Vera, E. Motta, and F. Ciravegna. Semantic annotation for knowledge management: Requirements and a survey of the state of the art. *Web Semantics: Science, Services and Agents on the World Wide Web*, 4(1):14–28, January 2006.
- [Ursu *et al.* 05]
C. Ursu, T. Tablan, H. Cunningham, and B. Popav. Digital media preservation and access through semantically enhanced web-annotation. In *Proceedings of the 2nd European Workshop on the Integration of Knowledge, Semantic and Digital Media Technologies (EWIMT 2005)*, London, UK, December 01 2005.
- [Usbeck *et al.* 14]
R. Usbeck, A.-C. Ngonga Ngomo, S. Auer, D. Gerber, and A. Both. Agdistis - graph-based disambiguation of named entities using linked data. In *International Semantic Web Conference*. 2014.
- [Ushioda 96]
A. Ushioda. Hierarchical clustering of words. In *Proc. of 16th International Conference on Computational Linguistics (COLING)*, pages 1159–1162, 1996.

- [Utiyama *et al.* 00]
M. Utiyama, M. Murata, and H. Isahara. Statistical approach to the processing of metonymy. 2000.
- [Uysal & Croft 11a]
I. Uysal and W. B. Croft. User oriented tweet ranking: a filtering approach to microblogs. In *Proceedings of the 20th ACM Conference on Information and Knowledge Management, CIKM 2011, Glasgow, United Kingdom, October 24-28, 2011*, pages 2261–2264, 2011.
- [Uysal & Croft 11b]
I. Uysal and W. B. Croft. User oriented tweet ranking: a filtering approach to microblogs. In *CIKM*, pages 2261–2264, 2011.
- [Uzdilli *et al.* 15]
F. Uzdilli, M. Jaggi, D. Egger, P. Julmy, L. Derczynski, and M. Cieliebak. Swiss-chocolate: Combining flipout regularization and random forest with artificially built subsystems to boost text-classification for sentiment. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*. ACL, 2015.
- [UzZaman & Allen 10]
N. UzZaman and J. Allen. TRIPS and TRIOS system for TempEval-2: Extracting temporal information from text. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pages 276–283. Association for Computational Linguistics, 2010.
- [UzZaman *et al.* 13]
N. UzZaman, H. Llorens, L. Derczynski, M. Verhagen, J. F. Allen, and J. Pustejovsky. SemEval-2013 Task 1: TempEval-3: Evaluating Time Expressions, Events, and Temporal Relations. In *Proceedings of the 7th International Workshop on Semantic Evaluations*, 2013.
- [Valarakos *et al.* 04]
A. G. Valarakos, G. Paliouras, V. Karkaletsis, and G. Vouros. Enhancing ontological knowledge through ontology population and enrichment. In *Engineering Knowledge in the Age of the Semantic Web*, pages 144–156, 2004.
- [Valderrabanos *et al.* 02]
A. Valderrabanos, A. Belskis, and L. Iraola. Textractor: a multilingual terminology extraction tool. In *Proceedings of Human Language Technologies Conference (HLT)*, San Diego, 2002.
- [Valiant 84]
L. Valiant. A theory of the learnable. *Communications of the ACM*, 27(11), 1984.
- [Van Atteveldt *et al.* 07]
W. Van Atteveldt, S. Schlobach, and F. Van Harmelen. Media, politics and the semantic web. In *European Semantic Web Conference*, pages 205–219. Springer, 2007.
- [van der Linden 94]
P. van der Linden. *Expert C Programming – Deep C Secrets*. Prentice Hall, New York, 1994.
- [Van derAalst 03]
T. H. A. K. B. B. A. Van der Aalst, W.M.P. Workflow patterns. *Distributed and Parallel Databases*, 14:5–51, 2003.
- [van derEijk 93]
P. van der Eijk. Automating the acquisition of bilingual terminology. In *Proc. of 6th Conference of the European Chapter of the Association for Computational Linguistics, EACL '93*, pages 113–119, 1993.
- [Van deVelde 12]
E. Van de Velde. *Feasibility study for an EU monitoring mechanism on key enabling technologies*. IDEA Consult, 2012.

- [Van Erp *et al.* 13]
M. Van Erp, A. Fokkens, P. Vossen, S. Tonelli, W. R. van Hage, L. Serafini, R. Sprugnoli, and J. Hoeksema. Denoting data in the Grounded Annotation Framework. In *International Semantic Web Conference (Posters & Demos)*, pages 189–192, 2013.
- [van Rijsbergen 79]
C. van Rijsbergen. *Information Retrieval*. Butterworths, London, 1979.
- [vanDen Bosch & Bogers 13]
A. van Den Bosch and T. Bogers. Memory-based named entity recognition in tweets. In *Making Sense of Microposts (#MSM2013) Concept Extraction Challenge*, pages 40–43, 2013.
- [Vanderheyden & Cohen 98]
P. Vanderheyden and R. Cohen. Information extraction and the casual user. *AAAI-98 Workshop on AI and Information Integration*, pages 137–142, 1998.
- [Vanderwende 95]
L. Vanderwende. Ambiguity in the acquisition of lexical information. In *Proc. of the AAAI 1995 Spring Symposium Series: Working notes of the symposium on representation and acquisition of lexical knowledge*, pages 174–179, 1995.
- [vanEmdeBoas & vanEmdeBoas 86]
G. vanEmdeBoas and P. vanEmdeBoas. Storing and evaluating horn-clause rules in a relational database. *IBM Journal of Research and Development*, 30(1):80–92, 1986.
- [vanEmden & Kowalski 76]
M. vanEmden and R. Kowalski. The semantics of predicate logic as a programming language. *Journal of the ACM*, 23(4), 1976.
- [vanErp *et al.* 13]
M. van Erp, G. Rizzo, and R. Troncy. Learning with the Web: Spotting Named Entities on the intersection of NERD and Machine Learning. In *Proceedings of the 3rd Workshop on Making Sense of Microposts (#MSM2013)*, 2013.
- [vanFraassen 71]
B. vanFraassen. *Formal Semantics and Logic*. The Macmillan Company, New York, 1971.
- [vanLehn 91]
K. vanLehn, editor. *Architectures for Intelligence*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1991.
- [vanRijsbergen 04]
K. van Rijsbergen. *The Geometry of Information Retrieval*. Cambridge, 2004.
- [Vapnik 98]
V. Vapnik. *Statistical Learning Theory*. John Wiley & Sons, 1998.
- [Vargas-Vera *et al.* 07]
M. Vargas-Vera, E. Moreale, A. Stutt, E. Motta, and F. Ciravegna. MNM: Semi-Automatic Ontology Population from Text. In *Ontologies*, volume 14 of *Integrated Series in Information Systems*, chapter 13, pages 373–402. Springer US, 2007.
- [Varol *et al.* 17]
O. Varol, E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini. Online human-bot interactions: Detection, estimation, and characterization. In *ICWSM*, pages 280–289. AAAI Press, 2017.

- [Vasiliki *et al.* 06]
 Z. Vasiliki, Y. Kompatsiaris, V. Papastathis, H. Eleftherohorinou, L. Wanner, G. Piella, E. Pianta, L. Serafini, and M. Giereth. Specification of the semantics-based content- and meta- representation formalism for patent documentation. Technical report, PATExpert EU Project Deliverable D2.1, 2006.
- [Vehviläinen *et al.* 06]
 A. Vehviläinen, E. Hyvönen, and O. Alm. A semi-automatic semantic annotation and authoring tool for a library help desk service. In *Proceedings of the first Semantic Authoring and Annotation Workshop*, November 2006.
- [Velardi & Pazienza 89]
 P. Velardi and M. T. Pazienza. Acquisition of lexical semantic knowledge from text. In G. Valle, editor, *Computational Intelligence '88*. North-Holland, 1989.
- [Velardi 91]
 P. Velardi. Acquiring a semantic lexicon for natural language processing. In U. Zernik, editor, *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*, pages 341–367. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1991.
- [Velardi *et al.* 89]
 P. Velardi, M. T. Pazienza, and S. Magrini. Acquisition of semantic patterns from a natural corpus of texts. *SIGART Newsletter, Special Issue on Knowledge Acquisition*, (108):115–123, 1989.
- [Velardi *et al.* 91]
 P. Velardi, M. T. Pazienza, and M. Fasolo. How to encode semantic knowledge: a method for learning representation. *Computational Linguistics*, 2(17), 1991.
- [Vendler 57]
 Z. Vendler. Verbs and times. *The philosophical review*, pages 143–160, 1957.
- [Venners 98]
 B. Venners. Introduction to Design Techniques. <http://www.javaworld.com/jw-02-1998/-jw-02-techniques.html>, February 1998.
- [Verhagen & Pustejovsky 08]
 M. Verhagen and J. Pustejovsky. Temporal Processing with the TARSQI Toolkit. In *CoLing 2008: Posters and Demonstrations*, pages 189–192, 2008.
- [Verma *et al.* 11]
 S. Verma, S. Vieweg, W. Corvey, L. Palen, J. H. Martin, M. Palmer, A. Schram, and K. M. Anderson. Natural language processing to the rescue? extracting “situational awareness” tweets during mass emergency. In *Fifth International AAAI Conference on Weblogs and Social Media*, 2011.
- [Veronis & Ide 91]
 J. Veronis and N. Ide. An assessment of semantic information automatically extracted from machine readable dictionaries. In *Proceedings of EACL '91*, pages 227–232, Berlin, 1991.
- [Veronis & Ide 96]
 J. Veronis and N. Ide. Considerations for the Reusability of Linguistic Software. Technical report, EAGLES, April 1996. <http://w3.lpl.univ-aix.fr/projects/multext/LSD/LSD1.html>.
- [Veronis 96]
 J. Veronis. Guidelines for Linguistic Software Development. Technical report, MULTEXT/EAGLES, April 1996. <http://w3.lpl.univ-aix.fr/projects/multext/LSD/LSD2.html>.
- [Veronis 98]
 J. Veronis. A study of polysemy judgements and inter-annotator agreement. In *Programme and Advanced Papers of the Senseval Workshop*, Herstmonceux Castle, England, 1998.

