

# Xiaozhong Liu

Luddy Hall (700 N. Woodlawn Ave) 2120, Bloomington, IN 47408

🌐 <http://xiaozhong.website2.me> ✉ [liu237@indiana.edu](mailto:liu237@indiana.edu) 📞 812-855-2793

## APPOINTMENTS

---

- **Associate Professor (joint appointment of Computer Science and Data Science)**, Worcester Polytechnic Institute, *Aug 2021 – Present*
- **Associate Professor (joint appointment of Computer Science and Information and Library Science)**, School of Informatics, Computing, and Engineering, Indiana University Bloomington, *May 2017 – Present*
- **Senior Consultant and NLP Director (lead NLP and Legal AI projects)**, Alibaba DAMO Academy, *Dec 2017 – May 2021*
- **Associate Professor and DS Curriculum/Admission Committee**, Data Science Program, School of Informatics, Computing, and Engineering, Indiana University Bloomington, *May 2013 – June 2021*.
- **Information Science PhD Program Director**, Indiana University Bloomington, *May 2020 – June 2021*
- **Assistant Professor**, School of Informatics and Computing, Indiana University Bloomington, *Jan 2011 – May 2017*.
- **Affiliate Faculty**, Indiana University Network Science Institute, *Jan 2014 – Present*.
- **Affiliate Faculty**, Center for Complex Networks and Systems Research (CNetS), School of Informatics and Computing, Indiana University Bloomington, *July 2011 – Present*.
- **Affiliate Faculty**, Data 2 Insight Center, School of Informatics and Computing, Indiana University Bloomington, *Jul 2011 – Present*.
- **Research Intern**, Yahoo!, Sunnyvale, CA, *Jul 2008 – Dec 2008*.
- **Research Intern**, Yahoo!, Sunnyvale, CA, *Jul 2007 – Sep 2007*.

## EDUCATION

---

**Syracuse University**

*Ph.D. in Information Science and Technology*

**Syracuse, NY**

*Jul 2011*

**Dissertation title:** - Community Interest as An Indicator for Ranking.

**Advisor:** Elizabeth Liddy.

## RESEARCH HIGHLIGHT ([CSrankings.org](http://CSrankings.org) - AI No.1; CS No.2 at IU)

---

- Recognized as a high impact scholar focusing on data science, natural language processing (NLP), machine/deep learning, information retrieval, AI fairness/transparency, computational social science, and graph mining (rank No.2 for CS and No.1 for AI at IU based on [CSrankings.org](http://CSrankings.org)).
- Led large-scale AI and data science projects in academia and industry with cutting-edge technologies.
- Published more than 100 peer-reviewed articles in top venues, e.g., top information science journal articles as well as highly regarded conferences in computer science.

- Leadership experience in a top industry research lab and a number of large-scale machine learning projects (e.g., de-biased legal AI, sentiment analysis, chatbot, information recommendation/retrieval, and collaborative intelligence).
- Developed algorithm APIs that are used more than 150 million times per day and outperform existing methods by 10% on the average.
- Led data mining projects that analyzed massive heterogeneous data from 0.5 billion users and 150 billion transactions.
- Interdisciplinary scholar across Computer and Information Science.

## PEER REVIEWED JOURNAL ARTICLES

---

- [J19] **Healthcare** **Recipe Mining** **Deep Learning** **IEEE Access** Kequan Li, Yan Chen, Hongsong Li, Xuhong Zhang Xiangwei Mu, and Xiaozhong Liu. Deeprecipes: Exploring massive online recipes and recovering food ingredient amounts. *IEEE Access*, 2021
- [J18] **Cybercrime Detection** **Misinformation** **Deep Learning** **IPM (pending)** Guoxiu He, Zhe Gao, Yangyang Kang Zhuoren Jiang, and Xiaozhong Liu. Understanding the Implied Textual Meaning via A Skim and Intensive Reading Model (In Review). *Journal of Information Processing and Management*, 2020b
- [J17] **Question-Answering** **NLP** **Deep Learning** **Neurocomputing (Accept)** Yongzhen Wang, Kaisong Song, Lidong Bing, and Xiaozhong Liu. Harvest Shopping Advice: Neural Question Generation from Multiple Information Sources in E-Commerce. *Neurocomputing*, 2020c
- [J16] **Scholarly Mining** **NLP** **Deep Learning** **Scientometrics (Accept)** Shutian Ma, Chengzhi Zhang, and Xiaozhong Liu. A Literature Review of Citation Recommendation: From Textual Content to Enriched Context. *Scientometrics*, 2020
- [J15] **Data Science** **Regression** **Social Data Mining** **DIM (Accept)** Han Jia, Chun Guo, and Xiaozhong Liu. Smartphone and Tablet Application (App) Life Cycle Characterization via Apple App Store Rank. *Data and Information Management*, 2020
- [J14] **Graph-Mining** **Bioinformatics** **BMC bioinformatics** Zheng Gao, Gang Fu, Chunping Ouyang, Satoshi Tsutsui, Xiaozhong Liu, Jeremy Yang, Christopher Gessner, Brian Foote, David Wild, and Ying Ding. edge2vec: Representation learning using edge semantics for biomedical knowledge discovery. *BMC bioinformatics*, 20(1):306, 2019a
- [J13] **Data Science** **Career Recommendation** **Text-Mining** **PNAS** Katy Börner, Olga Scrivner, Mike Gallant, Shutian Ma, Xiaozhong Liu, Keith Chewning, Lingfei Wu, and James A Evans. Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy. *Proceedings of the National Academy of Sciences*, 115(50):12630–12637, 2018
- [J12] **Data Science** **Graph Mining** **PLoS One** Yongzhen Wang, Xiaozhong Liu, and Yan Chen. Analyzing cross-college course enrollments via contextual graph mining. *PloS one*, 12(11):e0188577, 2017
- [J11] **Bibliometrics** **Graph Mining** **JOI** Scott Jensen, Xiaozhong Liu, Yingying Yu, and Staša Milojevic. Generation of topic evolution trees from heterogeneous bibliographic networks. *Journal of informetrics*, 10(2):606–621, 2016
- [J10] **Bibliometrics** **Graph Mining** **JASIST** Zhuoren Jiang, Xiaozhong Liu, and Yan Chen. Recovering uncaptured citations in a scholarly network: A two-step citation analysis to estimate publication

