



# 李直旭(ZHIXU LI)

博士、副教授、硕士生导师

地址: 苏州大学校本部理工楼 303

邮箱: [zhixuli@suda.edu.cn](mailto:zhixuli@suda.edu.cn)

主页: <http://ada.suda.edu.cn/staff/ZhixuLi>

电话:+86-17715140107

政治面貌: 中共党员

出生年月: 1983.11

## 教育工作履历

- 04/2014 至今: 苏州大学计算机科学与技术学院, 副教授、硕导
- 06/2013-06/2014: 沙特阿卜杜拉国王科技大学, 博士后研究员
- 10/2009-06/2013: 澳大利亚昆士兰大学, 获计算机博士学位(导师: 周晓方)
- 09/2006-07/2009: 中国人民大学信息学院, 获计算机硕士学位(导师: 杜小勇)
- 09/2002-07/2006: 中国人民大学信息学院, 获计算机学士学位

## 学术访问经历

- 04/2013-06/2013 香港中文大学(香港) 访问学者
- 10/2012-11/2012 阿卜杜拉国王科技大学(沙特) 访问学者
- 09/2011-02/2012 微软亚洲研究院(北京) 访问研究员

## 主要研究方向

- 知识图谱构建与管理 (Knowledge Graph Construction and Management)
- 数据质量, 数据清理 (Data Quality, Data Cleaning)
- 文本挖掘, 信息抽取 (Text Mining, Information Extraction)
- 众包技术, 群体智能 (Crowdsourcing, Swarm Intelligence)

## 主要荣誉奖励

- 2015 年度江苏省“双创博士”
- 2016 年 DASFAA 国际会议最佳论文
- 2015 年全国数据库大会萨师焯最佳论文
- 2015 年 WISE 国际会议最佳论文
- 2014 年度苏州市紧缺高层次引进人才
- 2013 年 WISE 国际会议最佳会议论文奖
- 2009-2014 留学基金委博士生奖学金
- 2009 届中国人民大学优秀毕业生

## 研究生指导情况

### 毕业学生 (4 人)

- M14 杨强, 一作完成 CCF 及核刊论文 5 篇, 拟毕业去向: 出国深造(澳大利亚)
- M14 顾斌斌, 一作完成 CCF A、B 期刊各 1 篇, 拟毕业去向: 出国深造(美国)
- M14 周剑, 一作完成 CCF 论文 2 篇, 拟毕业去向: 杭州娄文科技
- M13 蒋俊, 一作完成 CCF 论文 1 篇, 毕业去向: ORACLE(苏州)

### 在读学生 (10 人)

- M15 胡猛(信息抽取, 在投 CCF B 1 篇), 沈永新(文本解析, 在投 CCF B 1 篇)
- M16 杨佳莉(推荐系统, 在投核刊论文 1 篇), 何莹(关系抽取, 在投 SCI 2 区 1 篇)
- M17 李茂龙(知识图谱构建), 单双利(数据清洗), 王怡婷(深度学习强化学习), 何芙珍(众包质量控制), 胡嘉欣(区块链), 付晨鹏(大图管理与查询优化)

---

## 主持科研项目

- 校方主持（合作申请单位：清华大学，中国人民大学），国家自然科学基金（重点项目），“众包数据库的基础理论与关键技术研究”，320万（2017.01-2020.12）
- 主持，国家自然科学基金（青年项目），“基于互联网海量信息的数据库文本类型数据清洗研究”，24万(2015.01-2017.12)
- 主持，博士后科学基金第九批特别资助项目，“面向大数据应用的信息抽取关键技术研究”，15万（2016.01-2017.12）
- 主持，江苏省“双创博士”项目，“互联网大数据应用信息抽取和挖掘关键技术研究”，15万（2015.01-2016.12）
- 主持，教育部第58批博士后面基金项目，“互联网大数据实体信息识别与关联关系挖掘技术研究”，5万（2015.07-2017.06）
- 主持，江苏省博士后科研资助基金项目，“面向互联网大数据应用的信息抽取关键技术研究”，4万（2015.07-2017.06）

## 指导学生项目

- 江苏省大学生创新创业，省级重点项目，“面向多源房产信息融合的数据爬取、集成、存储与展示技术”，学生：李巡，邵健，沈涛，余金阳，张正齐（2015-2017）
- 江苏省大学生创新创业，省级重点项目，“基于数据融合的交互式二手车推荐系统”，学生：张文凌，牛雷，马疆楠，东野安然，徐昊（2016-2018）

## 大会报告及演讲

- 03/2017 “众包交互机制及查询接口”，清华大学
- 01/2017 “Interactive Data Cleaning Based On Web Knowledge And Crowd”，武汉理工大学
- 08/2016 “数据关联问题及方法概述”，Invited Talk，国家标准化院（北京）
- 04/2013 “Exploiting the Web for Solving Data Quality Problems”，Invited Talk，香港中文大学
- 11/2012 “WebPut: Efficient Web-based Data Imputation”，WISE Conference，塞浦路斯
- 03/2012 “Filling in Missing Attribute Values with Web Data”，DKE Day Talk，昆士兰大学
- 01/2012 “Cleaning Drifting Errors in ProBase”，Invited Talk，中科院软件所
- 05/2011 “Exploiting Web Search Engine For Relation Completion”，DKE Day Talk，昆士兰大学
- 10/2010 “Approximate Membership Localization (AML) for Web-Based Join”，CIKM，多伦多