- [Vert *et al.* 04]
 J. P. Vert, K. Tsuda, and B. Scholkopf. A primer on kernel methods. *Kernel Methods in Computational Biology*, pages 35–70, 2004.
- [Vertanen & Kristensson 11]
 K. Vertanen and P. O. Kristensson. The imagination of crowds: conversational aac language modeling using crowdsourcing and large data sources. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 700–711, 2011.
- [Vickrey *et al.* 08]
 D. Vickrey, A. Bronzan, W. Choi, A. Kumar, J. Turner-Maier, A. Wang, and D. Koller. Online word games for semantic data collection. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 533–542, 2008.
- [Vigna & Boldi 07]
 S. Vigna and P. Boldi. Efficient optimally lazy algorithms for minimal-interval semantics. *CoRR*, abs/0710.1525, 2007.
- [Vilain & Day 96]
 M. Vilain and D. Day. Finite-state phrase parsing by rule sequences. In *Proceedings of COLING-96*, Copenhagen, 1996.
- [Vilain 93]
 M. Vilain. Validation of terminological inference in an information extraction task. In *Proceedings of the 1993 ARPA Human Language Workshop*, 1993.
- [Vilain *et al.* 95]
 M. Vilain, J. Burger, J. Aberdeen, D. Connolly, and L. Hirschman. A model-theoretic coreference scoring scheme. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, pages 45–52. Morgan Kaufmann, California, 1995.
- [Villazón-Terrazas *et al.* 08]
 B. Villazón-Terrazas, M. Suárez-Figueroa, A. Gómez-Pérez, A. García-Silva, and D. Maynard. Methods and tools for re-engineering non-ontological resources. Technical Report D2.2.2, NeOn Project Deliverable, 2008.
- [Villemonde de la Clergerie & Lang 94]
 E. Villemonde de la Clergerie and B. Lang. LPDA: Another look at tabulation in logic programming. In V. Hentenryck, editor, *Proceedings of the 11th International Conference on Logic Programming (ICLP'94)*, pages 470–486. MIT Press, June 1994.
- [Vintsyuk 71]
 T. K. Vintsyuk. Element-wise recognition of continuous speech composed of words from a specified dictionary. *Kibernetika*, 7:133–143, 1971.
- [Viola & Jones 01]
 P. Viola and M. Jones. Robust real-time object detection. In *International Journal of Computer Vision*, 2001.
- [Vlachos 04]
 A. Vlachos. Active learning with support vector machines. MSc thesis, University of Edinburgh, 2004.
- [Vlachos 07]
 A. Vlachos. Evaluating and combining biomedical named entity recognition systems. In *Poster Proceedings of BioNLP at ACL*, 2007.

- [Voelkel *et al.* 06]
M. Voelkel, M. Kroetzsch, D. Vrandečić, H. Haller, and R. Studer. Semantic Wikipedia. In *Proceedings of the International World Wide Web Conference (WWW'06)*, pages 585–594, Edinburgh, Scotland, 2006. ACM.
- [Vogel *et al.* 96]
S. Vogel, H. Ney, and C. Tillmann. HMM-based Word Alignment in Statistical Translation. In *Int. Conf. on Computational Linguistics*, volume 2, pages 836–841, August 1996.
- [Volk *et al.* 03a]
M. Volk, B. Ripplinger, S. Vintar, P. Buitelaar, D. Raileanu, and B. Sacaleanu. Semantic annotation for concept-based cross-language medical information retrieval. *submitted to International Journal of Medical Informatics*, 2003.
- [Volk *et al.* 03b]
M. Volk, S. Vintar, and P. Buitelaar. Ontologies in cross-language information retrieval. In *Proceeding of the Workshop Ontologie-basieres Wissensmanagement (WOW2003)*, Luzern, Switzerland, 2003.
- [von Hahn 94]
W. von Hahn. The Architecture Problem in Natural Language Processing. *Prague Bulletin of Mathematical Linguistics*, 61:48–69, 1994.
- [vonAhn & Dabbish 04a]
L. von Ahn and L. Dabbish. Labeling images with a computer game. In *ACM Conference on Human Factors in Computing Systems, CHI 2004*, pages 319–326, Vienna, Austria, 2004. ACM.
- [vonAhn & Dabbish 04b]
L. von Ahn and L. Dabbish. Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '04*, pages 319–326, New York, NY, USA, 2004. ACM.
- [vonAhn & Dabbish 08]
L. von Ahn and L. Dabbish. Designing games with a purpose. *Commun. ACM*, 51(8):58–67, August 2008.
- [vonAhn 06]
L. von Ahn. Games with a purpose. *Computer*, 39(6):92–94, June 2006.
- [vonAhn *et al.* 06]
L. von Ahn, M. Kedia, and M. Blum. Verbosity: a game for collecting common-sense facts. In *Proceedings of the SIGCHI conference on Human Factors in computing systems, CHI '06*, pages 75–78, New York, NY, USA, 2006. ACM.
- [Voorhees & Harman 99]
E. M. Voorhees and D. Harman. Overview of the eighth Text REtrieval Conference (TREC-8). In *The Eighth Text REtrieval Conference (TREC-8), 16–19 November 1999*, pages 1–24, National Institute of Standards and Technology (NIST), 1999.
- [Voorhees 98a]
E. Voorhees. Using WordNet for Text Retrieval. In C. Fellbaum, editor, *WordNet: an electronic lexical database*. MIT Press, 1998.
- [Voorhees 98b]
E. M. Voorhees. Variations in Relevance Judgments and the Measurement of Retrieval Effectiveness. In *Proceedings of the 21st annual international ACM SIGIR*, 1998.
- [Vossen 97]
P. Vossen. EuroWordNet: a multilingual database for information retrieval. In *Proc. of the DELOS workshop on Cross-Language Information Retrieval*, Zurich, 1997.

- [Vossen *et al.* 97a]
P. Vossen, P. Diez-Orzas, and W. Peters. The Multilingual Design of EuroWordNet. In *Proceedings of the ACL/EACL-97 workshop on Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*, Madrid, 1997.
- [Vossen *et al.* 97b]
P. Vossen, W. Peters, and P. Diez-Orzas. The multilingual design of the EuroWordNet database. In K. Mahesh, editor, *Ontologies and multilingual NLP, Proceedings of workshop at IJCAI-97*, Nagoya, Japan, 1997.
- [Vouros & Spyropoulos 89]
G. Vouros and C. Spyropoulos. A methodology for conceptual representation of knowledge using attribute grammars. *Augewandte Informatik*, 9:287–293, 1989.
- [Vouros & Spyropoulos 91]
G. Vouros and C. Spyropoulos. The PHOS conceptual language for knowledge representation. In S. G. Tzafestas, editor, *Engineering Systems with Intelligence. Concepts, Tools and Applications*, pages 43–50. Kluwer, Academic, Amsterdam, 1991.
- [Vouros & Spyropoulos 94]
G. Vouros and C. Spyropoulos. The PHOS conceptual formalism for knowledge representation. Accepted under revision from the IEEE Trans. On Systems Man and Cybernetics, 1994.
- [Vouros 92]
G. Vouros. *Structure and Organisation of Conceptual Models - An Epistemological Approach*. Unpublished PhD thesis, Department of Informatics, University of Athens, April 1992. In Greek.
- [Voyer *et al.* 10]
R. Voyer, V. Nygaard, W. Fitzgerald, and H. Copperman. A hybrid model for annotating named entity training corpora. In *Proceedings of the fourth linguistic annotation workshop*, pages 243–246. Association for Computational Linguistics, 2010.
- [Vrandecic & Kroetzsch 06]
D. Vrandecic and M. Kroetzsch. Reusing Ontological Background Knowledge in Semantic Wikis. In *Proceedings of the First Workshop on Semantic Wikis (SemWiki'06)*, 2006.
- [Vrandecic *et al.* 05]
D. Vrandecic, H. S. Pinto, Y. Sure, and C. Tempich. The DILIGENT Knowledge Processes. *Journal of Knowledge Management*, 5(9):85–96, 2005.
- [Wache *et al.* 05]
H. Wache, L. Serafino, and R. G. Castro. Scalability - state of the art. Technical Report D2.1.1, KnowledgeWeb Deliverable, 2005.
- [Wacholder *et al.* 04]
S. Wacholder, S. Chanock, M. Garcia-Closas, L. El-guormli, and N. Rothman. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst*, 96(6):434–442, March 2004.
- [Wagner *et al.* 99]
J. Wagner, J. E. Rogers, R. H. Baud, and J.-R. Scherrer. Natural language generation of surgical procedures. *International Journal of Medical Informatics*, 53:175–192, 1999.
- [Wahlster 00]
W. Wahlster, editor. *Verbmobil: Foundations of speech-to-speech translations*. Springer Verlag, Berlin, Germany, July 2000.