- importance. *Journal of the Association for Information Science and Technology*, 67(7):1722–1735, 2016a
- [J9] **Social Media** **Metadata** **Machine Learning** **JASIST** Xiaozhong Liu, Chun Guo, and Lin Zhang. Scholar metadata and knowledge generation with human and artificial intelligence. *Journal of the Association for Information Science and Technology*, 65(6):1187–1201, 2014b
  - [J8] **Human Computing** **Machine Learning** **JASIST** Xiaozhong Liu and Jian Qin. An interactive metadata model for structural, descriptive, and referential representation of scholarly output. *Journal of the Association for Information Science and Technology*, 65(5):964–983, 2014
  - [J7] **Social Media** **Meta-search** **Cyberlearning** **JASIST** Xiaozhong Liu. Generating metadata for cyberlearning resources through information retrieval and meta-search. *Journal of the American Society for Information Science and Technology*, 64(4):771–786, 2013
  - [J6] **QA** **IR** **Cyberlearning** **JASIST** Xiaozhong Liu and Han Jia. Answering academic questions for education by recommending cyberlearning resources. *Journal of the American Society for Information Science and Technology*, 64(8):1707–1722, 2013
  - [J5] **Ranking** **IR** **Social Media** **JASIST** Xiaozhong Liu and Howard Turtle. Real-time user interest modeling for real-time ranking. *Journal of the American Society for Information Science and Technology*, 64(8):1557–1576, 2013
  - [J4] **Text-Mining** **IR** **JASIST** Xiaozhong Liu, Jinsong Zhang, and Chun Guo. Full-text citation analysis: A new method to enhance scholarly networks. *Journal of the American Society for Information Science and Technology*, 64(9):1852–1863, 2013
  - [J3] **Text-Mining** **Bibliometrics** **JASIST** Ying Ding, Xiaozhong Liu, Chun Guo, and Blaise Cronin. The distribution of references across texts: Some implications for citation analysis. *Journal of Informetrics*, 7(3):583–592, 2013
  - [J2] **Graph-Mining** **Bibliometrics** **Scientometrics** Lili Lin, Zhuoming Xu, Ying Ding, and Xiaozhong Liu. Finding topic-level experts in scholarly networks. *Scientometrics*, 97(3):797–819, 2013
  - [J1] **Policy** **info** Zixiang Tan, Hsiang Chen, and Xiaozhong Liu. Adoption of low mobility services: “Little-Smart” in China as a case. *info*, 8(6):69–78, 2006

## PEER REVIEWED CONFERENCE ARTICLES

---

- [C85] **Information Security** **NLP** **IEEE S&P2021** XiaoFeng Wang Dandan Xu Xiaozhong Liu Chang Yue Kai Chen Haixu Tang Baoxu Liu Yi Chen, Yepeng Yao. Bookworm Game: Automatic Discovery of LTE Vulnerabilities Through Documentation Analysis. In *Proceedings of the 42nd IEEE Symposium on Security and Privacy*. IEEE, 2021
- [C84] **NLG** **Chatbot** **CIKM2021** Changzhen Ji, Yating Zhang, Xiaozhong Liu, Changlong Sun Adam Jatowt, Conghui Zhu, and Tiejun Zhao. A Neural Conversation Generation Model via Equivalent Shared Memory Investigation. In *Proceedings of the 30th ACM International Conference on Information and Knowledge Management*. ACM, 2021
- [C83] **Legal Intelligence** **Information Summarization** **ACL2021** Xiyan Fu, Yating Zhang, Tianyi Wang, Xiaozhong Liu, Changlong Sun, and Zhenglu Yang. RepSum: Unsupervised Dialogue Summarization based on Replacement Strategy. In *Proceedings of the 59th Conference of the Association for Computational Linguistics*, 2021
- [C82] **Information Extraction** **NLP** **ACL2021** Fubang Zhao, Zhuoren Jiang, Yangyang Kang,

Changlong Sun, and Xiaozhong Liu. Adjacency List Oriented Relational Fact Extraction via Adaptive Multi-task Learning. In *Proceedings of the 59th Conference of the Association for Computational Linguistics*, 2021

- [C81] **Contract Mining** **NLP** **SIGIR2021** Zihan Wang, Hongye Song, Zhaochun Ren, Pengjie Ren, Zhumin Chen, Xiaozhong Liu, Hongsong Li, and Maarten de Rijke. Cross-Domain Contract Element Extraction with a Bi-directional Feedback Clause-Element Relation Network. In *Proceedings of the 44nd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2021
- [C80] **Legal Intelligence** **NLP** **SIGIR2021** Luyao Ma, Yating Zhang, Tianyi Wang, Xiaozhong Liu, Wei Ye, Changlong Sun, and Shikun Zhang. Legal Judgment Prediction with Multi-Stage Case Representation Learning in the Real Court Setting. In *Proceedings of the 44nd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2021
- [C79] **Cybercrime Detection** **NLP** **AAAI2021** Kaisong Song, Zhe Gao, Kangkang Yang, Changlong Sun, and Xiaozhong Liu. Evidence Aware Neural Pornographic Text Identification for Child Protection. In *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence*. AAAI, 2021
- [C78] **Collaborative Intelligence** **Chatbot** **AAAI2021** Jiawei Liu, Zhe Gao, Kangkang Yang, Zhuoren Jiang, Guoxiu He, Changlong Sun, Wei Lu, and Xiaozhong Liu. Time to Transfer: Predicting and Evaluating Machine-Human Chatting Handoff. In *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence*. AAAI, 2021
- [C77] **NLP** **Dialogue Mining** **AAAI2021** Yicheng Zou, Jun Lin, Lujun Zhao, Yangyang Kang, Zhuoren Jiang, Changlong Sun, Qi Zhang, Xuanjing Huang, and Xiaozhong Liu. Unsupervised Summarization for Chat Logs with Topic-Oriented Ranking and Context-Aware Auto-Encoders. In *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence*. AAAI, 2021a
- [C76] **NLP** **Dialogue Summarization** **AAAI2021** Yicheng Zou, Lujun Zhao, Yangyang Kang, Jun Lin, Minlong Peng, Zhuoren Jiang, Changlong Sun, Qi Zhang, Xuanjing Huang, and Xiaozhong Liu. Topic-Oriented Spoken Dialogue Summarization for Customer Service with Saliency-Aware Topic Modeling. In *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence*. AAAI, 2021b
- [C75] **Legal Intelligence** **Knowledge Graph** **EMNLP2020** Yiquan Wu, Kun Kuang, Yating Zhang, Xiaozhong Liu, Changlong Sun, Jun Xiao, Yueting Zhuang, Luo Si, and Fei Wu. De-biased Court's View Generation with Causality. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 2020
- [C74] **Content Generation** **Summarization** **EMNLP2020** Changzhen Ji, Xin Zhou, Yating Zhang, Xiaozhong Liu, Changlong Sun, Conghui Zhu, and Tiejun Zhao. Cross Copy Network for Dialogue Generation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 2020b
- [C73] **Bioinformatics** **PubMed Mining** **EMNLP2020** Qiao Jin, Chuanqi Tan, Mosha Chen, Xiaozhong Liu, and Songfang Huang. Predicting Clinical Trial Results by Implicit Evidence Integration. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 2020
- [C72] **Digital Library** **Graph Mining** **JCDL2020** Jinsong Zhang, Chun Guo, and Xiaozhong Liu. Characterize and Evaluate Scientific Domain and Domain Context Knowledge Map. In *Proceedings of ACM/IEEE Joint Conference on Digital Libraries*. ACM/IEEE, 2020

