University of Groningen

Geographically constrained information retrieval

Andogah, Geoffrey

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2010

[Link to publication in University of Groningen/UMCG research database](#)

Bibliography

Harith Alani, Christopher B. Jones, and Douglas Tudhope. Voronoi-based region approximation for geographical information retrieval with gazetteers. *International Journal of Geographical Information Science*, 15(4):287–306, 2001.

Einat Amitay, Nadav Har'El, Ron Sivan, and Aya Soffer. Web-a-Where: Geotagging Web Content. In *Proceedings of SIGIR-04, the 27th Conference on Research and Development in Information Retrieval*, Sheffield, South Yorkshire, UK, 2004.

Geoffrey Andogah. GIR Experimentation. In *Evaluation of Multilingual and Multi-modal Information Retrieval*, volume 4730/2009 of *Lecture Notes in Computer Science LNCS*, pages 881–888. Springer-Verlag Berlin/Heidelberg, 2007.

Geoffrey Andogah and Goose Bouma. Relevance measures using geographic scopes and types. In *Advances in Multilingual and Multimodel Information Retrieval*, volume 5152/2008 of *Lecture Notes in Computer Science LNCS*, pages 794–801. Springer-Verlag Berlin/Heidelberg, 2008.

Geoffrey Andogah and Gosse Bouma. University of Groningen at GeoCLEF 2007. In *Working Notes for CLEF 2007*, Budapest, Hungary, 2007.

Geoffrey Andogah, Goose Bouma, John Nerbonne, and Elwin Koster. Geographical scope resolution. In *Proceedings of the LREC 2008 Workshop on Methodologies and Resources for Processing Spatial Language*, Marrakech, Morocco, 2008.

Leonardo Andrade and Mario J. Silva. Relevance Ranking for Geographic IR. In *Proceedings of the Workshop on Geographical Information Retrieval, SIGIR'06*, Seattle, USA, August 2006.

Kate Beard and Vyjayanti Sharma. Multidimensional ranking for data in digital spatial libraries. *International Journal on Digital Libraries*, 1:153–160, 1997.

Davide Buscaldi and Paolo Rosso. The UPV at GeoCLEF 2007. In *Working Notes for CLEF 2007*, 2007.

Davide Buscaldi and Paolo Rosso. On the Relative Importance of Toponyms in GeoCLEF. In *Advances in Multilingual and Multimodel Information Retrieval*, volume 5152/2008 of *Lecture Notes in Computer Science LNCS*, pages 815–822. Springer-Verlag Berlin/Heidelberg, 2008.

Davide Buscaldi, Paolo Rosso, and Emilio Sanchis Arnal. Using the WordNet Ontology in the GeoCLEF Geographical Information Retrieval Task. In *Accessing Multilingual Information Repositories*, volume 4022/2006, pages 939–946. Springer (Lecture Notes in Computer Science LNCS), 2006.

Cláudio Elízio Calazans Campelo and Cláudio de Souza Baptista. Geographic scope modeling for web documents. In *Proceedings of Workshop on Geographic Information Retrieval (GIR'08)*, Napa Valley, California, USA, 2008.

Nuno Cardoso, David Cruz, Marcirio Chaves, and Mário J. Silva. The University of Lisbon at GeoCLEF 2007. In *Working Notes for CLEF 2007*, Budapest, Hungary, 2007.

Jean Carletta. Assessing agreement on classification tasks: The kappa statistic. In *Computational Linguistics*, volume 22 of *2*, pages 249–254, Cambridge, MA, USA, June 1996. MIT Press.

Xavier Carreras, Lluís Màrquez, and Lluís Padró. A Simple Named Entity Extractor using AdaBoost. In *Proceedings of CoNLL-2003*, Edmonton, Canada, 2003.

Hai Leong Chieu and Hwee Tou Ng. Named Entity Recognition with a Maximum Entropy Approach. In *Proceedings of CoNLL-2003*, Edmonton, Canada, 2003.

