

XIAO QIN

Homepage: <http://web.cs.wpi.edu/~xqin>

Office: Fuller Laboratories 319, 100 Institute Road, Worcester, MA 01609

Email: xqin@cs.wpi.edu

RESEARCH INTEREST

My broad research interests include data mining, big data analytics and information management.

EDUCATION

- Ph.D. Candidate**, Computer Science 2013 – Present
Worcester Polytechnic Institute, Worcester, MA, USA.
Advisor: Dr. Elke A. Rundensteiner
Dissertation: “Sequential Data Mining and its Applications to Pharmacovigilance”
- Master of Science**, Information Sciences 2011
University of Pittsburgh, Pittsburgh, PA, USA.
Advisor: Dr. Daqing He
Concentration: Information Retrieval/Online Question & Answering Systems
- Bachelor of Engineer**, Computer Science & Technology 2009
Harbin Institute of Technology, Harbin, Heilongjiang, China.
Thesis Advisor: Dr. Zhendong Niu
Thesis: “A Rule-based, Statistical Model for Chinese Archeology Term Extraction”

PUBLICATION

- [EDBT’19] *Scalable Kernel Density Estimation-based Local Outlier Detection over Large Data Streams*. **Xiao Qin**, Lei Cao, Elke A. Rundensteiner and Samuel Madden.
- [IVAPP’19] *MedViz: Visual Analytics for Medication Error Detection*. Tabassum Kakar, **Xiao Qin**, Cory Tapply, Derek Murphy, Daniel Yun, Oliver Spring, Elke A. Rundensteiner, Lane Harrison, Thang La, Sanjay K. Sahoo and Suranjan De.
- [Drug Safety’19] *Adverse Drug Event Detection from Electronic Health Records Using Hierarchical Recurrent Neural Networks with Dual-Level Embeddings*. Susmitha Wunnava, **Xiao Qin**, Tabassum Kakar, Cansu Sen, Elke A. Rundensteiner and Xiangnan Kong.
- [URTC’18] *Drug-Drug Interaction Signal Detection from Drug Safety Reports*. Brian Zylich, Brian McCarthy, Andrew Schade, Huy Quoc Tran, **Xiao Qin**, Tabassum Kakar and Elke A. Rundensteiner.
- [CIKM’18 Demo] *DEVES: Interactive Signal Analytics for Drug Safety*. Tabassum Kakar, **Xiao Qin**, Susmitha Wunnava, Brian McCarthy, Andrew Schade, Huy Quoc Tran, Brian Zylich, Elke A. Rundensteiner, Lane Harrison, Sanjay K. Sahoo and Suranjan De.
- [AMIA’18 Poster] *Deep Learning Strategies for the Automatic Detection of Medication and Adverse Drug Events from Electronic Health Records*. Susmitha Wunnava, **Xiao Qin**, Tabassum Kakar, Elke A. Rundensteiner and Xiangnan Kong.
- [MADE’18 Workshop] *Bidirectional LSTM-CRF for Adverse Drug Event Tagging in Electronic Health Records*. Susmitha Wunnava, **Xiao Qin**, Tabassum Kakar, Elke A. Rundensteiner and Xiangnan Kong.
- [ICDE’18 Demo] *Multi-Drug Adverse Reactions Analytics*. **Xiao Qin**, Tabassum Kakar, Susmitha Wunnava, Brian McCarthy, Andrew Schade, Huy Quoc Tran, Brian Zylich, Elke A. Rundensteiner, Lane Harrison, Sanjay K. Sahoo and Suranjan De.
- [HEALTHINF ’18] *One Size Does Not Fit All: An Ensemble Approach Towards Information Extraction from Adverse Drug Event Narratives*. Susmitha Wunnava, **Xiao Qin**, Tabassum Kakar, Xiangnan Kong, Elke Rundensteiner, Sanjay Sahoo and Suranjan De.
- [KDD’17] *MARAS: Signaling Multi-Drug Adverse Reactions*. **Xiao Qin**, Tabassum Kakar, Susmitha Wunnava, Elke Rundensteiner and Lei Cao.

11. [SAC'17] *Towards Transforming FDA Adverse Event Narratives into Actionable Structured Data for Improved Pharmacovigilance*. Susmitha Wunnava, **Xiao Qin**, Tabassum Kakar, Vimig Socrates, Amber Wallace and Elke Rundensteiner.
12. [EDBT'16] *Interactive Temporal Association Analytics*. **Xiao Qin**, Ramoza Ahsan, Xika Lin, Elke Rundensteiner, and Matthew Ward.
13. [Yahoo! Tech Pulse'15] *Death Hoax Detection in Query Suggestions*. **Xiao Qin**, Zhongqiang Chen, Yuan Zhang and Shenhong Zhu.
14. [KDD'14 Workshop] *iPARAS: Incremental Construction of Parameter Space for Online Association Mining*. **Xiao Qin**, Ramoza Ahsan, Xika Lin, Elke Rundensteiner, and Matthew Ward.
15. [CIKM'12 Short] *RESQ: Rank-Energy Selective Query Forwarding for Distributed Search Systems*. Amin Teymorian, **Xiao Qin**, and Ophir Frieder.
16. [IITA'18] *Chinese Unknown Word Recognition Based on Functional Applications of Type Theory*. Dongping Gao, Zhendong Niu, Lening Lv, Peng Jiang, **Xiao Qin**, and Jiahong Guo.