- [C71] Anti-Spam Deep-Learning Representation Learning SIGIR2020 Guoxiu He, Zhe Gao, Zhuoren Jiang, Yangyang Kang, Changlong Sun, Xiaozhong Liu, and Wei Lu. Think Beyond the Word: Understanding the Implied Textual Meaning by Digesting Context, Local, and Noise. In *Proceedings of the 43rd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2020a
- [C70] Summarization Review Mining Deep-Learning SIGIR2020 Haojie Pan, Xin Zhou, Changlong Sun, Fei Wu, and Xiaozhong Liu. Large Scale Abstractive Multi-Review Summarization (LSARS) via Aspect Alignment. In *Proceedings of the 43rd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2020
- [C69] Graph Mining Deep-Learning Cross-Domain SIGIR2020 Zheng Gao, Hongsong Li, Zhuoren Jiang, and Xiaozhong Liu. Detecting User Community in Sparse Domain via Cross-Graph Pairwise Learning. In *Proceedings of the 43rd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2020a
- [C68] Content Generation User Behavior Information Retrieval SIGIR2020 Yongzhen Wang, Jian Wang, Heng Huang, Hongsong Li, and Xiaozhong Liu. Evolutionary Product Description Generation: A Dynamic Fine-Tuning Approach Leveraging User Click Behavior. In *Proceedings of the 43rd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2020d
- [C67] Children Protection User Behavior Deep-Learning SIGIR2020 Guoxiu He, Yangyang Kang, Zhuoren Jiang, Jiawei Liu, Changlong Sun, Wei Lu, and Xiaozhong Liu. Creating a Children-Friendly Reading Environment via Joint Learning of Content and Human Attention. In *Proceedings of the 43rd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2020c
- [C66] Deep Graph Genetic Algorithm Data Heterogeneity WWW2020 Zhuoren Jiang, Zheng Gao, Jinjiong Lan, Hongxia Yang, Lu Yao, and Xiaozhong Liu. Task-Oriented Genetic Activation for Large-Scale Complex Heterogeneous Graph Embedding. In *The World Wide Web Conference*. ACM, 2020
- [C65] Summarization User Behavior Deep Graph ECAI2020 Zheng Gao, Lujun Zhao, Hongsong Li, Chenglong Sun, Luo Si, and Xiaozhong Liu. Behavior based Dynamic Summarization on Product Aspects via Reinforcement Neighbour Selection. In *Proceedings of the 24th European Conference on Artificial Intelligence*. IEEE, 2020b
- [C64] Legal Intelligence Dialog Mining Communication Technology AAI2020 Tianyi Wang, Yating Zhang, Changlong Sun, and Xiaozhong Liu. Masking Orchestration: Multi-task Pretraining for Multi-role Dialogue Representation Learning. In *Proceedings of the Thirty-Fourth AAI Conference on Artificial Intelligence*. AAI, 2020b
- [C63] Sentiment Mining Dialog Mining Communication Technology User-Behavior AAI2020 Kaisong Song, Changlong Sun, and Xiaozhong Liu. Sentiment Classification in Customer Service Dialogue with Topic-aware Multi-task Learning. In *Proceedings of the Thirty-Fourth AAI Conference on Artificial Intelligence*. AAI, 2020
- [C62] Natural Language Generation Cross-Language Deep Learning AAI2020 Juntao Li, Jian Wang, Rui Yan, and Xiaozhong Liu. Cross-Lingual Low-Resource Set-to-Description Retrieval for Global E-Commerce: New Data and New Challenges. In *Proceedings of the Thirty-Fourth AAI Conference on Artificial Intelligence*. AAI, 2020
- [C61] Cybercrime Misinformation User-Behavior SIGIR19 Guoxiu He, Yangyang Kang, Zhe Gao, Zhuoren Jiang, Changlong Sun, Xiaozhong Liu, Wei Lu, Qiong Zhang, and Luo Si. Finding

Camouflaged Needle in a Haystack?: Pornographic Products Detection via Berrypicking Tree Model. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 365–374. ACM, 2019

- [C60] Legal Intelligence Deep-Learning SIGIR19 Xin Zhou, Yating Zhang, Xiaozhong Liu, Changlong Sun, and Luo Si. Legal Intelligence for E-commerce: Multi-task Learning by Leveraging Multiview Dispute Representation. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 315–324. ACM, 2019
- [C59] Legal Intelligence Communication Technology Summarization CIKM19 Xinyu Duan, Yating Zhang, Xin Zhou, Changlong Sun, Fei Wu, Luo Si, and Xiaozhong Liu. Legal Summarization for Multi-role Debate Dialogue via Controversy Focus Mining and Multi-task Learning. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. ACM, 2019
- [C58] Computational Social Justice Misinformation Event Detection CIKM19 Yingchi Liu, Quanzhi Li, Xiaozhong Liu, Qiong Zhang, and Luo Si. Sexual Harassment Story Classification and Key Information Identification. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. ACM, 2019b
- [C57] User-Behavior Deep-Learning Sentiment Analysis CIKM19 Zhuoren Jiang, Jian Wang, Lujun Zhao, Changlong Sun, Yao Lu, and Xiaozhong Liu. Cross-domain Aspect Category Transfer and Detection via Traceable Heterogeneous Graph Representation Learning. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. ACM, 2019b
- [C56] Misinformation Cybercrime Deep Graph EMNLP19 Zhuoren Jiang, Zhe Gao, Guoxiu He, Changlong Sun, Qiong Zhang, Luo Si, and Xiaozhong Liu. Detect Camouflaged Spam Content via StoneSkipping: Graph and Text Joint Embedding for Chinese Character Variation Representation. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 2019a
- [C55] Legal Intelligence NLP Computational Social Justice EMNLP19 Yingchi Liu, Quanzhi Li, Marika Cifor, Xiaozhong Liu, Qiong Zhang, and Luo Si. Uncover Sexual Harassment Patterns from Personal Stories by Joint Key Element Extraction and Categorization. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 2019a
- [C54] Sentiment Analysis NLP Deep-Learning EMNLP19 Jingjing Wang, Changlong Sun, Shoushan Li, Jiancheng Wang, Luo Si, Min Zhang, Xiaozhong Liu, and Guodong Zhou. Human-Like Decision Making: Document-level Aspect Sentiment Classification via Hierarchical Reinforcement Learning. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 2019b
- [C53] Sentiment Analysis NLP Deep-Learning EMNLP19 Kaisong Song, Lidong Bing, Wei Gao, Jun Lin, Changlong Sun, Lujun Zhao, Xiaozhong Liu, and Qiong Zhang. Using Customer Service Dialogues for Satisfaction Analysis with Context-Assisted Multiple Instance Learning. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 2019a
- [C52] NLP Sentiment Analysis Deep-Learning IJCAI19 Kaisong Song, Wei Gao, Lujun Zhao, Jun Lin, Changlong Sun, and Xiaozhong Liu. Cold-Start Aware Deep Memory Network for Multi-Entity Aspect-Based Sentiment Analysis. In *Proceedings of the 2019 International Joint Conferences on Artificial Intelligence*, 2019b
- [C51] Sentiment Analysis NLP Deep-Learning ACL19 Jingjing Wang, Changlong Sun, Shoushan Li, Xiaozhong Liu, Luo Si, Min Zhang, and Guodong Zhou. Aspect Sentiment Classification Towards Question-Answering with Reinforced Bidirectional Attention Network. In *Proceedings of the 57th Conference of the Association for Computational Linguistics*, pages 3548–3557, 2019a

