Jack G. Conrad

Curriculum Vitae

Capabilities

Lead research & development and consulting efforts for technology transfer from research into new products and internal operations. Develop and evaluate state-of-art algorithms for real-time environments. Research and implement technologies for search and entity-resolution-related applications in large operational environments (harnessing technologies such as data mining, information clustering, information retrieval, and information extraction). Foster collaboration, innovation and the leveraging of technology resources across a spectrum of corporate business units.

Skills

Significant Java programming experience on Linux platforms. Eclipse programming and debugging proficiency. Perl language expertise. Appreciable SQL experience with DB2 and Oracle databases and DBVisualizer toolkit. Data analysis based on WEKA Machine Learning Toolkit and CLUTO Clustering Toolkit. Python and R programming. Windows tools.

Fluent in English. Proficient in French.

Research Interests

information retrieval: legal IR, e-discovery, question answering, document deduplication

data clustering: legal document organization

machine learning: public record matching

sentiment analysis: legal blogs, product and service opinion mining

Education

MS   **University of Massachusetts – Amherst**   Computer Science, June 1994
           "A System for Discovering Relationships by Feature Extraction from Text Databases"

MA   **University of British Columbia – Vancouver**   English/Linguistics, June 1984

BS   **Marquette University**   Electrical Engineering, *Summa Cum Laude*, Dec. 1980
                            Minor: English

Research Experience

Lead Research Scientist
Thomson Reuters Global Resources, Catalyst Lab — August 2011 - Present
Baar, Canton Zug Switzerland

I am working within the newly established Catalyst Lab with Thomson Reuters Global Resources. The focus of this research group is three-fold: to foster collaboration, innovation and resource leveraging across the spectrum of the company's Professional (Legal, Tax & Accounting, Scientific) and Markets (Financial, News) business units. We educate Thomson Reuters' business units about the breadth of enabling technologies available through assets like Corporate R&D; we also receive proposals for new products and services from these business units and identify resources and expertise to assist in the design and implementa-

tion of technology infrastructure to support such proposals. There is also an on-going R&D function associated with this position.

Senior Research Scientist

Thomson Reuters, Research & Development — April 2008 - July 2011

Saint Paul, Minnesota USA

I worked for Corporate R&D within Thomson Reuters. I focused on a broad range of technical application areas involving artificial intelligence and textual data processing. The technologies I have applied to a spectrum of production-oriented challenges include information retrieval, data clustering, and machine learning. I lead consulting efforts for technology transfer from research into new products and internal operations. I also foster cross-group collaboration and communication in the process of implementing and deploying technology to meet business needs. A recent example of such responsibilities put into action was the role I played for R&D in the research and development of clustering technology for WestlawNext, a next generation search engine for the legal domain (www.westlawnext.com).

Research Scientist and Senior Research Scientist

The Thomson Corporation, Research & Development — May 1996 - April 2008

Saint Paul, Minnesota USA

• Researched and implemented resource discovery technology for Global Search applications in a massive operational environment consisting of over 20,000 databases that resulted in a ten-fold increase in usage of company's resource discovery tools.
• Developed and evaluated state-of-art algorithms for real-time identical and non-identical duplicate document detection applications harnessing rapid and compact digital signatures in a production repository managing over 60 million documents.
• Developed core technology for new online legal citator service generating $32 million a year and growing 30% annually.

Research Scientist

West Publishing Company, New Search Technologies — February 1994 - May 1996

Eagan, Minnesota USA

• Researched new optimization algorithms for high-precision WIN natural language search engine, scaled to operate in full-text environments of several terabytes.
• Conducted and directed research in support of new information retrieval, data mining, text filtering, and text classification technologies.

Research Assistant

Center for Intelligent Information Retrieval — June 1992 - December 1994

Amherst, Massachusetts USA

• Designed and implemented a proper noun text-based extraction and presentation system / GUI which was reported on at SIGIR, the international search and retrieval conference.
• Studied the effect of variable stop word lists on the INQUERY retrieval engine's performance using standard online test collections. Performed independent research on stemming and case recovery algorithms.