- [Wahlster *et al.* 93]
W. Wahlster, E. Andre, W. Finkler, H.-J. Profitlich, and T. Rist. Plan-based integration of natural language and graphics generation. *Artificial Intelligence*, 63:387 – 427, 1993.
- [Waibel 88]
A. Waibel. *Prosody and Speech Recognition*. Pitman, London, 1988.
- [Waismann 65]
F. Waismann. *The principles of linguistic philosophy*. Macmillan, London, 1965.
- [Waitelonis *et al.* 11]
J. Waitelonis, N. Ludwig, M. Knuth, and H. Sack. WhoKnows? Evaluating linked data heuristics with a quiz that cleans up DBpedia. *Interactive Technology and Smart Education*, 8(4):236–248, 2011.
- [Wakao *et al.* 96a]
T. Wakao, R. Gaizauskas, and K. Humphries. Evaluation of an algorithm for the recognition and classification of proper names. In *Proceedings of the 16th International Conference on Computational Linguistics (COLING96)*, pages 418–423, Copenhagen, Denmark, 1996.
- [Wakao *et al.* 96b]
T. Wakao, R. Gaizauskas, and Y. Wilks. Evaluation of an algorithm for the recognition and classification of proper names. In *Proceedings of the 16th International Conference on Computational Linguistics (COLING96)*, pages 418–423, Copenhagen, 1996.
- [Wakefield 07]
J. Wakefield. A bayesian measure of the probability of false discovery in genetic epidemiology studies. *The American Journal of Human Genetics*, 81(2):208–227, August 2007.
- [Walker & Amsler 86a]
D. Walker and R. Amsler. The use of machine-readable dictionaries in sub-language analysis. In *Analysing language in restricted domains: sublanguage description and processing*, pages 69–84. Lawrence Erlbaum, 1986.
- [Walker & Amsler 86b]
D. E. Walker and R. A. Amsler. The use of machine-readable dictionaries in sublanguage analysis. In R. Grishman and R. Kittredge, editors, *Analyzing Language in Restricted Domains*. Lawrence Erlbaum, Hillsdale, NJ, 1986.
- [Walker & Amsler 86c]
D. E. Walker and R. A. Amsler. The use of machine-readable dictionaries in sublanguage analysis. In R. Grishman and R. Kittredge, editors, *Analyzing Language in Restricted Domains: Sublanguage Description and Processing*, pages 69–83. Lawrence Erlbaum Associates, Hillsdale, NJ, 1986.
- [Walker & Moore 97]
M. A. Walker and J. Moore. Empirical studies in discourse. *Computational Linguistics*, 23(1):1–12, 1997.
- [Waller 82]
R. Waller. Text as diagram: Using typography to improve access and understanding. In D. Jonassen, editor, *The Technology of Text*, pages 137–166. Educational Technology Publications, 1982.
- [Waltman & Eck 12]
L. Waltman and N. J. Eck. A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, 63(12):2378–2392, 2012.

- [Wang & Callison-Burch 10]
R. Wang and C. Callison-Burch. Cheap facts and counter-facts. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 163–167, 2010.
- [Wang & He 08]
W. Wang and Q. He. A survey on emotional semantic image retrieval. In *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, pages 117–120, 2008.
- [Wang *et al.* 05a]
T. Wang, D. Maynard, W. Peters, K. Bontcheva, and H. Cunningham. Extracting a domain ontology from linguistic resource based on relatedness measurements. In *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2005)*, pages 345–351, Compiègne, France, Septmeber 2005.
- [Wang *et al.* 05b]
W. Y. Wang, B. J. Barratt, D. G. Clayton, and J. A. Todd. Genome-wide association studies: Theoretical and practical concerns. *NATURE REVIEWS-GENETICS*, 6:109–118, February 2005.
- [Wang *et al.* 06]
T. Wang, Y. Li, K. Bontcheva, H. Cunningham, and J. Wang. Automatic Extraction of Hierarchical Relations from Text. In *Proceedings of the Third European Semantic Web Conference (ESWC 2006)*, Budva, Montenegro, 2006.
- [Wang *et al.* 08]
H. Wang, K. Zhang, Q. Liu, T. Tran, and Y. Yu. Q2semantic: A lightweight keyword interface to semantic search. In S. Bechhofer, M. Hauswirth, J. Hoffmann, and M. Koubarakis, editors, *The Semantic Web: Research and Applications*, volume 5021 of *Lecture Notes in Computer Science*, pages 584–598. Springer Berlin / Heidelberg, 2008.
- [Wang *et al.* 10a]
H. H. Wang, J. Sun, and D. Damjanovic. Enhanced Semantic Access to Formal Software Models. In *Proceedings of the 12th International Conference on Formal Engineering Methods*, pages 237–252. Springer-Verlag, November 2010.
- [Wang *et al.* 10b]
L. Wang, J. McGee, and S. Huang. Sumtweets. Technical report, Texas A & M University, 2010.
- [Wang *et al.* 11]
H. Wang, T. Tran, C. Liu, and L. Fu. Lightweight integration of ir & db for scalable hybrid search with integrated ranking support. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(4), 2011.
- [Wang *et al.* 12]
A. Wang, C. Hoang, and M. Y. Kan. Perspectives on Crowdsourcing Annotations for Natural Language Processing. *Language Resources and Evaluation*, Mar:1–23, 2012.
- [Wang *et al.* 13]
L. Wang, H. Raghavan, V. Castelli, R. Florian, and C. Cardie. A sentence compression based framework to query-focused multi-document summarization. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1384–1394, Sofia, Bulgaria, August 2013. Association for Computational Linguistics.
- [Wanner *et al.* 06]
L. Wanner, S. Brüggemann, B. Diallo, M. Giereth, Y. Kompatsiaris, E. Pianta, G. Rao, P. Schoester, and V. Zervaki. Patexpert: Semantic processing of patent documentation. In *SAMT (Posters and Demos)*, 2006.

- [Wanner *et al.* 08]
 L. Wanner, R. Baeza-Yates, S. Brugmann, J. Codina, B. Diallo, E. Escorsa, M. Giereth, Y. Kompatsiaris, S. Papadopoulos, E. Pianta, G. Piella, I. Puhmann, G. Rao, M. Rotard, P. Schoester, L. Serafini, and V. Zervaki. Towards Content-oriented Patent Document Processing. *World Patent Information*, 30(1):21–33, 2008.
- [Ward 94]
 N. Ward. An approach to tightly-coupled syntactic/semantic processing for speech understanding. In *Proceedings of the AAAI Workshop on Integration of Natural Language and Speech Processing*, pages 50–57, Seattle, 1994.
- [Warmuth *et al.* 02]
 M. Warmuth, G. Ratsch, M. Mathieson, J. Liao, and C. Lemmen. Active learning in the drug discovery process. In *Advances in Neural Information Processing Systems 15 (NIPS-2002)*, pages 1449–1456, 2002.
- [Warren 77]
 D. Warren. Implementing prolog - compiling predicate logic programs volumes 1 & 2. D.a.i. research reports nos. 39 & 40, department of artificial intelligence, University of Edinburgh, 1977.
- [Warren 82]
 D. Warren. Using semantics in non-context-free parsing of montague grammar. *American Journal of Computational Linguistics*, 8(3-4):123–138, 1982.
- [Warren *et al.* 06]
 P. Warren, I. Thurlow, and D. Alsmeyer. Applying Semantic Technology to a Digital Library. In J. Davies, R. Studer, and P. Warren, editors, *Semantic Web Technologies*. John Wiley and Sons, 2006.
- [Wartena *et al.* 07]
 C. Wartena, R. Brussee, L. Gazendam, and W.-O. Huijsen. Apolda: A practical tool for semantic annotation. In *In Proceedings of 18th International Workshop on Database and Expert Systems Applications*, pages 288–292, The Netherlands, September 2007.
- [Watanabe 96]
 H. Watanabe. A method for abstracting newspaper articles by using surface clues. In *Proc. of 16th International Conference on Computational Linguistics*, volume 2, pages 974–979, Copenhagen, Denmark, 1996.
- [Waterworth 87]
 T. Waterworth. *Speech and Language-Based Interaction With Machines*. Ellis Horwood, Chichester, 1987.
- [Webb 00]
 N. Webb. Rule-based dialogue management systems. In *Proceedings of the 3rd International Workshop on Human-Computer Conversation*, Bellagio, Italy, 2000.
- [Webber 79]
 B. Webber. *A Formal Approach to Discourse Anaphora*. Garland, New York, 1979.
- [Weeber *et al.* 05]
 M. Weeber, J. A. Kors, and B. Mons. Online tools to support literature-based discovery in the life sciences. *Briefings in Bioinformatics*, 6(3):277–286, 2005.
- [Wei-ning *et al.* 06]
 W. Wei-ning, Y. Ying-lin, and J. Sheng-ming. Image retrieval by emotional semantics: A study of emotional space and feature extraction. In *Systems, Man and Cybernetics, 2006. SMC '06. IEEE International Conference on*, volume 4, pages 3534–3539, 2006.