Nancy Chinchor. MUC-7 Named Entity Task Definition, 1997. 17 October 2007: `http://www.itl.nist.gov/iad/894.02/related_projects/muc/`.

Paul Clough. Extracting Metadata for Spatially-Aware Information Retrieval on the Internet. In *Proceedings of Workshop on Geographic Information Retrieval (GIR'05), CIKM2005*, Bremen, Germany, 2005.

James R. Curran and Stephen Clark. Language Independent NER using a Maximum Entropy Tagger. In *Proceedings of CoNLL-2003*, Edmonton, Canada, 2003.

Bruno Emanuel da Graca Martins. *Geographically Aware Web Text Mining*. PhD thesis, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, 2008.

Tiago M. Delboni, Karla A. V. Dorges, Alberto H. F. Laender, and Clodoveu A. Davis Jr. Semantic Expansion of Geographic Web Queries Based on Natural Language Positioning Expressions. In *Transactions in GIS*, pages 377–397, 2007.

Junyan Ding, Luis Gravano, and Narayanan Shivakumar. Computing Geographical Scopes of Web Resources. In *Proceedings of the 26th Very Large Data Bases (VLDB) Conference*, pages 545–556. Morgan Kaufmann Publishers Inc., 2000.

David Dubin. The most influential paper gerard salton never wrote. Technical report, Graduate School of Library and Information Science. University of Illinois at Urbana-Champaign., 2004.

Michael Ben Fleischman and Eduard Hovy. Multi-document person name resolution. In *ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (ACL), REFERENCE RESOLUTION WORKSHOP*, pages 66–82, 2004.

Radu Florian, Abe Ittycheriah, Hongyan Jing, and Tong Zhang. Named Entity Recognition through Classifier Combination. In *Proceedings of CoNLL-2003*, Edmonton, Canada, 2003.

Gaihua Fu, Christopher B. Jones, and Alia I. Abdelmonty. Ontology-based Spatial Query Expansion in Information Retrieval. In *On the Move to Meaningful Internet Systems 2005: CoopIS, DOA, and ODBASE*, volume 3761/2005, pages 1466–1482. Springer (Lecture Notes in Computer Science LNCS), 2005.

William A. Gale, Kenneth W. Chruch, and David Yarowsky. One Sense Per Discourse. In *Proceedings of the Fourth DARPA Speech and Natural Language Workshop*, pages 233–237, 1992.

Fredric Gey, Ray Larson, Mark Sanderson, Hideo Joho, Paul Clough, and Vivien Petras. GeoCLEF: the CLEF 2005 Cross-Language Geographic

Information Retrieval Track Overview. In *Cross-Language Evaluation Forum: CLEF 2005*. Springer (Lecture Notes in Computer Science LNCS 4022), 2006.

Fredric Gey, Ray Larson, Mark Sanderson, Kerstin Bischoff, Thomas Mandl, Christa Womser-Hacker, Diana Santos, Paulo Rocha, Giorgio M. Di Nunzio, and Nicola Ferro. GeoCLEF 2006: The CLEF 2006 Cross-Language Geographic Information Retrieval Track Overview. In *Evaluation of Multilingual and Multi-modal Information Retrieval*, volume 4730 of *Lecture Notes in Computer Science*, pages 852–876. Springer Berlin / Heidelberg, 2007.

Otis Gospodnetic and Eric Hatcher. *Lucene in Action*. Manning Publications Co., 206 Bruce Park Avenue, Greenwich, CT 06830, 2005.

Linda L. Hill, James Frew, and Qi Zheng. Geographic Names: The Implementation of a Gazetteer in a Georeferenced Digital Library. In *D-Lib Magazine*, volume 5 of *1*, January 1999.

Thorsten Joachims, Laura Granka, and Bing Pan. Accurately Interpreting Clickthrough Data as Implicit Feedback. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, Salvador, Brazil, 2005.

Christopher B. Jones and Ross Purves. GIR'05 2005 ACM workshop on geographical information retrieval. *ACM SIGIR Forum*, 40(1):34–37, 2006.