- [C50] Content Generation User Behavior Deep-Learning WWW19 Yongzhen Wang, Heng Huang, Yuliang Yan, and Xiaozhong Liu. Quality-Sensitive Training! Social Advertisement Generation by Leveraging User Click Behavior. In *The World Wide Web Conference*, pages 2045–2055. ACM, 2019c
- [C49] Content Generation NLP Communication Technology WWW19 Lujun Zhao, Kaisong Song, Changlong Sun, Qi Zhang, Xuanjing Huang, and Xiaozhong Liu. Review Response Generation in E-Commerce Platforms with External Product Information. In *The World Wide Web Conference*, pages 2425–2435. ACM, 2019
- [C48] Text Classification NLP Deep-Learning WWW19 Huang Heng, Liu Xiaozhong, and Yuliang Yan. Domain-aware Neural Model for Sequence Labeling using Joint Learning. In *The World Wide Web Conference*, pages 2837–2843. ACM, 2019
- [C47] Graph-Mining Genetic Algorithm GECCO19 Zheng Gao, Chun Guo, and Xiaozhong Liu. Efficient personalized community detection via genetic evolution. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 383–391. ACM, 2019b
- [C46] IR Graph-Mining Deep-Learning SIGIR18 Zhuoren Jiang, Yue Yin, Liangcai Gao, Yao Lu, and Xiaozhong Liu. Cross-language Citation Recommendation via Hierarchical Representation Learning on Heterogeneous Graph. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pages 635–644. ACM, 2018b
- [C45] QA Communication Technology Sentiment Analysis COLING18 Lu Wang, Shoushan Li, Changlong Sun, Luo Si, Xiaozhong Liu, Min Zhang, and Guodong Zhou. One vs. many QA matching with both word-level and sentence-level attention network. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 2540–2550, 2018
- [C44] Summarization Graph Mining Deep-Learning EMNLP18 Yongzhen Wang, Xiaozhong Liu, and Zheng Gao. Neural Related Work Summarization with a Joint Context-driven Attention Mechanism. *arXiv preprint arXiv:1901.09492*, 2019d
- [C43] QA Communication Technology Sentiment Analysis Deep-Learning EMNLP18 Chenlin Shen, Changlong Sun, Jingjing Wang, Yangyang Kang, Shoushan Li, Xiaozhong Liu, Luo Si, Min Zhang, and Guodong Zhou. Sentiment classification towards question-answering with hierarchical matching network. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3654–3663, 2018
- [C42] Cyberlearning Graph-Mining Formula-Mining CIKM18 Zhuoren Jiang, Liangcai Gao, Ke Yuan, Zheng Gao, Zhi Tang, and Xiaozhong Liu. Mathematics Content Understanding for Cyberlearning via Formula Evolution Map. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 37–46. ACM, 2018a
- [C41] Reinforcement Learning Multi-Agent RecSys17 Yang Zhang, Chenwei Zhang, and Xiaozhong Liu. Dynamic scholarly collaborator recommendation via competitive multi-agent reinforcement learning. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, pages 331–335. ACM, 2017
- [C40] Recommendation Graph-Mining ICTIR17 Xiaozhong Liu, Yingying Yu, Zhuoren Jiang, Chun Guo, and Scott Jensen. Personalized Navigation and Random Walk on a Complex Heterogeneous Graph. In *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval*, pages 217–224. ACM, 2017
- [C39] Career Recommendation Graph-Mining iConference17 Nan Li, Naren Suri, Zheng Gao, Tian Xia, Katy Börner, and Xiaozhong Liu. Enter a job, get course recommendations. *iConference 2017*

- [C38] Filter Bubble Social Media Text-Mining iConference17 Tian Xia, Xing Yu, Zheng Gao, Yijun Gu, and Xiaozhong Liu. Internal/External information access and information diffusion in social media. *iConference 2017 Proceedings Vol. 2, 2017*
- [C37] Graph-Mining Personalization iConference17 Zheng Gao and Xiaozhong Liu. Personalized community detection in scholarly network. *iConference 2017 Proceedings Vol. 2, 2017*
- [C36] Filter Bubble Social Media Graph-Mining ICWSM16 Xiaozhong Liu, Tian Xia, Yingying Yu, Chun Guo, and Yizhou Sun. Cross Social Media Recommendation. In *Tenth international aaai conference on web and social media, 2016a*
- [C35] Cyberlearning Social Media Personalization SIGIR16 Zhuoren Jiang, Xiaozhong Liu, Liangcai Gao, and Zhi Tang. Community-based Cyberreading for Information Understanding. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 789–792. ACM, 2016b
- [C34] Filter Bubble Social Media Graph-Mining Hypertext16 Xiaozhong Liu, Xing Yu, Zheng Gao, Tian Xia, and Johan Bollen. Comparing community-based information adoption and diffusion across different microblogging sites. In *Proceedings of the 27th ACM Conference on Hypertext and Social Media*, pages 103–112. ACM, 2016b
- [C33] Music Recommendation Graph-Mining Big Data16 Chun Guo and Xiaozhong Liu. Dynamic feature generation and selection on heterogeneous graph for music recommendation. In *2016 IEEE International Conference on Big Data (Big Data)*, pages 656–665. IEEE, 2016
- [C32] Summarization NLP Scholarly Text iConference16 Lin Zhang and Xiaozhong Liu. Synopsizing “literature review” for scientific publications. *iConference 2016 Proceedings, 2016 (Lee Dirks Award for Best paper Nominee)*
- [C31] Citation Recommendation Graph-Mining iConference16 Xiaozhong Liu, Jinsong Zhang, and Chun Guo. Citation recommendation via proximity full-text citation analysis and supervised topical prior. *ICConference 2016 Proceedings, 2016c*
- [C30] Filter Bubble Social Media Text-Mining iConference16 Tian Xia, Miao Chen, and Xiaozhong Liu. Twibo: comparing very large communities via massive social media datasets. *ICConference 2016 Proceedings, 2016*
- [C29] Cyberlearning IR Social Media SIGIR15 Xiaozhong Liu, Zhuoren Jiang, and Liangcai Gao. Scientific information understanding via open educational resources (OER). In *Proceedings of the 38th international acm sigir conference on research and development in information retrieval*, pages 645–654. ACM, 2015
- [C28] Music Recommendation Graph-Mining SIGIR15 Chun Guo and Xiaozhong Liu. Automatic feature generation on heterogeneous graph for music recommendation. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 807–810. ACM, 2015
- [C27] Graph-Mining Text-Mining SIGIR15 Workshop Yingying Yu, Zhuoren Jiang, and Xiaozhong Liu. Random Walk and Feedback on Scholarly Network. In *GSB@ SIGIR*, pages 33–37, 2015
- [C26] Formula-Mining IR JCDL15 Yuehan Wang, Liangcai Gao, Simeng Wang, Zhi Tang, Xiaozhong Liu, and Ke Yuan. WikiMirs 3.0: a hybrid MIR system based on the context, structure and importance of formulae in a document. In *Proceedings of the 15th ACM/IEEE-CS joint conference on*