- [Weichselbraun *et al.* 10]
 A. Weichselbraun, S. Gindl, and A. Scharl. A context-dependent supervised learning approach to sentiment detection in large textual databases. *Journal of Information and Data Management*, 1(3):329–342, 2010.
- [Weichselbraun *et al.* 11]
 A. Weichselbraun, S. Gindl, and A. Scharl. Using games with a purpose and bootstrapping to create domain-specific sentiment lexicons. In *Proceedings of the 20th ACM Conference on Information and Knowledge Management (CIKM)*, pages 1053–1060, 2011.
- [Weinstein *et al.* 99]
 P. Weinstein, W. Birmingham, and E. Durfee. Agent-Based Digital Libraries: Decentralization and Coordination. *IEEE Communication Magazine*, pages 110–115, 1999.
- [Weintraub *et al.* 97]
 M. Weintraub, F. Beaufays, Z. Rivlin, Y. Konig, and A. Stolcke. Neural-Network Based Measures of Confidence for Word Recognition. In *Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 2, pages 887–890, Munich, Germany, April 1997.
- [Weischedel 95]
 R. Weischedel. Description of the PLUM System as used for MUC-6. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, pages 55–70, San Francisco, 1995. Morgan Kaufmann, California.
- [Weischedel *et al.* 93]
 R. Weischedel, M. Meteer, R. Schwartz, L. Ramshaw, and J. Palmucci. Coping with ambiguity and unknown words through probabilistic models. *Computational Linguistics (Special Issue on Using Large Corpora: II)*, 19:359–382, 1993.
- [Wellner *et al.* 04]
 B. Wellner, A. McCallum, F. Peng, and M. Hay. An integrated, conditional model of information extraction and coreference with application to citation matching. In *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, pages 593 – 601, Banff, Canada, 2004.
- [Welty & Murdock 06]
 C. Welty and J. W. Murdock. Towards Knowledge Acquisition from Information Extraction. In *5th Internal Semantic Web Conference (ISWC'06)*. Springer, 2006.
- [Welty 98a]
 C. A. Welty. DLs for DLs : Description Logics for Digital Libraries. In *Proceedings of the 1998 International Workshop for Description Logics*, Trento, Italy, 1998.
- [Welty 98b]
 C. A. Welty. Towards an Ontology for Library Modalities. In *Proceedings of the AAAI-98 Workshop on Representations for Multi-Modal Human-Computer Interaction.*, July 1998.
- [Weng & Gennari 04]
 C. Weng and J. Gennari. Asynchronous collaborative writing through annotations. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 578 – 581, Chicago, Illinois, USA, 2004. ACM.
- [Weng *et al.* 11a]
 J. Y. Weng, C. L. Yang, B. N. Chen, Y. K. Wang, and S. D. Lin. IMASS: An Intelligent Microblog Analysis and Summarization System. In *Proceedings of the ACL-HLT 2011 System Demonstrations*, pages 133–138, Portland, Oregon, 2011.

- [Weng *et al.* 11b]
 J.-Y. Weng, C.-L. Yang, B.-N. Chen, Y.-K. Wang, and S.-D. Lin. IMASS: An Intelligent Microblog Analysis and Summarization System. In *Proceedings of the ACL-HLT 2011 System Demonstrations*, pages 133–138, Portland, Oregon, June 2011. Association for Computational Linguistics.
- [Wenger 98]
 E. Wenger. *Communities of Practice: The Key to Knowledge Strategy*. Cambridge University Press, Cambridge, England, 1998.
- [Wessa 07]
 Wessa. Kendall tau Rank Correlation (v1.0.7) in Free Statistics Software (v1.1.21). http://www.wessa.net/rwasp_kendall.wasp/, 2007.
- [Wessel *et al.* 98]
 F. Wessel, K. Macherey, and R. Schlüter. Using Word Probabilities as Confidence Measures. In *Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 1, pages 225–228, Seattle, Washington, USA, May 1998.
- [Wessel *et al.* 01]
 F. Wessel, R. Schlüter, K. Macherey, and H. Ney. Confidence Measures for Large Vocabulary Continuous Speech Recognition. *IEEE Transactions on Speech and Audio Processing*, 9(3):288–298, March 2001.
- [Westerski *et al.*]
 A. Westerski, C. A. Iglesias, and F. T. Rico. Linked Opinions: Describing Sentiments on the Structured Web of Data. In *Proceedings of the 4th International Workshop “Social Data on the Web” (SDoW2011)*.
- [Wexler & Culicover 80]
 K. Wexler and P. Culicover. *Formal Principles of Language Acquisition*. MIT Press, Cambridge MA, 1980.
- [Weyhrauch 80]
 R. Weyhrauch. Prolegomena to a theory of mechanized formal reasoning. *Artificial Intelligence*, 13, 1980.
- [Wheddon & Linggard 90]
 C. Wheddon and R. Linggard. *Speech and Language Processing*. Chapman and Hall, London, 1990.
- [White & Caldwell 98]
 M. White and D. E. Caldwell. EXEMPLARS: A practical, extensible framework for dynamic text generation. In *Proceedings of the 9th International Natural Language Generation Workshop (INLG’98)*, Niagara-on-the-Lake, Ontario, Canada, 1998.
- [White 09]
 T. White. *Hadoop: The Definitive Guide*. O’Reilly Media, Sebastopol, CA, May 2009.
- [White *et al.* 01]
 M. White, T. Korelsky, C. Cardie, V. Ng, D. Pierce, and K. Wagstaff. Multi-document Summarization via Information Extraction. In *Proceedings of the International Conference on Human Language Technology research*, 2001.
- [Whittemore & Macpherson 91]
 S. Whittemore and M. Macpherson. Event-building through role-filling and anaphora resolution. In *Proceedings of the 29th meetings of the Association for Computational Linguistics, Berkeley : California*, 1991.

- [Wibberley *et al.* 14]
S. Wibberley, J. Reffin, and D. Weir. Method51 for mining insight from social media datasets. *COLING 2014*, page 115, 2014.
- [Widdows & Bruza 07]
D. Widdows and P. Bruza. Quantum information dynamics and open world science. In *Proceedings of the First Quantum Interaction Symposium (QI-2007)*. AAAI Press, 2007.
- [Widdows & Peters 03]
D. Widdows and S. Peters. Word vectors and quantum logic: Experiments with negation and disjunction. In *Eighth Mathematics of Language Conference*, pages 141–154, 2003.
- [Widdows 03a]
D. Widdows. A Mathematical Model for Context and Word-Meaning. In *Proceedings of The Fourth International and Interdisciplinary Conference on Modeling and Using Context*, pages 369–382, Stanford, California, USA, 2003.
- [Widdows 03b]
D. Widdows. Orthogonal negation in vector spaces for modelling word-meanings and document retrieval. In *The 41st Annual Meeting of the Association for Computational Linguistics*, pages 136–143, 2003.
- [Widdows 04]
D. Widdows. *Geometry and Meaning*. CSLI Publications, 2004.
- [Widdows 08]
D. Widdows. Semantic vector products: Some initial investigations. In *Proceedings of the Second Quantum Interaction Symposium (QI-2008)*. College Publications, 2008.
- [Wiebe *et al.* 02]
J. Wiebe, E. Breck, C. Buckley, C. Cardie, P. Davis, B. Fraser, D. Litman, D. Pierce, E. Riloff, and T. Wilson. Nrrc summer workshop on multiple-perspective question answering final report. Technical report, 2002.
- [Wiebe *et al.* 04]
J. Wiebe, T. Wilson, R. Bruce, M. Bell, and M. Martin. Learning Subjective Language. *Computational Linguistics*, 30(3):277–308, 2004.
- [Wiebe *et al.* 05]
J. Wiebe, T. Wilson, and C. Cardie. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 39(2-3):165–210, 2005.
- [Wiener 93]
L. Wiener. *Digital Woes: Why We Should Not Depend on Software*. Addison-Wesley, Reading, Ma, 1993.
- [Wierenga & Eaton 93]
D. Wierenga and C. Eaton. The drug development and approval process. Technical report, Pharmaceutical Manufacturers Association, Washington. D.C., 1993.
- [Wierzbicka 89]
A. Wierzbicka. *Semantics, Culture and Cognition*. Oxford University Press, 1989.
- [Wightman & Ostendorf 92]
C. W. Wightman and M. Ostendorf. Automatic recognition of intonational features. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 1, pages 221–234, San Francisco, 1992.

