



Avishek Anand

Assistant Professor

Research Interests and Profile

I work on problems relating to Information retrieval, NLP and Web Mining. I am currently an Assistant Professor in the Intelligent Systems department in Leibniz University, Hannover. I am recently interested in topic involving **interpretability of Information Retrieval systems** and **Scaling unsupervised representations of discrete structures**. I teach topics relating to information retrieval, big data and machine learning and also advise students at the Masters and PhD level.

Education

- 2009–2013 **Ph.D. Student**, Max Planck Institute for Informatics (MPI) and Saarland University, Saarbrücken, Germany.
Thesis Title: **Indexing Methods for Web Archives**
- 2007–2008 **M.Sc. in Computer Science**, MPII and Saarland University, Saarbrücken, Germany.
Graduated with Honors
- 2002–2005 **B.Tech. in Information Technology**, Indian Institute of Information technology, Allahabad, India.
Graduated with Honors

Professional Experience

- 2014–2017 **Post-doctoral researcher** and **Project Leader** in **L3S Research Center** headed by Prof. Wolfgang Nejdl.
- 2005–2007 **Microsoft Research and Development Center, India**
Software Engineer in the Internet Explorer 7 team, Windows Group.
- 2004–2005 **Lucent Technologies, India.**
Research Intern, Core Switching Division.

Publications (Selected Full Papers)

- 2019 *Avishek Anand, Megha Khosla, Jaspreet Singh, Jan-Hendrik Zab and Zijian Zhang. **Asynchronous Training of Word Embeddings for large text corpora**. In: *Proceedings of the Conference on Web Search and Data Mining (WSDM 2019)*.*

- Jaspreet Singh, Avishek Anand. **EXS: Explainable Search Using Local Model Agnostic Interpretability**. In: *Proceedings of the Conference on Web Search and Data Mining (WSDM 2019)*.
- Besnik Fetahu, Avishek Anand, Maria Koutraki. . **TableNet: An approach for determining fine-grained relations for Wikipedia tables**. In: *Proceedings of Conference on The World Wide Web. (The WebConf 2019)*.
- 2018 Jaspreet Singh, Avishek Anand. **Posthoc Interpretability of Learning to Rank Models using Secondary Training Data..** In: *Proceedings of Workshop on Explainable Recommendation and Search. 2018*.
- Helge Holzmann, Megha Khosla, Avishek Anand. **Delusive Page Rank of Incomplete Graphs**. In: *Proceedings of Complex Networks. 2018*.
- Avishek Anand Kilian Bizer, Alexander Erlei, Ujwal Gadiraju, Christian Heinze, Lukas Meub, Wolfgang Nejdl and Bjoern Steinroetter. **Effects of Algorithmic Decision-Making and Interpretability on Human Behavior: Experiments using Crowdsourcing**. In: *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing. (HCOMP 2018)*.
- 2017 Besnik Fetahu, Katja Markert, Wolfgang Nejdl, Avishek Anand. **Fine-Grained Citation Span Detection for References in Wikipedia**. . In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2017)*.
- Vinay Setty, Abhijit Anand, Arunav Mishra, Avishek Anand. **Modeling Event Importance for Ranking Daily News Events**. . In: *Proceedings of the Conference on Web Search and Data Mining (WSDM 2017)*.
- Helge Holzmann, Wolfgang Nejdl, **Avishek Anand**. *Exploring Web Archives Through Temporal Anchor Texts*. Published in: *Proceedings of WebScience 2017*.
- 2016 Besnik Fetahu, Katja Markert, Wolfgang Nejdl, Avishek Anand. **Finding News Citations for Wikipedia**. In: *Proceedings of the 19th ACM Conference on Information and Knowledge Management (CIKM 2016)*.
- Jaspreet Singh, Johannes Hoffart, Avishek Anand. **Discovering Entities with Just a Little Help from You**. In: *Proceedings of the 19th ACM Conference on Information and Knowledge Management (CIKM 2016)*.
- Helge Holzmann, Vinay Goel, Avishek Anand. **ArchiveSpark: Efficient Web Archive Access, Extraction and Derivation**. In: *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2016)*
- Helge Holzmann, Wolfgang Nejdl, Avishek Anand. **The Dawn of Today's Popular Domains: A Study of the Archived German Web over 18 Years**. In: *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2016)*
- Jaspreet Singh, Wolfgang Nejdl, Avishek Anand. **History by Diversity: Helping Historians Search News Archives**. In: *Proceedings of the ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR 2016)*.
- 2015 Besnik Fetahu, Katja Markert, Avishek Anand. **Automated News Suggestions for Populating Wikipedia Entity Pages**. In: *Proceedings of the 19th ACM Conference on Information and Knowledge Management (CIKM 2015)*.

Tuan Tran, Nattiya Kahnabua, Claudia Niederee, *Avishek Anand*. **Balancing Novelty and Salience: Adaptive Learning to Rank Entities for Timeline Summarization of High-impact Events**. In: *Proceedings of the 19th ACM Conference on Information and Knowledge Management (CIKM 2015)*.

Besnik Fetahu, Abhijit Anand, **Avishek Anand**. *How much is Wikipedia Lagging Behind News?*. Published in: *Proceedings of WebScience 2015, Oxford*.

Evica Milchevski, *Avishek Anand*, Sebastian Michel. **The Sweet Spot between Inverted Indices and Metric-Space Indexing for Top-K-List Similarity Search**. Published in: *Proceedings of the Extended Databases Technology (EDBT 2015)*.

