

Dr. Stefan Siersdorfer

Curriculum Vitae

Hannover, Germany
✉ siersdorfer@outlook.de
🌐 www.l3s.de/~siersdorfer



Employment and Education

- 2009 – 2015 Senior Researcher and Project Leader, L3S Research Center, Hannover, Germany
 - 2012 Visiting Researcher, Yahoo! Research Labs, Barcelona, Spain (March – May 2012)
- 2007 – 2009 EU Marie Curie Research Fellow, Information Retrieval Group, University of Sheffield, UK
- 2005 – 2006 Postdoctoral Researcher, Database and Information Systems Group, Max-Planck Institute for Computer Science (MPII), Germany
- 2002 – 2005 PhD candidate, Database and Information Systems Group, Max-Planck Institute for Computer Science, Saarbrücken, Germany
DFG funded graduate program "Quality Guarantees for Computer Systems"
2005: PhD in Computer Science
- 1999 – 2002 University of the Saarland, Germany
 - 2000: Bachelors Degree in Computer Science (German "Vordiplom")
 - 2002: Masters Degree in Computer Science (German "Diplom")
- 1996 – 1999 University of the Saarland, Germany
 - 1998: Bachelors Degree in Physics (German "Vordiplom")

Research Interests

- Social Web Mining and Analysis
- Sentiment Classification and Analysis
- Multimedia Retrieval
- Text Mining and Retrieval
- Ensemble-based and Distributed Machine Learning
- Classification and Clustering
- Crawling and Analysis of Web Data

Research Statement

The long-term objective of my research is to better understand how to incorporate, adapt, and tune state-of-the-art machine learning and information retrieval methods into more intelligent next-generation systems for data mining and search on the Web.

Machine Learning and Web Mining: I investigate the engineering and in particular tuning issues of using classification and clustering algorithms for automatic data organization. I developed a constructive and efficient methodology for tuning a repertoire of classifiers and clustering methods to an application's specific goals. I designed, implemented and tested restrictive ensemble based learning techniques for automatic document organization, developed an evaluation methodology for this framework, and performed large-scale experiments on numerous test collections. In collaborative environments and in the context of personal information management, I have shown how multiple independently learned models from several users can be used to construct an advanced ensemble based decision model by taking the knowledge of multiple users into account in a decentralized manner, e.g. in a peer-to-peer overlay network. Many of the described concepts are embedded in BINGO!, a thematic Web information system based on focused crawling that uses classification methods for filtering of fetched documents and identifying relevant Web sources for further downloads. It comprises features such as refinements of the topic structure, document filtering and topic-specific search.

Data Analysis and Mining in Social Web Environments: Social Web applications like Flickr, YouTube, or Twitter for sharing multimedia content and information about events have become more and more popular. The rapidly increasing size of online communities and the huge amount of shared data makes discovering relevant content and finding related users a difficult task. My research in this area focuses on the application of machine learning and data mining techniques to the design of scalable and reliable assistance methods that individually guide particular users through large-scale multi-dimensional Social Web data to locate promising information. I have designed and tested social recommender systems that predict the utility of items, users, or groups based on the social environment of a given user. Using similar methods, I revealed new links between videos in a popular sharing platform using robust content-based video analysis techniques and exploited them for generating new tag assignments leading to enhanced feature representations for applications such as automatic data organization and search. I also conducted large-scale studies on commenting and comment rating behavior on the Social Web and analyzed dependencies between comments, views, comment ratings and topic categories, as well as sentiment expressed in comments.