- [Wikipedia 06]
Wikipedia. Sumerian language — wikipedia, the free encyclopedia, 2006.
- [Wilcock & Jokinen 03]
G. Wilcock and K. Jokinen. Generating Responses and Explanations from RDF/XML and DAML+OIL. In *Knowledge and Reasoning in Practical Dialogue Systems, IJCAI-2003*, pages 58–63, Acapulco, 2003.
- [Wilcock 03]
G. Wilcock. Talking OWLs: Towards an Ontology Verbalizer. In *Human Language Technology for the Semantic Web and Web Services, ISWC'03*, pages 109–112, Sanibel Island, Florida, 2003.
- [Wilcoxon 45]
F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):pp. 80–83, 1945.
- [Wilkins 86]
M. Wilkins. A model-theoretic approach to updating logical databases. Report no. stan-cs-86-1096, Department of Computer Science, Stanford University, Stanford, CA, 1986.
- [Wilks & Catizone 99a]
Y. Wilks and R. Catizone. Can We Make Information Extraction More Adaptive? In *M. Pazienza (ed.) Proceedings of the SCIE99 Workshop*, pages 1–16, Rome, Italy, 1999.
- [Wilks & Catizone 99b]
Y. Wilks and R. Catizone. Can we make Information Extraction more Adaptive? In *Proceedings of SCIE'99 Workshop*, 1999.
- [Wilks & Catizone 02]
Y. Wilks and R. Catizone. What is lexical tuning? *Journal of Semantics*, 2002.
- [Wilks & Cunningham 99]
Y. Wilks and H. Cunningham. Tender for Adaptive Information Extraction. Unpublished Project Proposal, 1999.
- [Wilks & Gaizauskas 97a]
Y. Wilks and R. Gaizauskas. LaSIE Jumps the GATE. In T. Strazlkowski, editor, *Natural Language Information Retrieval*. Kluwer, 1997. In press.
- [Wilks & Gaizauskas 97b]
Y. Wilks and R. Gaizauskas. LaSIE jumps the GATE. In *Natural Language Processing: Extracting Information for Business Needs*, UNICOM, 1997.
- [Wilks & Gaizauskas 99]
Y. Wilks and R. Gaizauskas. Report on epsrc research grant on the large scale information extraction research project. Technical Report GR/K25267, University of Sheffield, 1999.
- [Wilks & Nirenburg 93]
Y. Wilks and S. Nirenburg. Large-scale knowledge base acquisition. In *Proceedings of the Conference on very large knowledge bases*. Tokyo, 1993.
- [Wilks & Stevenson]
Y. Wilks and M. Stevenson. The grammar of sense: Is word-sense tagging much more than part-of-speech tagging? Submitted to Computational Linguistics.
- [Wilks & Stevenson 97a]
Y. Wilks and M. Stevenson. Combining Independent Knowledge Sources for Word Sense Disambiguation. In *Proceedings of the Conference 'Recent Advances in Natural Language Processing'*, pages 1–7, Tzigrav Chark, Bulgaria, 1997.

- [Wilks & Stevenson 97b]
Y. Wilks and M. Stevenson. Sense Tagging: Semantic Tagging with a Lexicon. In *Proceedings of the SIGLEX Workshop on Tagging Text with Lexical Semantics*, pages 74–78, Washington, DC, 1997.
- [Wilks & Stevenson 97c]
Y. Wilks and M. Stevenson. Sense tagging: Semantic tagging with a lexicon. In *Proceedings of the SIGLEX Workshop "Tagging Text with Lexical Semantics: What, why and how?"*. ANLP, 1997. <http://xxx.lanl.gov/abs/cs.CL/9705016>.
- [Wilks & Stevenson 97d]
Y. Wilks and M. Stevenson. Sense tagging: Semantic tagging with a lexicon. In *Proceedings of the SIGLEX Workshop "Tagging Text with Lexical Semantics: What, why and how?"*. ANLP, 1997. <http://xxx.lanl.gov/abs/cs.CL/9705016>.
- [Wilks & Stevenson 97e]
Y. Wilks and M. Stevenson. Combining independent knowledge sources for word sense disambiguation. In *Proceedings of Recent Advances in Natural Language Processing Conference (RANLP)*, pages 1–7, Bulgaria, 1997.
- [Wilks & Stevenson 98]
Y. Wilks and M. Stevenson. Word sense disambiguation using optimised combinations of knowledge sources. In *Proceedings of COLING-ACL98*, Montreal, 1998.
- [Wilks & Stevenson 99]
Y. Wilks and M. Stevenson. The Grammar of Sense: Using part-of-speech tags as a first step in semantic disambiguation. *Journal of Natural Language Engineering* 4(3), 4(3), 1999.
- [Wilks 72]
Y. Wilks. *Grammar, Meaning and the Machine Analysis of Meaning*. Routledge and Kegan Paul, 1972.
- [Wilks 91a]
Y. Wilks. Diderot: a text extraction system. In *Proceedings of DARPA Speech and Language Workshop*, pages 79–86, San Mateo, CA, 1991. Morgan Kaufmann, California.
- [Wilks 91b]
Y. Wilks. Diderot: a text extraction system. In *DARPA Speech and Natural Language Workshop*. Morgan Kaufmann, California, San Mateo, CA, 1991.
- [Wilks 92]
Y. Wilks. Where am I coming from: The reversibility of analysis and generation in natural language processing. In M. Puetz, editor, *Thirty Years of Linguistic Evolution*. John Benjamins, 1992.
- [Wilks 94]
Y. Wilks. Developments in MT Research in the US. *Proceedings of ASLIB*, 46(4), 1994.
- [Wilks 95]
Y. Wilks. ECRAN: Extraction of Content: Research at Near-market. EU Research Proposal, 1995.
- [Wilks 96a]
Y. Wilks. Natural Language Processing, guest editorial. *Communications of the ACM*, 39(1), January 1996.
- [Wilks 96b]
Y. A. Wilks. Senses and texts. Submitted to *Computational Linguistics*, 1996.
- [Wilks 97]
Y. Wilks. Language processing and the thesaurus. Technical report, University of Sheffield, Department of Computer Science, 1997. CS-97-13.

- [Wilks 98]
 Y. Wilks. Is word-sense disambiguation just one more NLP task? In *Proceedings of the SENSEVAL Conference*, Herstmonceux, Sussex, 1998. Also appears as Technical Report CS-98-12, Department of Computer Science, University of Sheffield, 1998.
- [Wilks *et al.* 87]
 Y. Wilks, D. C. Fass, C. Guo, J. E. McDonald, T. Plate, and B. M. Slator. A tractable machine dictionary as a resource for computational semantics. In B. Boguraev and T. Briscoe, editors, *Computational Lexicography for Natural Language Processing*. Longman Group Limited, Harlow, Essex, 1987. Also MCCS-87-105, CRL/NMSU.
- [Wilks *et al.* 92]
 Y. Wilks, L. Guthrie, J. Guthrie, and J. Cowie. *Combining Weak Methods in Large-Scale Text Processing, in Jacobs 1992, Text-Based Intelligent Systems*. Lawrence Erlbaum, 1992.
- [Wilks *et al.* 93a]
 Y. Wilks, D. Fass, C. Guo, J. McDonald, T. Plate, and B. Slator. Providing machine tractable dictionary tools. In J. Pustejovsky, editor, *Semantics and the Lexicon*, pages 341–401. Kluwer Academic Press, 1993.
- [Wilks *et al.* 93b]
 Y. Wilks, L. Guthrie, and D. Farwell. The automatic acquisition of lexical entries for machine translation. *Journal of Machine Translation*, 1993.
- [Wilks *et al.* 95]
 Y. Wilks, L. Guthrie, and B. Slator. *Electric Words: Lexicons, Dictionaries and Meanings*. MIT Press, Cambridge, MA, 1995.
- [Wilks *et al.* 96a]
 Y. Wilks, L. Guthrie, and B. Slator. *Electric Words*. MIT Press, Cambridge, MA, 1996.
- [Wilks *et al.* 96b]
 Y. Wilks, B. M. Slator, and L. M. Guthrie. *Electric Words: Dictionaries, Computers and Meanings*. MIT Press, 1996.
- [Wilks *et al.* 97]
 Y. Wilks, R. Catizone, M. Stevenson, R. Collier, R. Basili, P. Velardi, and M. Paizienza. Towards a Well-Tuned lexicon for Information Extraction. *submitted*, 1997.
- [Willett 88]
 P. Willett. Recent trends in hierarchic document clustering: A critical review. *Information Processing & Management*, 24(5):577–597, 1988.
- [Wilms 96]
 G. J. Wilms. Using a hybrid system of corpus- and knowledge-based techniques to automate the induction of a lexical sublanguage grammar. In *Proc. of 16th International Conference on Computational Linguistics*, Copenhagen, Denmark, 1996.
- [Wilson 99]
 M. Wilson. Standards and Reference Architectures: Their Purpose and Establishment. In *Proceedings of the AISB'99 Workshop on Reference Architectures and Data Standards for NLP*, Edinburgh, April 1999. The Society for the Study of Artificial Intelligence and Simulation of Behaviour.
- [Wilson *et al.* 02]
 A. Wilson, P. Rayson, and T. McEnery, editors. *A Rainbow of Corpora: Corpus Linguistics and the Languages of the World*. Lincom-Europa, Munich, 2002.