Teaching Experience

Teaching Assistant

Dept. of Computer Science, University of Massachusetts — June 1992 - December 1994

Amherst, Massachusetts USA

Taught Intro to Computers Course (solely responsible for two weekly labs).

Foreign Instructor
East China University of Science & Technology — September 1990 - August 1991
Shanghai, People's Republic of China

Trained graduate students and taught undergraduate courses in technical and business-related areas, sponsored by Department of Foreign Languages.

High School Math & Science Instructor
Lycée Kara and Ekukhanyeni High School — September 1984 - August 1989
Kara, Togo (West Africa) and Ekukhanyeni, Swaziland (Southern Africa)

Instructed High School and Lycée math classes, university-level calculus courses, and introductory science in preparation for strict national exams, while serving as a U.S. Peace Corps volunteer.

Publications
(Selected)

Journal Papers

Jack G. Conrad, E-Discovery Revisited: The Need for Artificial Intelligence Beyond Information Retrieval, *Journal of Artificial Intelligence and Law, Special Issue on E-Discovery,* 18(4), Springer, Netherlands, pp. 321-345, Dec. 2010.

Jack G. Conrad and Cindy P. Schriber, Managing Déjà Vu: Collection Building for Identifying Non-Identical Duplicate Documents, *Journal of the American Society for Information Science and Technology (JASIST)*, 57(7), John Wiley & Sons, Hoboken, NJ, pp. 919-930, 2006.

Jack G. Conrad and Joanne R.S. Claussen, Early User-System Interaction for Database Selection in Massive Domain-specific Online Environments, *Transactions on Information Systems (TOIS)*, 21(1), ACM Press, New York, pp. 94-131, 2003.

Conference Papers

Qiang Lu and Jack G. Conrad, Bringing Order to Legal Documents: An Issue-based Recommendation System via Cluster Association, In *Proceedings of the Fourth International Conference on Knowledge Engineering and Ontology Development (KEOD 2012)* (Barcelona, Spain), SciTePress DL, pp. 76-88, 2012.

Qiang Lu, Jack G. Conrad, Khalid Al-Kofahi, William Keenan, Legal Document Clustering with Built-In Topic Segmentation, In *Proceedings of the 2011 ACM-CIKM Twentieth International Conference on Information and Knowledge Management (CIKM 2011)* (Glasgow, Scotland), ACM Press, pp. 383-392, 2011.

Jack G. Conrad, Christopher Dozier, Hugo Molina-Salgado, Merine Thomas and Sriharsha Veeramachaneni, Public Record Aggregation Using Semi-supervised Entity Resolution, Public Record Aggregation Using Semi-supervised Entity Resolution, In *Proceedings of the 13th International Conference on Artificial Intelligence and Law, (ICAIL 2011) (*Pittsburgh, PA), ACM Press, pp. 239-248, 2011.

Jack G. Conrad, Jochen L. Leidner, Frank Schilder and Ravi Kondadadi, Query-based Opinion Summarization for Legal Blog Entries, In *Proceedings of the 12th International Conference on Artificial Intelligence and Law (ICAIL 2009)* (Barcelona, Spain), ACM Press, pp. 167-176, 2009.

Frank Schilder, Ravi Kondadadi, Jochen Leidner and Jack G. Conrad, Thomson Reuters at TAC 2008: Aggressive Filtering with FastSum for Update and Opinion Summarization, In *Proceedings of the First Text Analysis Conference (TAC08)* (Gaithersburg, Maryland), NIST, Washington, D.C., Nov. 2008.

Jack G. Conrad and Edward L. Raymond, Jr., Essential Deduplication Functions for Transactional Databases in Law Firms, In *Proceedings of the 11th International Conference on Artificial Intelligence and Law (ICAIL 2007)* (Stanford University, Palo Alto, CA), ACM Press, New York, pp. 231-236, 2007.

Jack G. Conrad and Frank Schilder, Opinion Mining in Legal Blogs, In *Proceedings of the 11th International Conference on Artificial Intelligence and Law (ICAIL 2007)* (Stanford University, Palo Alto, CA), ACM Press, New York, pp. 231-236, 2007.