*digital libraries*, pages 173–182. ACM, 2015 (*Best paper Nominee*)

- [C25] **Dynamic Topic Modeling** **Text-Mining** **CIKM15** Zhuoren Jiang, Xiaozhong Liu, and Liangcai Gao. Chronological citation recommendation with information-need shifting. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pages 1291–1300. ACM, 2015
- [C24] **Digital Library** **iConference15** Siyuan Guo, Trevor Edelblute, Bin Dai, Miao Chen, and Xiaozhong Liu. Toward enhanced metadata quality of large-scale digital libraries: Estimating volume time range. *iConference 2015 Proceedings*, 2015
- [C23] **Graph-Mining** **Topic Modeling** **CIKM14** Xiaozhong Liu, Yingying Yu, Chun Guo, and Yizhou Sun. Meta-path-based ranking with pseudo relevance feedback on heterogeneous graph for citation recommendation. In *Proceedings of the 23rd acm international conference on conference on information and knowledge management*, pages 121–130. ACM, 2014c
- [C22] **Graph-Mining** **Digital Library** **JCDL14** Xiaozhong Liu, Yingying Yu, Chun Guo, Yizhou Sun, and Liangcai Gao. Full-text based context-rich heterogeneous network mining approach for citation recommendation. In *IEEE/ACM Joint Conference on Digital Libraries*, pages 361–370. IEEE, 2014d
- [C21] **Metadata** **DC14** Xiaozhong Liu, Miao Chen, and Jian Qin. Interlinking cross language metadata using heterogeneous graphs and Wikipedia. In *International Conference on Dublin Core and Metadata Applications*, pages 157–166, 2014a
- [C20] **Filter Bubble** **Social Media** **Data Science** **Hypertext14** Xin Shuai, Xiaozhong Liu, Tian Xia, Yuqing Wu, and Chun Guo. Comparing the pulses of categorical hot events in twitter and weibo. In *Proceedings of the 25th ACM conference on Hypertext and social media*, pages 126–135. ACM, 2014
- [C19] **IR** **Formula-Mining** **SIGIR14** Xiaoyan Lin, Liangcai Gao, Xuan Hu, Zhi Tang, Yingnan Xiao, and Xiaozhong Liu. A mathematics retrieval system for formulae in layout presentations. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*, pages 697–706. ACM, 2014
- [C18] **Citation Mining** **Machine Learning** **ASIST14** Chun Guo, Yingying Yu, Azade Sanjari, and Xiaozhong Liu. Citation role labeling via local, pairwise, and global features. *Proceedings of the American Society for Information Science and Technology*, 51(1):1–10, 2014
- [C17] **Graph-Mining** **Topic Modeling** **CIKM14** Xiaozhong Liu, Yingying Yu, Chun Guo, and Yizhou Sun. Meta-path-based ranking with pseudo relevance feedback on heterogeneous graph for citation recommendation. In *Proceedings of the 23rd acm international conference on conference on information and knowledge management*, pages 121–130. ACM, 2014c
- [C16] **Wikipedia** **Text-Mining** **CIKM14 Workshop** Zhuoren Jiang, Miao Chen, and Xiaozhong Liu. Semantic annotation with rescoredesa: Rescoring concept features generated from explicit semantic analysis. In *Proceedings of the 7th International Workshop on Exploiting Semantic Annotations in Information Retrieval*, pages 25–27. ACM, 2014
- [C15] **Wikipedia** **Bibliometrics** **JCDL13** Xin Shuai, Zhuoren Jiang, Xiaozhong Liu, and Johan Bollen. A comparative study of academic and Wikipedia ranking. In *Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries*, pages 25–28. ACM, 2013
- [C14] **Social Media** **Cyberlearning** **iConference13** Han Jia and Xiaozhong Liu. Scientific referential metadata creation with information retrieval and labeled topic modeling. 2013

- [C13] **Social Media** **Text-Mining** **CIKM13 Workshop** Eric Baucom, Azade Sanjari, Xiaozhong Liu, and Miao Chen. Mirroring the real world in social media: twitter, geolocation, and sentiment analysis. In *Proceedings of the 2013 international workshop on Mining unstructured big data using natural language processing*, pages 61–68. ACM, 2013
- [C12] **Graph-Mining** **IR** **CIKM12** Xiaozhong Liu, Jinsong Zhang, and Chun Guo. Full-text citation analysis: enhancing bibliometric and scientific publication ranking. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, pages 1975–1979. ACM, 2012
- [C11] **IR** **Knowledge Management** **ASIST12** Chun Guo, Renuka Chinchankar, and Xiaozhong Liu. Knowledge retrieval for scientific literatures. *Proceedings of the American Society for Information Science and Technology*, 49(1):1–7, 2012
- [C10] **IR** **Wikipedia** **SIGIR12 Workshop** Le Zhao, Xiaozhong Liu, and Jamie Callan. WikiQuery—An Interactive Collaboration Interface for Creating, Storing and Sharing Effective CNF Queries. In *SIGIR 2012 Workshop on Open Source Information Retrieval*, pages 1–8, 2012
- [C9] **IR** **Social Media** **WWW12 Workshop** Xin Shuai, Xiaozhong Liu, and Johan Bollen. Improving news ranking by community tweets. In *Proceedings of the 21st international conference on world wide web*, pages 1227–1232. ACM, 2012
- [C8] **eScience12** Scott Jensen, Beth Plale, Xiaozhong Liu, Miao Chen, David Leake, and Julie England. Generalized representation and mapping for social-ecological data: Freeing data from the database. In *2012 IEEE 8th International Conference on E-Science*, pages 1–8. IEEE, 2012
- [C7] **DC10** Jian Qin, Miao Chen, Xiaozhong Liu, and Andrea Kathleen Wiggins. Linking Entities in Scientific Metadata. In *International Conference on Dublin Core and Metadata Applications*, pages 128–136, 2010
- [C6] **SIGIR09** Xiaozhong Liu. Using Computational Community Interest as an Indicator for Ranking. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 851–851. ACM, 2009
- [C5] **CIKM09** Xiaozhong Liu and Vadim von Brzeski. Computational community interest for ranking. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pages 245–254. ACM, 2009
- [C4] **DC09** Jian Qin, Xiaozhong Liu, Xia Lin, and Miao Chen. Using metadata for query refinement and recommendation. In *International Conference on Dublin Core and Metadata Applications*, pages 15–20, 2009
- [C3] **ASIST08** Xiaozhong Liu, Jian Qin, Miao Chen, and Ji-Hong Park. Automatic semantic mapping between query terms and controlled vocabulary through using WordNet and Wikipedia. *Proceedings of the American Society for Information Science and Technology*, 45(1):1–10, 2008
- [C2] **DC08** Miao Chen, Xiaozhong Liu, and Jian Qin. Semantic relation extraction from socially-generated tags: a methodology for metadata generation. In *International Conference on Dublin Core and Metadata Applications*, pages 117–127, 2008
- [C1] **HICSS07** Xiaozhong Liu, Robert Benjamin, and Ping Zhang. Re-Examining IT Enabled Change with a New Model of the Information Field: The Tiger Creek Case. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*, pages 248a–248a. IEEE, 2007