## 学术兼职与服务

- 计算机学会数据库专委会通讯委员
- 人工智能学会智能服务委员会委员
- ACM SIGMOD CHINA 委员会委员
- IEEE TCDE CHINA 委员会委员
- 国际期刊评审：IEEE TKDE, VLDBJ, WWWJ, NeuroComputing 等
- 国际会议承办 Local Organization Chair: DASFAA2017, APWeb2016
- 国际会议程序委员会委员：DSC, APWEB, CIKM, ADC, BigData, InfoSys, MTAP, FCS, ICPCSEE 等

---

## 主要论文列表（一作及通讯）

其中：SCI 论文 10 篇(其中 TKDE 5 篇)，中文核刊 1 篇，CCF A/B 类会议论文 8 篇  
期刊论文 Journal Papers:

1. Binbin Gu, **Zhixu Li\***, Xiangliang Zhang, An Liu, Guanfeng Liu, Kai Zheng, Lei Zhao and Xiaofang Zhou, "The Interaction between Schema Matching and Record Matching in Data Integration". *IEEE Transaction on Knowledge and Data Engineering (TKDE)* 29(1): 186-199 (2017) (SCI 2 区 , CCF A 类)
2. Qing Xie, Xiangliang Zhang, **Zhixu Li\*** and Xiaofang Zhou, "Optimizing Cost of Continuous Overlapping Queries over Data Streams by Filter Adaption", *IEEE Transaction on Knowledge and Data Engineering (TKDE)* 28(5): 1258-1271 (2016) (SCI, CCF A 类)
3. **Zhixu Li**, Lu Qin, Hong Cheng, Xiangliang Zhang and Xiaofang Zhou, "TRIP: An Interactive Retrieving-Infering Data Imputation Approach". *IEEE Transaction on Knowledge and Data Engineering (TKDE)* 27(9): 2550-2563 (2015) (SCI, CCF A)
4. **Zhixu Li**, Mohamed A. Sharaf, Laurianne Sitbon, Xiaoyong Du and Xiaofang Zhou, "CoRE: A Context-Aware Relation Extraction Method for Relation Completion". *IEEE Transaction on Knowledge and Data Engineering (TKDE)* 26(4): 836-849 (2014) (SCI, CCF A)
5. **Zhixu Li**, Laurianne Sitbon, Liwei Wang, Xiaofang Zhou and Xiaoyong Du, "AML:Efficient Approximate Membership Localization within a Web-Based Join Framework", *IEEE Transaction on Knowledge and Data Engineering (TKDE)* 25(2): 298-310 (2013) (SCI, CCF A)
6. Binbin Gu, **Zhixu Li\***, Qiang Yang, Qing Xie, An Liu, Guanfeng Liu, Kai Zheng, Xiangliang Zhang, "Web-ADARE: A Web-Aided Data Repairing System", *Neurocomputing* 253: 201-214 (2017) (SCI 2 区)
7. **Zhixu Li**, Qiang Yang, An Liu, Guanfeng Liu, Jia Zhu, Jiajie Xu and Min Zhang, "Crowd-Guided Entity Matching with Consolidated Textual Data", *JSCT* (SCI, CCF B, Accepted in June 2017)
8. An Liu, **Zhixu Li\***, Guanfeng Liu, Kai Zheng, Min Zhang, Qing Li, and Xiangliang Zhang, "Efficient Privacy-preserving Task Assignment in Spatial Crowdsourcing", *JSCT* (SCI, CCF B, Accepted in June 2017)
9. **Zhixu Li**, Qiang Yang, An Liu, Guanfeng Liu, Lei Zhao, XiangLiang Zhang, Xiaofang Zhou, "PubTeller: A Compound Paper Recommendation Approach", *World Wide Web Journal (WWWJ)* (SCI, CCF B, Accepted in May 2017)
10. **Zhixu Li**, Mohamed A. Sharaf, Laurianne Sitbon, Shazia Wasim Sadiq, Marta Indulska and Xiaofang Zhou, "A Web-based Approach to Data Imputation". *World Wide Web Journal (WWWJ)* 17(5): 873-897 (2014) (SCI, CCF B)
11. 杨强, **李直旭\***, 蒋俊, 赵朋朋, 刘冠峰, 刘安, 周晓方, "基于非主属性值的实体匹配", *计算机学报*. (Accepted in Jan. 2016) (中文核心期刊)