- [Wilson *et al.* 09]
T. Wilson, J. Wiebe, and P. Hoffmann. Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. *Computational Linguistics*, 35(3):399–433, 2009.
- [Winnenburg *et al.* 08]
R. Winnenburg, T. Wächter, C. Plake, A. Doms, and M. Schroeder. Facts from text: can text mining help to scale-up high-quality manual curation of gene products with ontologies? *Briefings in Bioinformatics*, 9(6):466–478, 2008.
- [Winograd 72]
T. Winograd. *Understanding Natural Language*. Academic Press, New York, 1972.
- [Winograd 83]
T. Winograd. *Language as a Cognitive Process Volume 1: Syntax*. Addison-Wesley, Reading, MA, 1983.
- [Winograd 85]
T. Winograd. Moving the semantic fulcrum. *Linguistics and Philosophy*, 8(1):91–104, 1985.
- [Witte & Gitzinger 08]
R. Witte and T. Gitzinger. Semantic Assistants – User-Centric Natural Language Processing Services for Desktop Clients. In *3rd Asian Semantic Web Conference (ASWC 2008)*. Springer LNCS 5367, pages 360–374, 2008.
- [Witte *et al.* 07a]
R. Witte, Q. Li, Y. Zhang, and J. Rilling. Ontological text mining of software documents. In *NLDB*, volume 4592 of *Lecture Notes in Computer Science*, pages 168–180. Springer, 2007.
- [Witte *et al.* 07b]
R. Witte, Y. Zhang, and J. Rilling. Empowering Software Maintainers with Semantic Web Technologies. In E. Franconi, M. Kifer, and W. May, editors, *4th European Semantic Web Conference (ESWC 2007)*, number 4519 in LNCS, pages 37–52, Innsbruck, Austria, June 3–7 2007. Springer-Verlag Berlin Heidelberg. <http://www.eswc2007.org/pdf/eswc07-witte.pdf>.
- [Witten & Frank 99a]
I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. The Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann, San Francisco, CA, October 1999.
- [Witten & Frank 99b]
I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann, October 1999.
- [Witten *et al.* 94]
I. Witten, A. Moffat, and T. Bell. *Managing Gigabytes*. Van Nostrand Reinhold, New York, 1994.
- [Witten *et al.* 96]
I. Witten, C. Nevill-Manning, and S.J.Cunningham. Digital Libraries based on full-text retrieval. In *Proceedings of WebNet 96*, San Francisco, October 1996.
- [Witten *et al.* 03]
I. H. Witten, K. J. Don, M. Dewsnip, and V. Tablan. Text mining in a digital library. *Journal of Digital Libraries special issue 'Digital Libraries as experienced by the editors of the Journal'*, 2003.
- [Wolf *et al.* 11]
L. Wolf, M. Knuth, J. Osterhoff, and H. Sack. Risq! renowned individuals semantic quiz: a jeopardy like quiz game for ranking facts. In *Proceedings of the 7th International Conference on Semantic Systems, I-Semantics '11*, pages 71–78, New York, NY, USA, 2011. ACM.

- [Wolff 80a]
J. Wolff. Data compression, generalisation and overgeneralisation in an evolving theory of language development. *Artificial Intelligence and the Simulation of Behaviour*, 1980.
- [Wolff 80b]
J. Wolff. Language acquisition and the discovery of phrase structure. *Language and Speech*, 23 Part 3, 1980.
- [Wolff 84]
J. Wolff. Cognitive development as optimisation. In L. Bolc, editor, *Knowledge Based Learning Systems*. Springer-Verlag., 1984.
- [Wolff-Terroine 77]
M. Wolff-Terroine. Terminology and nomenclatures. In W. Schneider and A. Sagvall Hein, editors, *Computational Linguistics in Medicine*, pages 55–61. North Holland Publishing Company, 1977.
- [Wolinski *et al.* 98]
F. Wolinski, F. Vichot, and O. Gremont. Producing NLP-based On-line Contentware. In *Natural Language and Industrial Applications*, Moncton, Canada, 1998. <http://xxx.lanl.gov/abs/cs.CL/9809021>.
- [Wood *et al.* 03]
M. M. Wood, S. J. Lydon, V. Tablan, D. Maynard, and H. Cunningham. Using parallel texts to improve recall in IE. In *Recent Advances in Natural Language Processing*, Bulgaria, 2003.
- [Wood *et al.* 04]
M. Wood, S. Lydon, V. Tablan, D. Maynard, and H. Cunningham. Populating a Database from Parallel Texts using Ontology-based Information Extraction. In *Proceedings of NLDB 2004*, 2004. <http://gate.ac.uk/sale/nldb2004/NLDB.pdf>.
- [Woodland & Young 93]
P. C. Woodland and S. J. Young. The HTK tied-state continuous speech recogniser. In *Proceedings European Conference on Speech Communication and Technology*, Berlin, 1993.
- [Woodland *et al.* 94]
P. C. Woodland, J. J. Odell, V. Valtchev, and S. J. Young. Large vocabulary continuous speech recognition using HTK. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 2, pages 125–128, Adelaide, 1994.
- [Woodruff & Stonebreaker 95]
A. Woodruff and M. Stonebreaker. Buffering of Intermediate Results in Dataflow Diagrams. In *Proceedings VL'95 11th International IEEE Symposium on Visual Languages, Darmstadt*. IEEE Computer Society Press, 1995.
- [Wos *et al.* 84]
L. Wos, R. Overbeek, E. Lusk, and J. Boyle. *Automated Reasoning: Introduction and Applications*. Prentice-Hall, Englewood Cliffs, NJ, 1984.
- [Wright & Budin 97]
S. E. Wright and G. Budin, editors. *Handbook of Terminology Management*, volume 1 — Basic Aspects of Terminology Management. John Benjamins, Amsterdam, 1997.
- [Wright & Budin 01]
S. E. Wright and G. Budin, editors. *Handbook of Terminology Management*, volume 2 — Applications-Oriented Terminology Management. John Benjamins, Amsterdam, 2001.
- [Wright 78]
G. Wright. *The Illustrated Handbook of Sporting Terms*. Hampton House Productions Ltd., 1978.

- [Wright 90]
J. Wright. LR parsing of probabilistic grammars with input uncertainty for speech recognition. *Computer Speech and Language*, 4:297–323, 1990.
- [Wright *et al.* 02]
M. Wright, M. Marilino, and T. Sumner. Meta-design of a community digital library. *D-Lib Magazine*, 8(5), May 2002.
- [Wu & Palmer 94a]
Z. Wu and M. Palmer. Verbs semantics and lexical selection. In *Proceedings of the 32Nd Annual Meeting on Association for Computational Linguistics*, ACL '94, pages 133–138, Stroudsburg, PA, USA, 1994. Association for Computational Linguistics.
- [Wu & Palmer 94b]
Z. Wu and M. Palmer. Verbs semantics and lexical selection. In *ACL*, 1994.
- [Wu & Pottenger 05]
T. Wu and W. Pottenger. A Semi-Supervised Active Learning Algorithm for Information Extraction from Textual Data. *Journal of the American Society for Information Science and Technology*, 56(3):258–271, 2005.
- [Wu *et al.* 09]
Y. Wu, Q. Zhang, X. Huang, and L. Wu. Phrase dependency parsing for opinion mining. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 3 - Volume 3*, EMNLP '09, pages 1533–1541, Stroudsburg, PA, USA, 2009. Association for Computational Linguistics.
- [Wu *et al.* 10]
W. Wu, B. Zhang, and M. Ostendorf. Automatic generation of personalized annotation tags for twitter users. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 689–692, 2010.
- [Wu *et al.* 11]
S. Wu, J. M. Hofman, W. A. Mason, and D. J. Watts. Who says what to whom on twitter. In *Proceedings of the 20th International Conference on World Wide Web*, WWW '11, pages 705–714, New York, NY, USA, 2011. ACM.
- [Wyner *et al.* 10]
A. Wyner, K. Angelov, G. Barzdins, D. Damljanovic, B. Davis, N. Fuchs, S. Hoefler, K. Jones, K. Kaljurand, T. Kuhn, M. Luts, J. Pool, M. Rosner, R. Schwitter, and J. Sowa. On controlled natural languages: Properties and prospects. In N. Fuchs, editor, *Controlled Natural Language*, volume 5972 of *Lecture Notes in Computer Science*, pages 281–289. Springer Berlin / Heidelberg, 2010. 10.1007/978-3-642-14418-9_17.
- [Wynne 02]
M. Wynne. The language resource archive of the 21st century. In *Proceedings of 3rd Language Resources and Evaluation Conference (LREC'2002)*, Gran Canaria, Spain, 2002.
- [Xin *et al.* 11]
W. Xin, Z. Jing, J. Jing, H. Yang, S. Palakorn, W. X. Zhao, J. Jiang, J. He, Y. Song, P. Achananurparp, E. P. Lim, and X. Li. Topical keyphrase extraction from Twitter. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, HLT '11, pages 379–388, 2011.
- [Xu *et al.* 12]
Xu, Bhargava, Nowak, and Zhu. Socioscope: Spatio-temporal signal recovery from social media. In *Proceedings of the European Conference on Machine Learning: PKDD*, 2012.

