

## Resumé of Steven J. DeRose, Ph.D. (November, 2015)

1908 Wallace Ave., Silver Spring MD 20902  
301-318-3233 (c), 301-949-6544 (f), sderose@acm.org

### Executive Summary

World-class domain expert in scalable text systems. 25+ years experience in natural language processing, text analysis, and standards. Innovator focused on functionality, maintainability, and usability. Effective communicator across technical levels in person, on stage, and in print.

### Selected Core Competencies

Computational linguistics

- Natural language processing, lexicology, corpus and text analytics.
- Formal semantics, knowledge representation, and ontology.

Enterprise document systems

- Structured and semi-structured documents, hypertext, and retrieval.
- Digital libraries, metadata, and scalable document systems.

Software and standards

- Prefer Python, XSLT, C, Java, and \*nix for development.
- Extensive standards work, often as chair or editor: XML, Xpath, Xlink, Open eBook, OSIS, TEI, HyTime, Encoded Archival Description, NLM/NCBI JATS.

### Employment History

2014-2015 *Computational Linguist*, Codeq LLC

- Developed corpora and annotation systems for email and Twitter analytics, clustering, lexicology, summarization, and feature extraction.
- Strengthened engineering culture, use of factoring, versioning, unit testing,....

2009-2013 *Director of Linguistics*, then *Director of R&D*, OpenAmplify LLC

- Supervised development of high-volume text analytic tools. Feature extraction, social media monitoring, sentiment, task extraction, parsing, ontology,....
- Developed corpora, meta-languages, models, and data-management tools.
- Some work with Amazon s3 for Wikipedia extraction; Gate, NLTK, R.

2005-2008 *XML Specialist*, National Center for Biotechnology Information, NIH

1998-2003 *Member, Technical Advisory Board*, ContextMedia Corp.

1998-2001 *Chief Scientist*, Brown University Scholarly Technology Group;

*Adjunct Associate Professor*, Department of Computer Science

1996-1999 *Chief Scientist*, Inso Corporation Electronic Publishing Systems

1989-96 *Senior System Architect, co-founder, and board member*,

Electronic Book Technologies

1987-89 *Design Consultant*, CDWord Project

### Education

1982 Sc.B. Computer Science / A.B. Linguistics (honors)

1986 A.M. Computational Linguistics

1989 Ph.D. Computational Linguistics. *Dissertation*: Dynamic programming and Hidden Markov Models for part-of-speech in English and Greek.

### Grants, Honors, and Professional Activities

Eleven US Patents on document representation, IR, and rendering.

Principal Investigator or co-PI on multiple grants from NEH and NSF.

Keynote speaker at the ACM Conference on Very Large Databases 1999, Edinburgh.

Seybold Editors' "Achievement of the Year" and "Technology Trailblazer" Awards.

Bentley Historical Library Research Fellowship. SAA C.F.W. Coker prize.

Senior Member, ACM. Past vice-chair of SigWeb. Member, IEEE.

Technical Advisory Boards: ccel.org, ContextMedia, Velocity Equity Partners.

### Publications (see <http://scholar.google.com/citations?user=bjwfKeMAAAAJ&hl=en>)

Two books, 40+ articles. Impact rating: h-index 24, i10-index 48.

## Selected Publications

- 2014 S. J. DeRose. "What do we still lack? Or: Prolegomena to any future hypertext system." In *Symposium on HTML5 and XML, Washington, DC, August 4*.
- 2014 S. J. DeRose. "JSOX: A Justly Simple Objectization for XML: Or: How to do better with Python and XML." In *Proceedings of Balisage 2014: The Markup Conference, Washington, DC, August 5-8*.
- 2012 S. J. DeRose. "The structure of content." *International Symposium on Quality Assurance and Quality Control in XML, Montréal, Canada, August 6*.
- 2004 S. J. DeRose. "Markup Overlap: A Review and a Horse." In *Proceedings of Extreme Markup Languages, Montréal*.
- 1999 S. J. DeRose and A. van Dam. "Document structure and markup in the FRESS hypertext system." *Markup Languages* 1(1): 7-32.
- 1999 S. J. DeRose. "XML linking." In *ACM Computing Surveys* 31(4).
- 1997 S. J. DeRose. "Navigation, Access, and Control Using Structured Data." In *American Archivist* 60(33): 298-309.
- 1994 S. J. DeRose and D. G. Durand. *Making Hypermedia Work: A Users' Guide to HyTime*. Boston: Kluwer Academic.
- 1992 L. R. Reynolds and S. J. DeRose. "Electronic books." *Byte* 17(6), June.
- 1990 S. J. DeRose, D. G. Durand, E. Mylonas, and A. H. Renear. "What is Text, Really?" *Journal of Computing in Higher Education* 1(2): 3-26. Reprinted with commentaries as *Journal of Computer Documentation*, 21(3), 1997.
- 1988 S. J. DeRose. "Grammatical Category Disambiguation by Statistical Optimization." *Computational Linguistics* 14(1): 31-39.
- 1987 J. H. Coombs, A. H. Renear, and S. J. DeRose. "Markup Systems and the Future of Scholarly Text Processing." *Communications of the Association for Computing Machinery* 30(11): 933-947.