Reviewer

- PC Member of the International World Wide Web Conference (WWW 2012, 2013, 2014)
- PC Member of the ACM International Conference on Web Search and Data Mining (WSDM 2011, 2015)
- PC Member of the International Conference on Research and Development in Information Retrieval (SIGIR 2008)
- PC Member of ACM International Conference on Information and Knowledge Management (CIKM 2013)
- PC Member of International Conference on Weblogs and Social Media (ICWSM 2014)
- PC Member of the European Conference on Information Retrieval (ECIR 2007, 2008, 2009, 2010)
- PC Member of the European Conference on Machine Learning (ECML 2009)
- PC Member of the IEEE/WIC/ACM International Conference on Web Intelligence (WI 2010)
- PC Member of the ACM Conference on Hypertext and Hypermedia (HT 2010)
- Reviewer for IEEE Transactions on Knowledge and Data Engineering (TKDE 2006, 2007, 2008, 2009, 2015)
- Reviewer for IEEE Transactions on Big Data (2015)
- Reviewer for IEEE Intelligent Systems (2012)
- Reviewer for IEEE Transactions on Multimedia (2009, 2012, 2013, 2014)
- Reviewer for ACM Transactions on Information Systems (TOIS 2011)
- Reviewer for ACM Transactions on the Web (TWEB 2009, 2010, 2014, 2015)
- Reviewer for ACM Transactions on Interactive Intelligent Systems (TiiS 2012)
- Reviewer for Journal of the Association for Information Science and Technology (JASIST 2014)
- Reviewer for Elsevier Journal on Web Semantics (JWS 2014)
- Reviewer for Elsevier Journal on Data and Knowledge Engineering (DKE 2009, 2010)
- Reviewer for Elsevier Journal on Information Processing and Management (IPM 2011)

Teaching

- Teaching Assistant of the course “Selected Topics in Web IR and Mining” (MPII)
- Supervisor at the seminar “Distributed Systems” (MPII)
- Supervisor at the seminar “Organizing and Searching Information with XML” (MPII)
- Guest Lecturer in the course “Information Storage and Retrieval Research” (University of Sheffield)
- Lecturer of the course “Social Network Analysis” (L3S and Leibnitz University Hannover)

Supervised Theses

- Bachelor Thesis “Retraining of Text Classifiers and Meta Learning” (MPII)
- Master Thesis “Advanced Linguistic Models for Document Classification” (MPII)
- Master Thesis “Using Ontologies for Multilingual Classification” (MPII)
- Master Thesis “Linking Historical German Dictionaries and Encyclopedias” (MPII)
- Master Thesis “Reward Mechanisms for Crowdsourcing” (L3S)
- Master Thesis “Crowdsourcing based Fashion Trend Detection in Social Networks” (L3S)
- Master Thesis “Scalable Web Search Based Techniques for Efficient Mining of Large Social Graphs” (L3S)
- Mentor of PhD Thesis on topic “Feedback and Comments on the Social Web” (L3S)
- Mentor of PhD Thesis on topic “Privacy Aware Information Retrieval in collaborative environments” (L3S)

- Supervision of other PhD students in the EU Projects LivingKnowledge und CUBRIK (L3S)

Event Organization

- Organizer of International Workshop on Collaborative Knowledge Management for Web Information Systems at WISE 2007
- Organizer of International Workshop on Recommendation and Collaboration at IUI 2008
- Organizer of International Workshop on DETecting and Exploiting Cultural diversiTy on the Social Web (DETECT) at CIKM 2011
- Symposium Chair of European Summer School on Information Retrieval (ESSIR 2011)

Projects

- EU Marie Curie Project Memoir ("Digital Tools for Managing and Sharing Memories"), 2006 – 2008, Role: Marie Curie Research Fellow
- EU FET Project LivingKnowledge ("Facts, Opinions and Bias in Time"), 2009 – 2012, Role: Project Group Leader
- EU IP Project CUBRIK ("Developing Human-Enhanced Time-Aware Multimedia Search"), 2011 – 2014, Role: Project Group Leader
- EU IP Project Glocal ("Event-Based Retrieval of Networked Media"), 2010 – 2013, Role: Co-Author of Proposal
- ERC Advanced Grant ALEXANDRIA ("Foundations for Temporal Retrieval, Exploration and Analytics in Web Archives"), 2014 – 2019, Role: Contributions to Proposal, responsible for Workpackage "Aggregating Social Networks and Streams"

Selected Full Paper Publications

Markus Rokicki, Sergej Zerr, Stefan Siersdorfer *Just in Time: Controlling Temporal Performance in Crowdsourcing Competitions*, 25th International World Wide Web Conference, WWW 2016, Montreal, Canada, 2016