Christopher B. Jones, Harith Alani, and Douglas Tudhope. Geographic Information Retrieval with Ontologies of Place. In *Proceedings of the International Conference on Spatial Information Theory: Foundations of Geographic Information Science*, volume 2205/2001, pages 322–335. Springer (Lecture Notes in Computer Science LNCS), 2001.

Diane Kelly and Nicholas J. Belkin. Reading Time, Scrolling and Interaction: Exploring Implicit Sources of User Preferences for Relevance Feedback During Interactive Information Retrieval. In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*, 2001.

Diane Kelly and Jaime Teevan. Implicit Feedback for Inferring User Preference: A Bibliography. *ACM SIGIR Forum*, 37(2):18–28, 2003.

Dan Klein, Joseph Smarr, Huy Nguyen, and Christopher D. Manning. Named Entity Recognition with Character-Level Models. In *Proceedings of CoNLL-2003*, Edmonton, Canada, 2003.

Jon M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46:604–632, 1999.

Ray R. Larson and Patricia Frontiera. Spatial Ranking Methods for Geographic Information Retrieval (GIR) in Digital Libraries. In *Research and Advanced Technology for Digital Libraries*, volume 3232 of *Lecture Notes in Computer Science*, pages 45–56. Springer Berlin / Heidelberg, 2004.

Ray R. Larson, Fredric C. Gey, and Vivien Petras. Berkeley at GeoCLEF: Logistic Regression and Fusion for Geographic Information Retrieval. In *Accessing Multilingual Information Repositories*, volume 4022/2006, pages 963–976. Springer (Lecture Notes in Computer Science LNCS), 2006.

Jochen L. Leidner. Experiments with Geo-Filtering Predicates for Geographic IR. In *Accessing Multilingual Information Repositories*, volume 4022/2006, pages 987–996. Springer (Lecture Notes in Computer Science LNCS), 2006.

Jochen L. Leidner, Gail Sinclair, and Bonnie Webber. Grounding spatial named entities for information extraction and question answering. In *Kornai, A. and Sundheim, B. (eds) Proceedings of the HTL-NAACL 2003 Workshop on Analysis of Geographic References*, pages 31–38, Alberta, Canada, 2003.

Jochen Lothar Leidner. *Toponym Resolution in Text: Annotation, Evaluation and Applications of Spatial Grounding of Place Names*. PhD thesis, Institute for Communicating and Collaborative Systems, School of Informatics, University of Edinburgh, 2007.

Huifeng Li, Rohini K. Srihari, Cheng Niu, and Wei Li. InfoXtract location normalization: a hybrid approach to geographic references in information extraction. In *Kornai, A. and Sundheim, B. (eds) Proceedings of the HTL-NAACL 2003 Workshop on Analysis of Geographic References*, pages 39–44, Alberta, Canada, 2003.

Eugene E. Loos, Susan Anderson, Jr. Dwight H., Day, Paul C. Jordan, and J. Douglas Wingate. Glossary of linguistic terms, 2004. 5 October 2007: `http://www.sil.org/linguistics/`.

Thomas Mandl, Fredric Gey, Giorgio Di Nunzio, Nicola Ferro, Ray Larson, Mark Sanderson, Diana Santos, Christa Womser-Hacker, and Xing Xie. GeoCLEF 2007: the CLEF 2007 Cross-Language Geographic Information Retrieval Track Overview . In *Working Notes for CLEF 2007*, 2007.

Thomas Mandl, Fredric Gey, Giorgio Di Nunzio, Nicola Ferro, Ray Larson, Mark Sanderson, Diana Santos, Christa Womser-Hacker, and Xing Xie. GeoCLEF 2007: the CLEF 2007 Cross-Language Geographic Information Retrieval Track Overview . In *Advances in Multilingual and Multimodel Information Retrieval*, volume 5152/2008 of *Lecture Notes in Computer Science LNCS*, pages 674–686. Springer-Verlag Berlin/Heidelberg, 2008.