## PATENT

---

- [P9] Hot Term Prediction for Contextual Shortcuts, US Patent.
- [P8] Method and device for determining keyword variants, Chinese Patent.
- [P7] Knowledge based Cross-Domain Personalized Travel Information Recommendation, Chinese Patent.
- [P6] Personalized Travel-Guide Generation, Chinese Patent, Chinese Patent.
- [P5] Legal Document Retrieval with Legal Knowledge Integration and Deep Representation Learning, Chinese Patent.
- [P4] Automatic Controversy Focus Mining and Identification from Legal Documents, Chinese Patent.
- [P3] Cross-domain Aspect Detection and System based on Heterogeneous Graph Representation Learning, Chinese Patent.
- [P2] A Commodity Attribute Extraction Self-learning System with Fusion Behavior and Feedback Information, Chinese Patent.
- [P1] Dynamic Interpretations on Item Chronological Aspect-Sentiment via Multi-Type User Behavior Data, Chinese Patent.

## GRANTS

---

- 2021–2022 **Co-principal investigator** DoD, \$2,100,000, for “Embedded Systems Security / Artificial Intelligence (ESS/ AI) Research and Workforce Development” (David Crandall, Andrew Lukefahr, Gregory Lewis, Bennett Bertenthal, Katy Borner, Jean Camp, Lauren Ann Christopher, Chunfeng Huang, Apu Kapadia, David B Leake and Donald Williamson).
- 2020–2022 **Senior Personnel** ARO-IARPA, \$614,996, for “TrojAI: “Statistical Methods for Backdoor Detection in Deep Neural Networks”, part of the Project Iliad of Perspecta Labs Role” (with Xiaofeng Wang and Haixu Tang).
- 2020–2022 **Co-principal investigator** NSF, Convergence Accelerator Phase II (RAISE) Grant, \$5,000,000, for “Situational Awareness for Employment Resilience (SAFER) (invited submission, pending)” (with Katy Borner, Kemi Jona, Martin Storksdieck, André Skupin and Kosali Simon).
- 2020 - 2021 **Principal investigator** Racial Justice Research Fund (RJRF), \$14,078, “[Chinese Student  $\approx$  Spy?] [Black Man  $\approx$  Predator?] Reducing Racial Bias in Legal Artificial Intelligence” (with Devan Donaldson).
- 2019 - 2020 **Senior Personnel** NSF, Convergence Accelerator Phase I (RAISE) Grant, \$995,000, “Analytics-Driven Accessible Pathways To Impacts-Validated Education (ADAPTIVE)” (with Katy Borner, Kemi Jona, Martin Storksdieck and Ariel Anbar).
- 2018 - 2021 **Principal investigator** Alibaba DAMO Academy, \$182,409, “Legal Intelligence: Legal Outcome Prediction with Court Debate Data”.
- 2019 - 2022 **Principal investigator** Alibaba DAMO Academy, \$134,327, “Cybercrime Detection and eCommerce Security with Deep Learning”.
- 2019 - 2020 **Principal investigator** Institute for Advanced Study (IAS) Grant, \$9,000, “Automatic Scholarly Information Summarization via NLP and Graph Mining” (with Stasa Milojevic).

- 2018 - 2021 **Co-principal investigator** Institute of Museum and Library Services (IMLS) Grant, \$249,908 "Enabling Accessibility and Linking Digital Media Collections in Academic Libraries" (with Yun Huang and Jian Qin) .
- 2016 - 2018 **Principal investigator** Social Science Research Commons (SSRC) Grant, \$14,800, "Cross Twitter and Weibo Information Mining and Analysis: Infrastructure, Methodology, and Applications" (with Weihua An and Hsien-Chang Lin)
- 2014 - 2015 **Principal investigator** Scholarship of Teaching and Learning (SoTL) Grant, \$5,000 "Scholarship of Teaching and Learning, OER-based Collaborative Scaffolding: Assisting In-Depth Understanding of Scientific Publications via Open Educational Resources"
- 2014 - 2015 **Principal investigator** Amazon Web Services Education Grant, \$2,600 "Z604: Big Data Analytics for Web and Text"
- 2013 - 2014 **Principal investigator** Amazon Web Services Education Grant, \$4,800 "Z534 Information Retrieval: Theory and Practice"
- 2012 - 2014 **Principal investigator** Lilly Grant, \$26,432 "Text Mining and Natural Language Processing Research via Data 2 Insight Center"
- 2012 - 2013 **Co-principal investigator** Faculty Research Support Program (FRSP) Grant, \$73,625 "Socio-eco-informatics: enhancing predictive capability of social ecological systems research" (with Beth Plale and David Leake)

## INVITED PRESENTATION

---

- 2019, CS Colloquium: Legal Intelligence and Cybercrime Detection via Natural Language Processing, Indiana University Bloomington
- 2019, Dark Business Detection via NLP and User Behavior Data Mining, Wuhan University
- 2019, Cybercrime Detection via Artificial Intelligence, Zhejiang University
- 2018, Natural Language Processing with User Behavior Data, University of Maryland (College Park)
- 2018, Legal Intelligence Challenge and Opportunity, The Computing Conference 2018 of Alibaba Cloud
- 2016, Scientific Information Understanding, University of Pittsburgh
- 2014, Scientific literature recommendation via full-text publication data and heterogeneous graph mining. School of Information, The University of Texas at Austin.
- 2014, ILS Colloquium: Connect Language, Network and Culture Bubbles: Cross-Twitter and Weibo Information Recommendation and Community Comparison, Indiana University Bloomington
- Keynote Talk: Fulltext Citation Analysis and its Applications. International Workshop on Computational Scientometrics: Theory and Applications (CIKM2013)
- December 2013, Catapult Center workshop: Open Educational Resources based Scaffolding for Cyberlearning, Indiana University Bloomington
- 2013, Fulltext citation analysis and cyberlearning. Yahoo!-DAIS Seminar. Department of Computer Science, University of Illinois at Urbana-Champaign