- [Xue & Shen 03]
N. Xue and L. Shen. Chinese Word Segmentation as LMR Tagging. In *Proceedings of the 2nd SIGHAN Workshop on Chinese Language Processing*, 2003.
- [Yamada & Matsumoto 03]
H. Yamada and Y. Matsumoto. Statistical Dependency Analysis with Support Vector machines. In *The 8th International Workshop of Parsing Technologies (IWPT2003)*, 2003.
- [Yan *et al.* 12a]
R. Yan, M. Lapata, and X. Li. Tweet recommendation with graph co-ranking. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, pages 516–525, Jeju Island, Korea, 2012.
- [Yan *et al.* 12b]
R. Yan, M. Lapata, and X. Li. Tweet recommendation with graph co-ranking. In *ACL*, 2012.
- [Yang & Chute 94]
Y. Yang and C. Chute. Words or concepts: the features of indexing units and their optimal use in information retrieval. In *Proc. of SCAMC '94*, pages 685–689, 1994.
- [Yang & Liu 99]
Y. Yang and X. Liu. A Re-Examination of Text Categorization Methods. In *Proceedings of ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 42–49, 1999.
- [Yang 98]
Y. Yang. An evaluation of statistical approaches to text categorization. *Journal of Information Retrieval*, 1:67–88, 1998.
- [Yang 01]
Y. Yang. A study on thresholding strategies for text categorization. In *Proceedings of ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'01)*, pages 137–145, New York, 2001.
- [Yang *et al.* 02]
Y. Yang, J. Zhang, J. Carbonell, and C. Jin. Topic conditioned novelty detection. In *Proceedings 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'02)*, pages 688–693, Edmonton, Canada, July 2002.
- [Yangarber 01]
R. Yangarber. *Scenario Customization for Information Extraction*. Unpublished PhD thesis, New York University, USA, Jan 2001.
- [Yangarber 03]
R. Yangarber. Counter-Training in Discovery of Semantic Patterns. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL'03)*, 2003.
- [Yangarber *et al.* 00]
R. Yangarber, R. Grishman, P. Tapanainen, and S. Huttunen. Unsupervised Discovery of Scenario-level Patterns for Information Extraction. In *Proceedings of ANLP-NAACL'00*, Seattle, WA, 2000.
- [Yankova *et al.* 08]
M. Yankova, H. Saggion, and H. Cunningham. Adopting ontologies for multisource identity resolution. In A. Duke, M. Hepp, K. Bontcheva, and M. B. Vilain, editors, *OBI*, volume 308 of *ACM International Conference Proceeding Series*, page 6. ACM, 2008.
- [Yano *et al.* 10]
T. Yano, P. Resnik, and N. A. Smith. Shedding (a thousand points of) light on biased language. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 152–158, 2010.

- [Yanulevskaya *et al.* 08]
V. Yanulevskaya, J. Van Gemert, K. Roth, A. K. Herbold, N. Sebe, and J.-M. Geusebroek. Emotional valence categorization using holistic image features. In *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, pages 101–104, 2008.
- [Yardi & Boyd 10]
S. Yardi and D. Boyd. Tweeting from the town square: Measuring geographic local networks. In *Proceedings of ICWSM*, 2010.
- [Yarowski 92]
D. Yarowski. Word-sense disambiguation using statistical models of Roget’s categories trained on large corpora. In *COLING*, pages 454–460, 1992.
- [Yarowski 95]
D. Yarowski. Unsupervised Word Sense Disambiguation Rivaling Supervised Methods. In *Proceedings of the Association for Computational Linguistics*, pages 189–196, June 1995.
- [Yarowsky & Ngai 01]
D. Yarowsky and G. Ngai. Inducing multilingual POS taggers and NP bracketers via robust projection across aligned corpora. In *Proceedings of NAACL-01*, 2001.
- [Yarowsky 92a]
D. Yarowsky. Word-sense disambiguation using statistical models of Roget’s categories trained on large corpora. In *COLING-92*, 1992.
- [Yarowsky 92b]
D. Yarowsky. Word sense disambiguation using statistical models of Roget’s categories trained on large corpora. In *Proc. of 14th International Conference on Computational Linguistics*, pages 454–460, 1992.
- [Yarowsky 93]
D. Yarowsky. One sense per collocation. In *Proceedings ARPA Human Language Technology Workshop*, pages 266–271, 1993.
- [Yarowsky 95a]
D. Yarowsky. Unsupervised word-sense disambiguation rivaling supervised methods. In *Proceedings of ACL95*, 1995.
- [Yarowsky 95b]
D. Yarowsky. Unsupervised word sense disambiguation rivaling supervised methods. In *Proc. of 33rd Annual Meeting of the Association for Computational Linguistics, ACL’95*, pages 189–196, 1995.
- [Yarowsky 00]
D. Yarowsky. Hierarchical decision lists for word sense disambiguation. *Computers and the Humanities*, 34(1-2), 2000. <http://citeseer.nj.nec.com/yarowsky99hierarchicall.html>.
- [Ye *et al.* 02]
S. Ye, T.-S. Chua, and L. Jimin. An agent-based approach to chinese named entity recognition. In *Proceedings of 19th International Conference on Computational Linguistics (COLING-02)*, Taipei, Taiwan, 2002.
- [Yimam *et al.* 14]
S. M. Yimam, C. Biemann, R. Eckart de Castilho, and I. Gurevych. Automatic annotation suggestions and custom annotation layers in webanno. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 91–96, Baltimore, Maryland, June 2014. Association for Computational Linguistics.

- [Yin *et al.* 13]
P. Yin, H. Wang, and K. Guo. Feature–opinion pair identification of product reviews in chinese: a domain ontology modeling method. *New Review of Hypermedia and Multimedia*, 19(1):3–24, 2013.
- [Yokoi & Suematsu 94]
T. Yokoi and H. Suematsu. The future of the edr electronic dictionary. In *Proceedings of the International Post-Coling Workshop on Lexical Research*, Beijing, 1994.
- [Yokoi 95]
T. Yokoi. The EDR Electronic Dictionary. *Communications of the ACM*, 38(11), November 1995.
- [Young *et al.* 99]
S. Young, D. Kershaw, J. Odell, D. Ollason, V. Valtchev, and P. Woodland. *The HTK Book (Version 2.2)*. Entropic Ltd., Cambridge, 1999. <ftp://ftp.entropic.com/pub/htk/>.
- [Yourdon 89]
E. Yourdon. *Modern Structured Analysis*. Prentice Hall, New York, 1989.
- [Yourdon 96]
E. Yourdon. *The Rise and Resurrection of the American Programmer*. Prentice Hall, New York, 1996.
- [Yu & Gish 93]
G. Yu and H. Gish. Identification of speakers engaged in dialog. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 2, pages 383–386, Minneapolis MN, 1993.
- [Yu *et al.* 90]
G. Yu, W. Russell, R. Schwartz, and J. Makhoul. Discriminant analysis and supervised vector quantization for continuous speech recognition. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 685–688, 1990.
- [Yu *et al.* 98]
S. Yu, S. Bai, and P. Wu. Description of the Kent Ridge Labs System used for MUC-7. In *Proceedings of the MUC-7 Conference*, 1998.
- [Yu *et al.* 07]
J. Yu, E. Reiter, J. Hunter, and C. Mellish. Choosing the content of textual summaries of large time-series data sets. *Natural Language Engineering*, 13(1):25–49, March 2007.
- [Yu *et al.* 09]
J. X. Yu, L. Qin, and L. Chang. Keyword search in databases. *Synthesis Lectures on Data Management*, 1(1):1–155, 2009.
- [Yue *et al.* 07]
Y. Yue, T. Finley, F. Radlinski, and T. Joachims. A support vector method for optimizing average precision. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '07, pages 271–278, New York, NY, USA, 2007. ACM.
- [Yuen *et al.* 09]
M.-C. Yuen, L.-J. Chen, and I. King. A survey of human computation systems. In *Computational Science and Engineering, 2009. CSE '09. International Conference on*, volume 4, pages 723 –728, aug. 2009.
- [Zagibalov & Carroll 08]
T. Zagibalov and J. Carroll. Unsupervised classification of sentiment and objectivity in chinese text. In *Proceedings of IJCNLP 2008*, Hyderabad, India, January 2008.