Jack G. Conrad, Khalid Al-Kofahi, Ying Zhao and George Karypis, Effective Document Clustering for Large Heterogeneous Law Firm Collections, In *Proceedings of the 10th International Conference on Artificial Intelligence and Law (ICAIL 2005)* (Bologna, Italy), ACM Press, New York, pp. 177-187, 2005.

Jack G. Conrad, Changwen Yang and Joanne S. Claussen, Effective Collection Metasearch in a Hierarchical Environment: Global vs. Localized Retrieval Performance, In *Proceedings of the 25th Annual International ACM-SIGIR Conference on Research & Development in Information Retrieval (SIGIR 2002)* (Tampere, Finland), Springer-Verlag, London, pp. 371-372, 2002.

Jack G. Conrad, Xi S. Guo, and Cindy P. Schriber Online Duplicate Document Detection: Signature Reliability in a Dynamic Retrieval Environment, In *Proceedings of the 2003 ACM-CIKM Twelfth International Conference on Information and Knowledge Management (CIKM03)* (New Orleans, Louisiana), ACM Press, New York, pp. 243-252, 2003.

Jack G. Conrad and Joanne S. Claussen, Client-System Collaboration in Legal Corpus Selection in a Large Production Environment, In *Proceedings of 9th International Conference on Artificial Intelligence and Law (ICAIL 2003)* (Edinburgh, Scotland), ACM Press, pp. 262-273, 2003.

Jack G. Conrad, Changwen Yang, and Joanne S. Claussen, Global vs. Localized Search: A Comparison of Database Selection Methods in a Hierarchical Environment, In *Proceedings of the 2002 Annual Meeting of the American Society for Information Science and Technology (ASIST 2002)* (Washington, D.C.), Information Today, Medford, NJ, pp. 130-139, 2002.

Jack G. Conrad, Xi S. Guo, Peter Jackson, and Monem Meziou, Database Selection Using Complete Physical and Acquired Logical Collection Resources in a Massive Domain-Specific Operational Environment, In *Proceedings of the 28th International Conference on Very Large Databases (VLDB02)* (Hong Kong), pp. 71-82, 2002.

Jack G. Conrad and Daniel P. Dabney, Automatic Recognition of Distinguishing Negative Indirect History Language in Judicial Opinions, In *Proceedings of the 2001 ACM-CIKM Tenth International Conference on Information and Knowledge Management (CIKM 2001)* (Atlanta, Georgia), ACM Press, New York, pp. 287-294, 2001.

Jack G. Conrad and Daniel P. Dabney, A Cognitive Approach to Judicial Opinion Structure: Applying Domain Expertise to Component Analysis, In *Proceedings of 8th International Conference on Artificial Intelligence and Law (ICAIL 2001)* (St. Louis, Missouri), ACM Press, New York, pp. 1-11, 2001.

Jack G. Conrad and Mary H. Utt, A System for Discovering Relationships by Feature Extraction from Text Databases, In W.B. Croft and C.J. van Rijsbergen (Eds.), In *Proceedings of the 17th Annual International ACM-SIGIR Conference on Research & Development in Information Retrieval (SIGIR 1994)* (Dublin, Ireland), Springer-Verlag, London, pp. 260-270, 1994.

Workshop Papers

Jack G. Conrad, Jochen L. Leidner, Frank Schilder, Professional Credibility: Authority on the Web, *Proceedings of the Second Workshop on Information Credibility on the Web (CIKM08 Credibility Workshop -- WICOW08)* (Napa Valley, CA), ACM Press, pp. 85-88, 2008.

Jack G. Conrad, Tonya Custis, Christopher Dozier, Terry Heinze, Marc Light, Sriharsha Veeramachaneni, Linking, Mapping, and Clustering Entity Records in Information-based Solutions for Business and Professional Customers, In *Proceedings of the Workshop on Resources and Evaluation for Identity Matching, Entity Resolution and Entity Management at the Sixth International Conference on Language Resources and Evaluation (LREC 2008 -- Linking Workshop),* (Marrakech, Morocco), ELRA Press, 8 pgs, 2008.