---

## 国际会议论文 Conference Papers:

1. Binbin Gu, **Zhixu Li\***, Xiangliang Zhang, An Liu, Guanfeng Liu, Kai Zheng, Lei Zhao, Xiaofang Zhou, “The Interaction Between Schema Matching and Record Matching in Data Integration (Extended Abstract)”. *ICDE 2017*, May 2017 (EI, CCF A)
2. **Zhixu Li**, Lu Qin, Hong Cheng, Xiangliang Zhang and Xiaofang Zhou, "TRIP: An Interactive Retrieving-Inferring Data Imputation Approach (Extended Abstract)". *ICDE 2016*, May 2016, Helsinki, Finland (EI, CCF A)
3. Qiang Yang, **Zhixu Li\***, Binbin Gu, An Liu, Guanfeng Liu, Pengpeng Zhao, Lei Zhao: CTextEM: Using Consolidated Textual Data for Entity Matching. *DASFAA 2016*: 117-132 (EI, CCF B 类)
4. Jian Zhou, **Zhixu Li\***, Binbin Gu, Qing Xie, Jia Zhu, Xiangliang Zhang, Guoliang Li: CrowdAidRepair: A Crowd-Aided Interactive Data Repairing Method. *DASFAA 2016*: 51-66 (EI, CCF B)
5. Jian Zhou, **Zhixu Li\***, Qiang Yang, Jun Jiang, An Liu, Guanfeng Liu, Lei Zhao and Jia Zhu, “HouseIn: A Housing Rental Platform with Non-Redundant Information Integrated from Multiple Sources“. *In Proceedings of the 17th Asia Pacific Web Conference (APWeb)*, September 2015, Guangzhou, China(EI, CCF C)
6. **Zhixu Li**, Xiangliang Zhang, Qing Xie, Hai Huang, Jia Zhu and Xiaofang Zhou, “Addressing Instance Ambiguity in Web Harvesting”. *In Proceedings of the 18th International Workshop on the Web and Databases (WebDB)*, May 2015, Melbourne, Australia (EI, CCF C)
7. Qiang Yang, **Zhixu Li\***, Jun Jiang, Pengpeng Zhao, Guanfeng Liu, An Liu and Zhu Jia, “NokeaRM: “Employing Non-Key Attributes in Record Matching”. *In Proceedings of the 16th International Conference on Web-Age Information Management (WAIM 2015)*, June 2015, Qingdao, China (EI , CCF C)
8. Jun Jiang, **Zhixu Li\***, Qiang Yang, Pengpeng Zhao, Guanfeng Liu and Lei Zhao, “SmartInt: A Demonstration System for the Interaction between Schema Mapping and Record Matching”. *In Proceedings of the 16th International Conference on Web-Age Information Management (WAIM 2015)*, June 2015, Qingdao, China (EI , CCF C 类) (DEMO)
9. **Zhixu Li**, Hongsong Li, Haixun Wang, Yi Yang, Xiangliang Zhang and Xiaofang Zhou, “Overcoming Semantic Drift in Information Extraction”. *In Proceedings of the 17th International Conference on Extending Database Technology (EDBT 2014)*, March, 2014, Greece (EI, CCF B)
10. **Zhixu Li**, Shuo Shang, Qing Xie, Xiangliang Zhang, “Cost Reduction for Web-Based Data Imputation”. (*DASFAA 2014*) (EI, CCF B)
11. **Zhixu Li**, Mohamed Sharaf, Laurianne Sitbon and Xiaofang Zhou, “WebPut: Efficient Web-based Data Imputation”. *In Proceedings of the 13th International Conference on Web Information System Engineering (WISE 2012 Best Paper Award)* (EI, CCF C)
12. **Zhixu Li**, Laurianne Sitbon and Xiaofang Zhou, “Learning-based Relevance Feedback for Web-based Relation Completion”. *In Proceedings of the 20th ACM International Conference on Information and Knowledge Management (CIKM 2011)*, October 2011, Glasgow, UK (EI, CCF B)

- 
13. **Zhixu Li**, Laurianne Sitbon and Xiaofang Zhou, “PartSS: An Efficient Partition-based Filtering for Edit Distance Constraints”. *In Proceedings of the 22nd Australasian Database Conference (ADC 2011)*, Jan 2011, Perth, Australia (EI)
  14. **Zhixu Li**, Laurianne Sitbon, Liwei Wang, Xiaofang Zhou and Xiaoyong Du, “Approximate Membership Localization (AML) for Web-Based Join”. *In Proceedings of 19th ACM International Conference on Information and Knowledge Management (CIKM 2010)*, October 2010, Toronto, Canada (EI, CCF B)
  15. **Zhixu Li**, Xiaoyong Du, Hongyan Liu, Jun He and Xiaofang Zhou, “IRank: A Term-based Innovation Ranking System for Conferences and Scholars”. *The Joint International Conferences on Asia-Pacific Web Conference and Web-Age Information Management (APWEB/WAIM 2009)* (EI, CCF C)
  16. **Zhixu Li**, Pei Li, Wei Wei, Hongyan Liu, Jun He, Tao Liu and Xiaoyong Du, “AutoPCS: A Phrase-based Text Categorization System for Similar Texts”. *The Joint International Conferences on Asia-Pacific Web Conference and Web-Age Information Management (APWEB/WAIM 2009)* (EI, CCF C)