- [Zaidan & Callison-Burch 11]
 O. F. Zaidan and C. Callison-Burch. Crowdsourcing translation: professional quality from non-professionals. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 1220–1229, 2011.
- [Zajac 92]
 R. Zajac. Towards Computer-Aided Linguistic Engineering. In *Proceedings of COLING '92*, pages 828–834, Nantes, France, 1992.
- [Zajac 97]
 R. Zajac. An Open Distributed Architecture for Reuse and Integration of Heterogenous NLP Components. In *Proceedings of the 5th conference on Applied Natural Language Processing (ANLP-97)*, 1997.
- [Zajac 98a]
 R. Zajac. Feature Structures, Unification and Finite-State Transducers. In *International Workshop on Finite State Methods in Natural Language Processing*, Ankara, Turkey, 1998.
- [Zajac 98b]
 R. Zajac. Reuse and Integration of NLP Components in the Calypso Architecture. In *Workshop on Distributing and Accessing Linguistic Resources*, pages 34–40, Granada, Spain, 1998. <http://www.dcs.shef.ac.uk/~hamish/dalr/>.
- [Zajac et al. 97]
 R. Zajac, V. Mahesh, H. Pfeiffer, and M. Casper. The Corelli Document Processing Architecture. Technical report, Computing Research Lab, New Mexico State University, 1997.
- [Zanzotto et al. 11]
 F. Zanzotto, M. Pennacchiotti, and K. Tsioutsoulouliklis. Linguistic Redundancy in Twitter. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 659–669, Edinburgh, UK, 2011. Association for Computational Linguistics.
- [Zarri 92]
 G. Zarri. Semantic modeling of the content of (normative) natural language documents. In *Proceedings of the specialized Conference on Natural Language Processing and its applications - Avignon 92*, 1992.
- [Zavitsanos et al. 11]
 E. Zavitsanos, G. A. Vouros, and G. Paliouras. Classifying users and identifying user interests in folksonomies. In *Proceedings of the 2nd Workshop on Semantic Personalized Information Management: Retrieval and Recommendation*, 2011.
- [Zeng et al. 10]
 Y. Zeng, Y. Wang, Z. Huang, D. Damljanovic, N. Zhong, , and C. Wang. User interests: Definition, vocabulary, and utilization in unifying search and reasoning. In *Proceedings of the 2010 International Conference on Active Media Technology, Lecture Notes in Computer Science 6335*, pages 98–107. Springer, 2010.
- [Zeng et al. 16]
 L. Zeng, K. Starbird, and E. S. Spiro. # unconfirmed: Classifying rumor stance in crisis-related social media messages. In *Tenth International AAAI Conference on Web and Social Media*, 2016.
- [Zenz et al. 09]
 G. Zenz, X. Zhou, E. Minack, W. Siberski, and W. Nejdl. From keywords to semantic queries - incremental query construction on the semantic web. *Web Semantics*, 7(3):166–176, September 2009.

- [Zesch *et al.* 08]
T. Zesch, C. Müller, and I. Gurevych. Extracting Lexical Semantic Knowledge from Wikipedia and Wiktionary. In *Proceedings of the Conference on Language Resources and Evaluation (LREC)*, 2008.
- [Zhai 97]
C. Zhai. Exploiting context to identify lexical atoms - a statistical view of linguistic context. In *Proc. of International & Interdisciplinary Conference on Modifying and Using Context (CONTEXT-97)*, pages 119–129, Rio de Janeiro, Brazil, 1997.
- [Zhang & Koppaka 07]
P. Zhang and L. Koppaka. Semantics-based legal citation network. In *ICAAIL '07: Proceedings of the 11th international conference on Artificial intelligence and law*, pages 123–130, New York, NY, USA, 2007. ACM.
- [Zhang & Mani 03]
J. Zhang and I. Mani. KNN Approach to Unbalanced Data Distributions: A Case Study Involving Information Extraction. In *Proceedings of the ICML'2003 Workshop on Learning from Imbalanced Datasets*, 2003.
- [Zhang *et al.* 07]
L. Zhang, Q. Liu, J. Zhang, H. Wang, Y. Pan, and Y. Yu. Semplore: An ir approach to scalable hybrid query of semantic web data. In *The Semantic Web, 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007*, pages 652–665, 2007.
- [Zhang *et al.* 13]
Z. Zhang, A. L. Gentile, I. Augenstein, E. Blomqvist, and F. Ciravegna. Mining equivalent relations from linked data. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 289–293, Sofia, Bulgaria, August 2013. Association for Computational Linguistics.
- [Zhao & Rosson 09]
D. Zhao and M. B. Rosson. How and why people Twitter: the role that micro-blogging plays in informal communication at work. In *Proceedings of the ACM 2009 international conference on Supporting group work, GROUP '09*, pages 243–252, New York, NY, USA, 2009. ACM.
- [Zhao 96]
G. Zhao. *Analogical Translator: Experience-Guided Transfer in Machine Translation*. Unpublished PhD thesis, Dept. of Language Engineering, UMIST, Manchester, England, 1996.
- [Zhao *et al.* 11]
W. X. Zhao, J. Jiang, J. Weng, J. He, E. Lim, H. Yan, and X. Li. Comparing Twitter and traditional media using topic models. In *Proceedings of the 33rd European Conference on Advances in Information Retrieval (ECIR)*, pages 338–349, 2011.
- [Zhao *et al.* 15]
Z. Zhao, P. Resnick, and Q. Mei. Early detection of rumors in social media from enquiry posts. In *International World Wide Web Conference Committee (IW3C2)*, 2015.
- [Zheleva & Getoor 09]
E. Zheleva and L. Getoor. To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles. In *Proceedings of the 18th international conference on World wide web, WWW '09*, pages 531–540, New York, NY, USA, 2009. ACM.
- [Zheng *et al.* 13]
Z. Zheng, X. Si, F. Li, E. Chang, and X. Zhu. Entity Disambiguation with Freebase. In *Proceedings of the Conference on Web Intelligence (WI-IAT'13)*, 2013.

- [Zho04]
Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communications. *Group Decision and Negotiation*, 13(1), 2004.
- [Zhong *et al.* 02]
J. Zhong, H. Zhu, J. Li, and Y. Yu. Conceptual graph matching for semantic search. In U. Priss, D. Corbett, and G. Angelova, editors, *Conceptual Structures: Integration and Interfaces*, volume 2393 of *Lecture Notes in Computer Science*, pages 92–106. Springer Berlin / Heidelberg, 2002.
- [Zhou & Hovy 06]
L. Zhou and E. Hovy. On the summarization of dynamically introduced information: Online discussions and blogs. In *AAAI Symposium on Computational Approaches to Analysing Weblogs (AAAI-CAAW)*, pages 237–242, 2006.
- [Zhou & Su 02]
G. Zhou and J. Su. Named entity recognition using an hmm-based chunk tagger. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL'02)*, Philadelphia, PA, 2002.
- [Zhou *et al.* 05]
G. Zhou, J. Su, J. Zhang, and M. Zhang. Exploring Various Knowledge in Relation Extraction. In *Proceedings of the 43rd Annual Meeting of the ACL*, pages 427–434, 2005.
- [Zhou *et al.* 07a]
M. Zhou, S. Bao, X. Wu, and Y. Yu. An unsupervised model for exploring hierarchical semantics from social annotations. In *Proceedings of the 6th International Semantic Web Conference, ISWC'07*, pages 680–693. Springer-Verlag, 2007.
- [Zhou *et al.* 07b]
Q. Zhou, C. Wang, M. Xiong, H. Wang, and Y. Yu. Spark: adapting keyword query to semantic search. In *Proceedings of the 6th international The semantic web and 2nd Asian conference on Asian semantic web conference, ISWC'07/ASWC'07*, pages 694–707, Berlin, Heidelberg, 2007. Springer-Verlag.
- [Zhou *et al.* 10a]
B. Zhou, Y. Jia, C. Liu, and X. Zhang. A distributed text mining system for online web textual data analysis. In *Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), 2010 International Conference on*, pages 1–4, Los Alamitos, CA, USA, October 2010. IEEE Computer Society.
- [Zhou *et al.* 10b]
W. Zhou, Y. Lu, H. Li, Y. Song, and Q. Tian. Spatial coding for large scale partial-duplicate web image search. In *ACM Multimedia*, pages 511–520, 2010.
- [Zhuang *et al.* 06]
L. Zhuang, F. Jing, and X.-Y. Zhu. Movie review mining and summarization. In *CIKM '06: Proceedings of the 15th ACM international conference on Information and knowledge management*, pages 43–50, New York, NY, USA, 2006. ACM.
- [Zipf 35]
G. K. Zipf. *The Psycho-Biology of Language*. Houghton Mifflin, Boston, 1935.
- [Zolton-Ford 84]
E. Zolton-Ford. Reducing Variability in Natural-Language Interactions with Computers. In *Proceedings of the Human Factors Society 28th Annual Meeting*, pages 768–772. The Human Factors Society, 1984.

- [Zubiaga *et al.* 12]
A. Zubiaga, D. Spina, E. Amigó, and J. Gonzalo. Towards real-time summarization of scheduled events from twitter streams. *CoRR*, abs/1204.3731, 2012.
- [Zubiaga *et al.* 15]
A. Zubiaga, M. Liakata, R. Procter, K. Bontcheva, and P. Tolmie. Crowdsourcing the annotation of rumourous conversations in social media. In *WWW 2015 Companion*, 2015.
- [Zubiaga *et al.* 16a]
A. Zubiaga, M. Liakata, R. Procter, G. Wong Sak Hoi, and P. Tolmie. Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PLoS ONE*, 11(3):1–29, 03 2016.
- [Zubiaga *et al.* 16b]
A. Zubiaga, M. Liakata, R. Procter, G. Wong Sak Hoi, and P. Tolmie. Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PLoS ONE*, 11(3):1–29, 03 2016.
- [Zweigenbaum *et al.* 07]
P. Zweigenbaum, D. Demner-Fushman, H. Yu, and K. B. Cohen. Frontiers of biomedical text mining: current progress. *Briefings in Bioinformatics*, 8(5):358–375, 2007.
- [Zwicker 61]
E. Zwicker. Subdivision of the audible frequency range into critical bands (frequenzgruppen). *Journal of the Acoustical Society of America*, 33:248–249, 1961.
- [Zwicky & Pullum 83]
A. Zwicky and G. Pullum. Cliticization vs. inflection: English n't. *Language*, 59(3), 1983.