Patents & Patent Applications

Qiang Lu, Jack G. Conrad, Michael Dahn, William M. Keenan, Systems, Methods and Interfaces for Extending Legal Search Results, Thomson Reuters, Research & Development, U.S. Patent App. Number 2010332520, December 30, 2010.

Jack G. Conrad, Joanne C. Claussen, Jie Lin, Information Retrieval Systems with Duplicate Document Detection and Presentation Functions, Thomson Reuters, Research & Development, U.S. Patent Number 7,809,695, October 5, 2010.

Robert Haschart, Jack G. Conrad, Forrest Rhoads, Dan Dabney, Howard Turtle, Kinley Larntz, Po Yang, and Dan Gannon, System and Method of Processing Formatted Text Documents in a Database, Thomson Reuters, Research & Development, U.S. Patent Number 7,529,756, May 5, 2009.

Jack G. Conrad, Christopher C. Dozier, Sriharsha Veeramachaneni, Systems, Methods and Software for Entity Relationship Resolution, Thomson Reuters, Research & Development, U.S. Patent App. Number 20090198678, August 6, 2009.

Committees & Service

Executive Committee, International Association for Artificial Intelligence and Law (IAAIL)

Editorial Board, Journal of Artificial Intelligence and Law

Information Retrieval Track Chairperson on Program Committee and Mentoring Program Co-Coordinator, International Conference on Artificial Intelligence and Law (ICAIL)

Co-organizer, DESI III Workshop on Global E-Discovery/E-Disclosure, Third in a series of workshops on "Supporting Search and Sensemaking for Electronically Stored Information in Discovery Proceedings," co-located with ICAIL 2009, Barcelona, Spain

Program Committees for SIGIR, CIKM, ICAIL and JURIX conferences

Awards

Thomson Reuters Community Champion Award, 1st Place (2009), in support of the Young Heroes Foundation of Swaziland.

EarthShares Champion of the Year (2000), for work to expand choice in company's Employee Giving Campaign.

Sargent Shriver Peace Worker of the Year, one of two finalists (1990), for secondary school teaching and rural water project successes while serving in the U.S. Peace Corps in southern and West Africa.

W.D. Bliss Cooperative Education Student of the Year, while an undergraduate at Marquette University interning with the 3M Company.

Honors / Societies

Eta Kappa Nu – National Electrical and Computer Engineering Honor Society (Past VP, Marquette Chapter)

Tau Beta Pi – National Engineering Honor Society

Extracurricular Boards

Friends of Swaziland, Communications Chair

Minnesota Returned Peace Corps Volunteers, Past President, authored successful proposal to host national conference in Saint Paul, MN (1999).

Extracurricular Activities

Cycled through 6 European countries in 6 days (750 kms / 465 miles) to raise funds for Young Heroes of Swaziland, an organization supporting AIDS orphans

Climbed Mount Kilimanjaro to Uhuru Peak (5,895m / 19,341 ft), Tanzania, East Africa

Ran the Chicago marathon, in under 3¾ hours

Contact

Jack G. Conrad

Lead Research Scientist
Catalyst Lab
Corporate Technology
Thomson Reuters Global Resources
Neuhofstrasse 1
6340 Baar, Switzerland
T (41) 41 709 04 44
F (41) 41 709 04 39

jack.conrad@acm.org
www.conradweb.org/~jackg/

References

Professor W. Bruce Croft

Distinguished University Professor
Department of Computer Science
University of Massachusetts
140 Governor's Drive
Amherst, MA 01003

T (413) 545-3240
F (413) 545-1789

croft@cs.umass.edu

Dr. Kevin Ashley

Professor of Law
Graduate Program in Intelligent Systems
University of Pittsburgh
3900 Forbes Avenue
Pittsburgh, PA 5260

T (412) 624-7496
F (412) 624-9249

ashley@pitt.edu

Dr. Khalid Al-Kofahi

Vice President
Research & Development
Thomson Reuters
610 Opperman Drive
Saint Paul, MN 55123

T (651) 687-3665
F (651) 848-3483

khalid.al-kofahi@thomsonreuters.com