- 2013, Catapult Center's workshop: Information retrieval to Knowledge retrieval, one more step, Indiana University Bloomington
- 2012, SLIS Brown Bag Series: Full Text Citation Analysis for Scientific Recommendation, Indiana University Bloomington
- 2011, Natural Language Processing, Zhejiang University, China
- 2011, Statistical Information Retrieval and Semantic Web, Zhejiang University, China
- 2011, Information Retrieval to Knowledge Retrieval, Chinese Academy of Sciences, China

## TEACHING

---

- B659 - Topics in Artificial Intelligence: Big Data Analytics, Indiana University Bloomington, January 2020 – Present.
- Z604 - Z604: Big Data Analytics for Web and Text, School of Informatics and Computing, Indiana University Bloomington, January 2014 – Present. (*IU Trustees Teaching Award*)
- Z517 - Web Programming, School of Informatics and Computing, Indiana University Bloomington, January 2011 – Present.
- Z534 - Information Retrieval: Theory and Practice (Search), School of Informatics and Computing, Indiana University Bloomington, January 2011 – Present. (*IU Trustees Teaching Award*)
- IST256 – Applied Programming For Information System, School of Information Studies, Syracuse University, August 2009 – December 2010
- IST400 – Natural Language Processing, School of Information Studies, Syracuse University, August 2008 – August 2009 (Teaching Assistant)
- IST400 – Natural Language Processing, School of Information Studies, Syracuse University (Teaching Assistant)
- IST400 – Search Engine Design & Implement, School of Information Studies, Syracuse University (Teaching Assistant)

## AWARDS AND ACHIEVEMENTS

---

- **Rank No. 1** (for 400+ teams) for Chinese AI and Law Challenge Competition (CAIL2018)
- **Best Reviewer Award (2016)**. Journal of Information Processing & Management
- **iConference Lee Dirks Award for Best paper Nominee (2016)** – Synopsizing “Literature Review” for Scientific Publications, iConference.
- **IU Trustees Teaching Award (2016)**: Indiana University established the Trustees Teaching Awards (TTA) to recognize and enhance excellent teaching at Indiana University.
- **IU Trustees Teaching Award (2015)**: Indiana University established the Trustees Teaching Awards (TTA) to recognize and enhance excellent teaching at Indiana University.
- **JCDL Best paper Nominee (2015)**: Wikimirs 3.0: a hybrid mir system based on the context, structure and importance of formulae in a document.
- **JCDL Best Poster Award (2013)**: for my paper Jiang, Z. & Liu, X. (2013) Recovering Missing Citations in Scholarly Network: A 2-Step Citation Analysis to Estimate Publication Importance,

Joint Conference on Digital Libraries (JCDL).

- **University Fellowship (2005-2010):** top university fellowship at Syracuse University.

## CONFERENCE PANEL

---

- Liu, X., Wu, D., He, D., & Li, Y. (2019). Breaking Social Media Bubbles for Information Globalization: A Cross-Cultural and Cross-Language User Centered Sense Making Approach. Proceedings of the American Society for Information Science & Technology (ASIST).
- Liu, X., & Hara, N., Kazmer, M., Subramaniam, M., & Twidale, M. (2015). Education in the Cyber-learning Era: New Challenges, Opportunities, and Applications. Proceedings of the American Society for Information Science & Technology (ASIST).
- Chen, H., Gruzd, A., Liu, X., & Meyers, E. (2012). Understanding Information and Knowledge Sharing in Online Communities: Emerging Research Approaches. Proceedings of the American Society for Information Science & Technology (ASIST).

## CONFERENCE WORKSHOP

---

- Liu, X., Sun, C., Zhang, Y. & Wu, F. (2020) Legal Intelligence: Algorithmic, Data, and Social Challenges. SIGIR 2020.
- Li, Q., & Liu, X. (2018). The 2nd International Workshop on Big Data for Financial News and Data. IEEE BigData 2018.
- Liu, X., Ding Y., Song, M. & Chen, M. (2013). Mining Unstructured Big Data using Natural Language, ACM International Conference on Information and Knowledge Management (CIKM)
- Caragea, C., Giles, L. & Liu, X. (2013). Computational Scientometrics: Theory and Applications, ACM International Conference on Information and Knowledge Management (CIKM)

## REVIEW and COMMITTEE

---

- Journal Editorial Board member: Journal of Information Processing & Management
- Journal Reviewer: Journal of the American Society for Information Science and Technology
- Journal Reviewer: Journal of Information Processing & Management
- Journal Reviewer: Journal of Informetrics
- Journal Reviewer: ACM Transactions on Information Systems
- Journal Reviewer: Journal of Library and Information Science
- Conference Planning Committee/Local Chair: 2016 ACM International Conference on Information and Knowledge Management (CIKM)
- Conference Senior PC Member: 2019 Joint Conference of Digital Library (JCDL)
- Conference PC Member: 2016, 2017, 2018, 2019 International ACM SIGIR Conference on Research and Development in Information Retrieval
- Conference PC Member: 2018, 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)
- Conference PC Member: 2018, 2019 Conference of the Association for Computational Linguistics

(ACL)

- Conference Planning Committee/PC Member: 2013, 2014, 2015, 2016, 2017, 2018 Joint Conference of Digital Library (JCDL)
- Publication Chair: 2013 ACM/IEEE Joint Conference of Digital Library (JCDL)
- Workshop Chair: 2017 ACM/IEEE Joint Conference of Digital Library (JCDL)
- Student Volunteer Chair: 2013 ACM/IEEE Joint Conference of Digital Library (JCDL)
- PC Member: 2014, 2015 ACM/IEEE Joint Conference of Digital Library (JCDL)
- PC Member: 2014, 2015 International Conference on Web Information Systems and Technologies (WEBIST)
- PC Member: 2014, 2015 European Conference on Information Retrieval (ECIR)
- Conference Reviewer: American Society for Information Science & Technology (ASIST)
- Conference Reviewer: iConference

## SELECTED PROJECT

---

### **Legal Intelligence**

**Bloomington, Indiana**

*Project PI*

*2015 - Current*

- This project makes original contributions on theories, methodologies, and applications of legal text/data mining, machine learning, legal knowledge graph generation/mining, natural language processing, and legal knowledge representation in the legal ecosystem.