Before 2014 Stephan Seufert, *Avishek Anand*, Srikanta Bedathur, Gerhard Weikum. **FERRARI: Flexible and Efficient Reachability Range Assignment for Graph Indexing**. To be published in: *International Conference on Data Engineering (ICDE 2013)*.

Avishek Anand, Srikanta Bedathur, Klaus Berberich, Ralf Schenkel. **Index Maintenance for Time-Travel Text Search**. Published in: *Conference on Research and Development in Information Retrieval (SIGIR 2012)*.

Avishek Anand, Srikanta Bedathur, Klaus Berberich, Ralf Schenkel. **Temporal Index Sharding for Space-Time Efficiency in Archive Search**. Published in: *Conference on Research and Development in Information Retrieval (SIGIR 2011)*.

Avishek Anand, Srikanta Bedathur, Klaus Berberich, Ralf Schenkel. **Efficient Temporal Keyword Queries over Versioned Text**. Published in: *Proceedings of the 19th ACM Conference on Information and Knowledge Management (CIKM 2010)*.

Avishek Anand, Srikanta Bedathur, Klaus Berberich, Ralf Schenkel, Christos Tryfonopoulos. **EverLast: A Distributed Architecture for Preserving the Web**. Published in: *Joint Conference for Digital Libraries (JCDL 2009)*.

Teaching Experience

- 2019 Graduate Lecture “**Deep Learning**”
Lecturer.
- 2018 Graduate Lecture “**Algorithms for Big Data**”
Lecturer.
- 2017-2018 Graduate Lecture “**Foundations of Probabilistic IR**”
Lecturer.
- 2016-2017, Graduate Lecture “**Temporal Information Retrieval**”
2015-2016 Lecturer.
- 2016 Graduate Lecture “**Large Scale Data Mining**”
Lecturer.
- 2015-2016 Undergraduate Lecture “**Puzzling problems in Computer Science**”
Lecturer.
- 2014 Graduate Lecture “**Web Science**”
Guest Lecturer.

PhD Students

- Current **Jaspreet Singh**
Thesis: "Search, Exploration and Interpretability of Temporal Text Collections"
- Zijian Zhang**
Thesis: "Post-hoc Interpretability of Learning Methods"
- Jurek Leonhardt**
Thesis: "Scalable Graph Representations"
- Graduated **Besnik Fetahu**
Thesis: "Automatic Wikipedia Recommendation and Population" (Co-advised with Prof. Wolfgang Nejdl)
- Helge Holzmann**
Thesis: "Large-scale Approaches for Web Archive analysis" (Co-advised with Prof. Wolfgang Nejdl)

Masters' Students Advised/Co-advised

- 2017 **Jurek Leonhardt** Thesis: "Deep Learning for Citation Detection". [currently a Phd student at L3S]
- 2017 **Nicolas Tempelmeir** Thesis: "Indexing Evolving Graphs". [currently a Phd student at L3S]
- 2014 **Jaspreet Singh** Thesis: "Diversification for searching News Archives". [currently a PhD student in L3S Research Center]
- 2011 **Yagiz Kargin** Thesis: "Distributed Analytics over Web Archives". [currently a PhD student in CWI Amsterdam]
- 2012 **Bibek Paudel** Master Thesis: "Redundancy Control in Web Archives". [currently a PhD student in University of Zurich]
- 2011-2012 **Sairam Gurajada** Research Topic: "Scaling Graph Clustering". [currently a Senior researcher in IBM Research.]

Professional Activities

- Organization **Dagstuhl Seminar '19** – Conversational Search
Committee
- Organization **WebScience 2016** – Hackathon Chair
Committee
- Organizer **Alexandria Workshop '16, '17, '18**
- PC Responsibilities **SIGIR '19, '18; WWW '19, 18, 17, WSDM '13, '17, '18 ; CIKM '18, '17 ; TKDE '12, '14, '15; TOIS '18**

Invited Talks/Tutorials

- 2018 **Split, Summer School** – "Scaling Text and Graph Representations"
ETH Zürich, Summer School – "Scalable Word Representations"
Weimar University, WEBIS group – "Interpretability of Rankers"

- KU Leuven** – “Temporal Text Indexing”
- 2017 **University of Glasgow, IR group** – “Wikipedia Enrichment”
- 2016 **SIGIR Conference (Tutorial)** – “Temporal Information Retrieval.”
- 2015 **University of Bolzano, Department of Database Systems** – “Temporal Indexing.”

Awards (selected)

- 2019 **Schufa Faculty Award.** Grant for investigating interpretability issues for Machine Learning Models. **Interpretability of Machine Learning models in Finance.**
- 2017-18 **Amazon Research Award.** Grant for investigating interpretability issues for Information Retrieval. **Interpretability of Neural Rankers.**
- 2016 **Best Paper Award Nomination in JCDL 2016.** Nominated for the Best Paper Award for the paper **ArchiveSpark: Efficient Web Archive Access, Extraction and Derivation.**
- 2016 **Hannover Impuls Ideenwettbewerb.** Runners up for Hochschul und Wissenschaftspreis for the commercialization potential of a research prototype.
- 2009 **International Max-Planck Research School (IMPRS) Ph.D. scholarship.** Awarded to pursue a Ph.D. program at Max Planck Institute for Informatics.
- 1999 **Indian National Mathematical Olympiad (INMO - 1999).** Selected to represent state of Orissa in INMO '99.
- Silver Medalist - Regional Mathematical Olympiad (RMO - 1999).** Silver Medalist and distinction for winning RMO three years in progression.