TALK/POSTER

- *Topically-Coherent Neural Language Model Conditioned on Arbitrary Features*. New England Machine Learning Day (NEML). Microsoft Research New England, Cambridge, MA. May 7, 2018.
- *Kernel Density Estimation-based Local Outlier Detection over High Volume Data Streams*. New England Database Day (NEDBD). MIT, Cambridge, MA. Jan 19, 2018.
- *Density-based Outlier Detection over Data Stream*. Graduate Research Innovation Exchange (GRIE). WPI, Worcester, MA. Apr 11, 2017.
- *Text Mining From Drug Surveillance Report Narratives*. The 5th Annual Community Engagement and Research Symposium. Umass Medical School, Worcester, MA. Mar 25, 2016.
- *Towards Pharmacovigilance Using Machine Learning To Identify Unknown Adverse Reactions Triggered By Drug-Drug Interaction*. The 5th Annual Community Engagement and Research Symposium. Umass Medical School, Worcester, MA. March 25, 2016.
- *Interactive Temporal Association Analytics*. Graduate Research Innovation Exchange (GRIE). WPI, Worcester, MA. Apr 11, 2016.
- *EPSTAR: An Evolving Parameter Space Framework for Interactive Temporal Association Rule Mining*. New England Database Summit (NEDB). MIT, Cambridge, MA. Jan 30, 2015.
- *Evolving Parameter Space Framework for Interactive Temporal Association Mining*. Graduate Research Innovation Exchange (GRIE). WPI, Worcester, MA. Dec 10, 2014.
- *IncPARAS: An Incremental Parameter Space Load Pipeline for Interactive Association Mining*. Graduate Research Innovation Exchange (GRIE). WPI, Worcester, MA. Mar 19, 2014.
- *IncPARAS: An Incremental Parameter Space Load Pipeline for Interactive Association Mining*. New England Database Summit (NEDB). MIT, Cambridge, MA. Jan 31, 2014.

WORKING EXPERIENCE

U.S. Food and Drug Administration

ORISE Fellow

9/2016 – Now
Silver Spring, MD

- *Center for Drug Evaluation and Research (CDER)*
Supervisor: Suranjan De, Sanjay Sahoo and Dr. Thang La
- Developed natural language processing solutions for the FDA Adverse Event Reporting System (FAERS).

Worcester Polytechnic Institute

Research Assistant

8/2013 – Now
Worcester, MA

- *Database System Research Group (DSRG)*
Supervisor: Dr. Elke A. Rundensteiner
- <http://davis.wpi.edu:8180/DSRG/>

- *Data Visualization Lab (Xmdv)*
Supervisor: Dr. Matthew O. Ward

<http://davis.wpi.edu/xmdv/>

Worcester Polytechnic Institute
Teaching Assistant

8/2015 – 5/2016
Worcester, MA

- CS 2223: Algorithms
- CS 4432: Database Systems II
- CS 542: Database Management Systems

Yahoo! Search
Technical Intern III

5/2015 – 8/2015
Sunnyvale, CA

- *Search Assist Team*, Manager : Shenhong Zhu, Mentor: Dr. Zhongqiang Chen
- Developed data mining algorithms and pipelines for filtering query suggestions with controlled contents.

SERVICE & OTHER ACTIVITY

External Reviewer

- **KDD** International Conference on Knowledge Discovery and Data Mining. **2018**.
- **ICDE** International Conference on Data Engineering. **2016**.
- **VLDB** International Conference on Very Large Data Bases. **2015, 2018**.
- **SIGMOD** International Conference on Management of Data. **2015, 2017, 2018**.
- **EDBT** International Conference on Extending Database Technology. **2014, 2017**.

Conference Volunteer

- **KDD** International Conference on Knowledge Discovery and Data Mining. **2017**.
- **ICDE** International Conference on Data Engineering. **2012**.

AWARD

- 1st place in named entity recognition task in 2018 NLP challenges for detecting medication and adverse drug events from electronic health records (MADE1.0).
- HEALTHINF 2018 Best Student Paper Award Candidate.
- Oak Ridge Institute for Science and Education (**ORISE**) Fellowship. 2016 – 2017, 2017 – 2018.
- WPI Travel Award 2014, 2016, 2017, 2018.
- ACM SIGKDD Travel Award 2017.
- 3rd place in 2016 WPI Graduate Research and Innovation Exchange (GRIE) Poster Competition.
- Honorable mention in 2017 WPI GRIE Poster Competition.

TECHNICAL STRENGTH

Computer Languages
Databases
Tools

Python, Java, C++, Ruby, Perl, C and SQL.
Oracle, MySQL, Microsoft Access and SQL Server.
Pytorch, Hadoop, Pig, Oozie, Mahout, MatLab, Weka, Lingpipe,
OpenNLP, Mallet, Lucene, Lemur, GNUPlot and \LaTeX .