### **Cybercrime and Dark Business Detection**

**Bloomington, Indiana**

*Project PI*

*2014 - Current*

- The Internet is not a safe place, especially for children and teenagers, to be roaming around. Cybercrime and harmful content/products, like erotica novels, pornography videos, illegal drugs, and violence information, are polluting the online ecosystems. The goal of this proposed project is to develop a theoretical, algorithmic, and systematic foundation for effective online dark business detection by means of massive user information-seeking-behavior, text mining and joint representation learning.

### **Auto-Travel Guide Generation**

**Hangzhou, China**

*Project PI*

*2017 - 2018*

- By using massive user behavior and user travel data, the proposed algorithm and system can automatically generate personalized travel guide information by leveraging novel deep learning algorithms.

### **Cyberlearning via Collaborative PDF Reader System**

**Bloomington, Indiana**

*Project PI*

*2013 - Current*

- Assist students to better understand scientific publications. Students can collaborate with others classmates and backend algorithms for scientific content understanding.

### **MatchMaker Scholar Recommendation System**

**Bloomington, Indiana**

*Project PI*

*2014 - 2018*

- Recommending high-quality scholar resources, i.e., publications, experts, venues, funding opportunities, via full-text citation network analysis and text mining techniques.

### **ScholarWiki Automated Learning Environments (SWALE)**

**Bloomington, Indiana**

*Project PI*

*2013 - 2017*

- ScholarWiki, an experimental system developed as a proof of innovative metadata model, was built over the wiki platform to allow user interaction with the metadata and the editing, deleting, adding metadata. By allowing low-barrier participation, scholars can participate in the knowledge and metadata editing and enhancing process and benefit from more accurate and effective information retrieval.

### **MatchMaker Scholar Recommendation System**

**Bloomington, Indiana**

*Project PI*

2014 - 2018

- Recommending high-quality scholar resources, i.e., publications, experts, venues, funding opportunities, via full-text citation network analysis and text mining techniques.

**Real-time News ranking Algorithm**

**Sunnyvale, California**

*Project PI*

2009 - 2013

- During my work in Yahoo! contextual search team, developed the news ranking system for Yahoo! Search, which can automatically extract real-time community interest from user-oriented text data and rank the news search results.

## PEER REVIEWED CONFERENCE POSTER

---

- Zhang, C., Gao, Z., & Liu, X., (2016) How others affect your Twitter hashtag adoption? Examination of community-based and context-based information diffusion in Twitter, Proceedings of the iConference.
- Guo, C., & Liu, X. (2015) Feature Selection on Heterogeneous Graph, Proceedings of the Association for Information Science and Technology (ASIST) Annual Meeting.
- Jensen, S., Yu, Y., Liu, H., & Liu, X. (2014) Query-Centric Scientific Topic Evolution Extraction, Proceedings of the Association for Information Science and Technology (ASIST) Annual Meeting.
- Pan, Y., & Liu, X., (2015) Using full-text citation network to enhance the keyword label performance, Proceedings of the iConference.
- Yu, Y., Guo, C., & Liu, X., (2015) Explicit Graphical Relevance Feedback for Scholarly Information Retrieval, Proceedings of the iConference.
- Guo, C., Staša, M., & Liu, X., (2015) How are academic articles cited over time, Proceedings of the iConference.
- Liu, X., Gao, L., Hara, L., & Sun, Y. (2014) Open Educational Resource based Information Understanding via PDF Document Interaction, Proceedings of the ACM Web Science.
- Landrum, J., Hara, N., & Liu, X. (2014) Scaffolding of Scientific Publications with Open Educational Resources (OER), Proceedings of the Association for Information Science and Technology (ASIST) Annual Meeting.
- Jia, H., Guo, C., & Liu, X. (2014) Mobile Application Life Cycle Characterization via Apple App Store Rank, Proceedings of the Association for Information Science and Technology (ASIST) Annual Meeting.
- Xia, T., Chen, M., & Liu, X. (2014) Explicit Semantic Path Mining via Wikipedia Knowledge Tree, Proceedings of the Association for Information Science and Technology (ASIST) Annual Meeting.
- Zhang, J., & Liu, X. (2013). Full-text and Topic based AuthoRank and Enhanced Publication Ranking. In Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries (pp. 393-394). ACM.
- Jiang, Z., & Liu, X. (2013) Recovering Missing Citations in Scholarly Network: A 2-Step Citation Analysis to Estimate Publication Importance, In Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries (pp. 419-420). ACM. (Best Poster Award)
- Guo, C., Zhang, J., & Liu, X. (2013). Scientific Metadata Quality Enhancement for Scholarly Publications, Proceedings of the iConference
- Zhao, X., Liu, X., & Ye, Y. (2013). Characterizing the Scholar H-Index via Full-Text Citation Analysis, Proceedings of the iConference



- Zhang, J., Guo, C.& Liu, X. (2012) Characterize Scientific Domain and Domain Context. Proceedings of the 12th ACM/IEEE Joint Conference on Digital Libraries
- Liu, X.,& Jia, H. (2012) Scientific Cyberlearning Resources Referential Metadata Creation Via Information Retrieval. Proceedings of the 12th ACM/IEEE Joint Conference on Digital Libraries
- Jensen, S., Chen, M., Liu, X., Plale, B.,& Leake, D. (2012) Mining Classifications from Social-Ecological Databases. Proceedings of the American Society for Information Science and Technology (ASIST).
- Chen, M.,& Liu, X., (2011) Predicting popularity of online distributed applications: iTunes app store case analysis. Proceedings of the iConference.
- Finlay, S. C., Johnson, M., Malic, V., Hank, C., Liu, X., Ni, C.,& Sugimoto, C. R. (2011) Exploring connections of the biblioblogosphere. Proceedings of the American Society for Information Science and Technology (ASIST).
- Liu, X., Qin, J.,& Chen, M. (2011). ScholarWiki system for knowledge indexing and retrieval. Proceedings of the American Society for Information Science& Technology (ASIST).
- Crowston, K., Liu, X.,& Allen E. E. (2010) Machine learning and rule based automated coding of qualitative data. Proceedings of the American Society for Information Science and Technology (ASIST).
- Liu, X. (2010). Building Folk UMLS: An Approach to Finding Meaning of Folk Terms in Medical Domain. Proceedings of the iConference.
- Chen, M., Yu, B.,& Liu, X. (2010). Building Folk UMLS: An Approach to Finding Meaning of Folk Terms in Medical Domain. Proceedings of the iConference.
- Liu, X., (2009) Comment centric news analysis for ranking. Proceedings of the American Society for Information Science and Technology (ASIST).
- Liu, X. & von Brzeski, V. (2009) Computational Community Interest and Comments Centric Analysis Ranking, Proceedings of the iConference
- Liu, X. (2008) Proposal of Document Classification with Word Sense Disambiguation, Proceedings of the iConference