Stefan Siersdorfer, Hanno Ackermann, Philipp Kemkes, Sergej Zerr *Who with Whom and How? - Extracting Large Social Networks using Search Engines*, 24th ACM Conference on Information and Knowledge Management (CIKM), Melbourne, Australia, 2015

Markus Rokicki, Sergej Zerr, Stefan Siersdorfer *Groupsourcing: Team Competition Designs for Crowdsourcing*, 24th International World Wide Web Conference, WWW 2015, Florence, Italy

Markus Rokicki, Sergiu Chelaru, Sergej Zerr, Stefan Siersdorfer *Competitive Game Designs for Improving the Cost Effectiveness of Crowdsourcing*, 23rd ACM Conference on Information and Knowledge Management (CIKM), Shanghai, China, 2014

Stefan Siersdorfer, Sergiu Chelaru, Jose San Pedro, Sengor Altingovde, Wolfgang Nejdl, *Analyzing and Mining Comments and Comment Ratings on the Social Web*, ACM Transactions on the Web (TWEB), 2014

Odysseas Papapetrou, Wolf Siberski, Stefan Siersdorfer, *Efficient Model Sharing for Scalable*

Collaborative Classification, Peer-to-Peer Networking and Applications (PPNA), Springer, 2014

Sergiu Chelaru, Sengör Altingövde, Stefan Siersdorfer, Wolfgang Nejdl *Analyzing, Detecting and Exploiting Sentiment in Web Queries*, ACM Transactions on the Web (TWEB), 2013

Fan Deng, Stefan Siersdorfer, Sergej Zerr, *Efficient Jaccard-based Diversity Analysis of Large Document Collections*, 21st ACM Conference on Information and Knowledge Management (CIKM), Maui, Hawaii, USA, 2012

Sergej Zerr, Stefan Siersdorfer, Jonathon Hare, Elena Demidova, *Privacy-Aware Image Classification and Search*, 35th ACM SIGIR Conference, Portland, Oregon, USA, 2012

Stefan Siersdorfer, Jose San Pedro, Mark Sanderson, *Content Redundancy in YouTube and its Application to Video Tagging*, ACM Transactions on Information Systems (TOIS), 2011

Paul Clough, Andrew Cox, Stefan Siersdorfer, *Developing Metrics to Characterize Flickr Groups*, Journal of the American Society for Information Science and Technology (JASIST), 2011

Stefan Siersdorfer, Jose San Pedro, Sergiu Chelaru, Wolfgang Nejdl, *How useful are your comments?- Analyzing and Predicting YouTube Comments and Comment Ratings*, 19th International World Wide Web Conference, WWW 2010, Raleigh, USA

Stefan Siersdorfer, Jose San Pedro, Mark Sanderson, *Automatic Video Tagging using Content Redundancy*, 32nd ACM SIGIR Conference, Boston, USA, 2009

Stefan Siersdorfer, Sergej Sizov, *Social Recommender Systems for Web 2.0 Folksonomies*, 20th ACM Conference on Hypertext and Hypermedia, Hypertext 2009, Torino, Italy

Stefan Siersdorfer, Jose San Pedro *Ranking and Classifying Attractiveness of Photos in Folksonomies*, 18th International World Wide Web Conference, WWW 2009, Madrid, Spain

Stefan Siersdorfer, Sergej Sizov, *Meta Methods for Model Sharing in Personal Information Systems*, ACM Transactions on Information Systems (TOIS), 2008.

Stefan Siersdorfer, Sergej Sizov, *Restrictive Clustering and Metaclustering for self-organizing Document Collections*, 27th Annual International ACM Conference on Research and Development in Information Retrieval (SIGIR), Sheffield, UK, 2004.

Stefan Siersdorfer, Sergej Sizov, Gerhard Weikum, *Goal-oriented Methods and Meta Methods for Document Classification and their Parameter Tuning*, 13th Conference on Information and Knowledge Management (CIKM), Washington D.C., USA, 2004.

Sergej Sizov, Michael Biber, Jens Graupmann, Stefan Siersdorfer, Martin Theobald, Gerhard Weikum, Patrick Zimmer, *The BINGO! System for Information Portal Generation and Expert Web Search*, 1st Semiannual Conference on Innovative Data Systems Research (CIDR), Asilomar, CA, USA, 2003.