Thomas Mandl, Paula Carvalho, Giorgio Maria Di Nunzio, Fredric Gey, Ray R. Larson, Diana Santos, and Christa Womser-Hacker. GeoCLEF 2008: The CLEF 2008 Cross-Language Geographic Information Retrieval Track Overview . In *Evaluating Systems for Multilingual and Multimodal Information Access*, volume 5706/2009 of *Lecture Notes in Computer Science LNCS*, pages 808–821. Springer-Verlag Berlin/Heidelberg, 2009.

Gideon S. Mann and David Yarowsky. Unsupervised personal name disambiguation. In *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003*, pages 33–40. Association for Computational Linguistics, 2003.

Christopher D. Manning, Prabhakar Raghven, and Hinrich Schütze. *An Intoducation To Infromation Retrieval*. Cambridge University Press, Cambridge, England, November 2007. Draft.

Alexander Markowetz, Yen-Yu Chen, Torsten Suel, Xiaohui Long, and Bernhard Seefer. Design and Implementation of a Geographic Search Engine. In *Eighth International Workshop on the Web and Databases (WebDB 2005)*, 2005.

Bruno Martins and Mário J. Silva. A Graph-Ranking Algorithm for Geo-Referencing Documents. In *Proceedings of ICDM-05, the 5th IEEE International Conference on Data Mining*, Texas, USA, Novermber 2005.

Bruno Martins, Mário J. Silva, and Leonardo Andrade. Indexing and Ranking in Geo-IR Systems. In *Proceedings of the ACM Workshop on Geographic Information Retrieval*, pages 31–34. ACM New York, NY, USA, 2005.

Bruno Martins, Mário J. Silva, Sérgio Freitas, and Ana Paula Afonso. Handling Locations in Search Engine Queries. In *Proceedings of the 3rd Workshop on Geographic Information Retrieval held at The 29th Annual International ACM SIGIR Conference*, Seattle, WA, USA, 2006.

Bruno Martins, Nuno Cardoso, Marcirio Chaves, Leonardo Andrade, and Mário J. Silva. The University of Lisbon at GeoCLEF 2006. In *Evaluation of Multilingual and Multi-modal Information Retrieval, 7th Workshop of the Cross-Language Evaluation Forum, CLEF 2006, Alicante, Spain, September 20-22, 2006, Revised Selected Papers*, volume 4730/2007, pages 986–994. Springer (Lecture Notes in Computer Science LNCS), 2007.

Simon E Overell. *Geographic Information Retrieval: Classification, Disambiguation and Modelling*. PhD thesis, Department of Computing, Imperial College London, 2009.

Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The PageRank citation ranking: Bringing order to the Web, 1999. Technical Report SIDL-WP-1999-0120, Stanford Digital Library.

Bruno Pouliquen, Ralf Steinberger, Camelia Ignat, and Tom De Groeve. Geographical Information Recognition and Visualisation in Texts Written in Various Languages. In *Proceedings of ACM (SAC2004), Nicosia, Cyprus*, 2004.

Bruno Pouliquen, Marco Kimler, Ralf Steinberger, Camelia Ignat, Tamara Oellinger, Ken Blackler, Flavio Fluart, Wajdi Zaghouani, Anna Widiger, Ann-Charlotte Forslund, and Clive Best. Geocoding Multilingual Texts: Recognition, Disambiguation and Visualisation. In *Proceedings of The Fifth International Conference on Language Resources and Evaluation (LREC)*, pages 53–58, 2006.

Erik Rauch, Michael Bukatin, and Kenneth Baker. A Confidence-Based Framework for Disambiguating Geographic Terms. In *Kornai, A. and Sundheim, B. (eds) Proceedings of the HTL-NAACL 2003 Workshop on Analysis of Geographic References*, pages 50–54, Alberta, Canada, 2003.

Tony Rose, Mark Stevenson, and Miles Whitehead. Reuters Corpus Volume 1 - from yesterday?s news to tomorrow?s language resources. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC-2002)*, volume 3, pages 827–833, 2002.

Gerard Salton. *Automatic Text Processing: The Transformation Analysis and Retrieval of Information by Computer.* Addison-Wesley Series in Computer Science, 1989.

Gerard Salton and Chris Buckley. Improving retrieval performance by relevance feedback. In *Journal of the American Society for Information Science*, volume 41 of *4*, pages 288–297. John Wiley & Sons, Inc., 1990.

Mark Sanderson and Janet Kohler. Analyzing geographic queries. In *Proceedings of the Workshop on Geographic Information Retrieval held at The 27th Annual International ACM SIGIR Conference*, Sheffield, Engand, UK, 2004.

Christoph Schlieder, Thomas Vogele, and Ubbo Visser. Qualitative Spatial Representation for Information Retrieval by Gazetteers. In *Spatial Information Theory. Foundations of Geographic Information Science : International Conference, COSIT 2001 Morro Bay, CA, USA, September 19-23, 2001. Proceedings*, volume 2205/2001, pages 336–351. Springer (Lecture Notes in Computer Science LNCS), 2001.

David A. Smith and Gideon S. Mann. Bootstrapping Toponym Classifiers. In *Kornai, A. and Sundheim, B. (eds) Proceedings of the HTL-NAACL 2003 Workshop on Analysis of Geographic References*, pages 45–49, Alberta, Canada, 2003.

Amanda Spink, Dietmar Wolfram, Major B. J. Jansen, and Tefko Saracevic. Searching the web: The public and their queries. *Journal of the American Society for Information Science and Technology*, 52(2):226–234, 2001.

Beth M. Sundheim. Overview of the Fourth Message Understanding Evaluation and Conference. In *Proceedings of the Fourth Message Understanding Conference (MUC-4)*, pages 3–21. Morgan Kaufmann Publishers, 1992.

Erik F. Tjong Kim Sang and Fien De Meulder. Introduction to the CoNLL-2003 Shared Task: Language independent named entity recognition. In *Walter Daelemans and Miles Osborne, Editors, Proceedings of CoNLL-2003*, pages 142–147, 2003.

Subodh Vaid, Christopher B. Jones, Hideo Joho, and Mark Sanderson. Spatio-textual Indexing for Geographical Search on the Web. In *Advances in Spatial and Temporal Databases*, volume 3633 of *Lecture Notes in Computer Science*, pages 218–235. Springer Berlin / Heidelberg, 2005.

Marc van Kreveld and Iris Reinbacher. Good NEWS: partitioning a simple polygon by compass directions. In *Proceedings of the Nineteenth Annual Symposium on Computational geometry*, pages 78–87. ACM New York, NY, USA, 2003.

Cornelis Joost van Rijsbergen. *Information Retrieval*. Butterworths, 2nd edition, 1979. 7:112-140.

Raphael Volz, Joachim Kleb, and Wolfgang Mueller. Towards ontology-based disambiguation of geographical identifiers. In *Proceedings of WWW2007*, May 2007.

David R. F. Walker, Ian A. Newman, David J. Medyckyj-Scott, and Clive L. N. Ruggles. A system for identifying datasets for GIS users. *International Journal of Geographical Information Systems*, 6(6):511–527, 1992.

Jinxi Xu and W. Bruce Croft. Query expansion using local and global document analysis. In *ACM SIGIR International Conference on Research and Development in Information Retrieval*, pages 4–11, New York, NY, USA, 1996. ACM.

Tong Zhang and David Johnson. A Robust Risk Minimization based Named Entity Recognition System. In *Proceedings of CoNLL-2003*, Edmonton, Canada, 2003.

Wenbo Zong, Dan Wu, Aixin Sun, Ee-Peng Lim, and Dion Hoe-Lian Goh. On Assigning Place Names to Geography Related Web Pages. In *Proceedings of the Fifth ACM/IEEE-CS Joint Conference on Digital Libraries*, pages 354–362, 2005.