

# Binbin Gu

Department of Computer Science, University of California, Irvine

binbing@uci.edu | +1 (831) 400-9628  
<https://tuzijun111.github.io/>

---

EDUCATION	<b>University of California, Irvine, USA</b> <i>Ph.D.</i> , Information and Computer Science Advisor: Faisal Nawab	June 2024 ( <i>Expected</i> )
	<b>University of California, Santa Cruz, USA</b> <i>Ph.D. student</i> , Computer Science and Engineering Advisor: Faisal Nawab	December, 2020
	<b>Soochow University, China</b> <i>Masters</i> in Software Engineering	June, 2017
	<b>Qilu University of Technology, China</b> <i>Bachelors</i> in Information and Computing Science	June, 2014
RESEARCH EXPERIENCE	<b>University of California, Santa Cruz</b> Solving the balanced partitioning problem in distributed systems under dynamic scenarios (with machine learning techniques). Efficient dynamic record linkage with machine learning models	September, 2019-Present
	<b>Soochow University</b> Data fusion on big data Data Imputation and truth finding from the web	July, 2014-June, 2019
EMPLOYMENT EXPERIENCE	<b>University of California, Santa Cruz</b> Teaching Assistant. Teaching the course: Database Systems	April, 2020-June, 2020
	<b>Soochow University, China</b> Research Assistant. Working on data cleaning, data fusion, and truth discovery together with Prof. Zhixu Li and Prof. Xiaofang Zhou	July, 2017-June, 2019
	<b>King Abdullah University of Science and Technology</b> Internship in King Abdullah University of Science and Technology (KAUST). Working on data fusion project with Prof. Xiangliang Zhang	April, 2016-October, 2016
SOFTWARE SKILLS	<b>Programming:</b> Java, Python, C++, Jupyter <b>Other:</b> L <sup>A</sup> T <sub>E</sub> X	<b>Databases:</b> PostgreSQL, MySQL
INTERESTS	<b>Sports:</b> Basketball, Swimming, Hiking	
PUBLICATIONS	<ol style="list-style-type: none"><li>Dawei Chen, Zhixu Li, <b>Binbin Gu</b>, An Liu, Lei Zhao, Zhigang Chen, Guoping Hu: Multimodal Named Entity Recognition with Image Attributes and Image Knowledge. DASFAA (Accepted in January-2020)</li><li>Jia Chen, <b>Binbin Gu</b>, Zhixu Li, Pengpeng Zhao, An Liu, Lei Zhao: SAEA: Self-Attentive Heterogeneous Sequence Learning Model for Entity Alignment. DASFAA (1) 2020: 452-467</li><li>Lei Niu, <b>Binbin Gu</b>, Zhixu Li, Wei Chen, Ying He, Zhaoyin Zhang, Zhigang Chen: Fine-grained Entity Typing for Relation-Sparsity Entities. DASFAA (2) 2020: 141-157</li><li><b>Binbin Gu</b>, Zhixu Li, An Liu, Jiajie Xu, Lei Zhao, Xiaofang Zhou: Improving the Quality of Web-based Data Imputation with Crowd Intervention. IEEE Trans. Knowl. Data Eng. doi: 10.1109/TKDE.2019.2954087</li></ol>	

5. Jiapeng Li, Wei Chen, **Binbin Gu**, Junhua Fang, Zhixu Li, Lei Zhao: Measuring Semantic Relatedness with Knowledge Association Network. DASFAA (1): 676-691 (2019)
6. Zhixu Li, Ying He, **Binbin Gu**, An Liu, Hongsong Li, Haixun Wang, Xiaofang Zhou: Diagnosing and Minimizing Semantic Drift in Iterative Bootstrapping Extraction. IEEE Trans. Knowl. Data Eng. 30(5): 852-865 (2018)
7. **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)
8. **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. IEEE Trans. Knowl. Data Eng. 29(1): 186-199 (2017)
9. **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 : 33-34 (2017)
10. Jian Zhou, Zhixu Li, **Binbin Gu**, Qing Xie, Jia Zhu, Xiangliang Zhang, Guoliang Li: CrowdAidRepair: A Crowd-Aided Interactive Data Repairing Method. DASFAA (1) : 51-66 (2016)
11. Qiang Yang, Zhixu Li, **Binbin Gu**, An Liu, Guanfeng Liu, Pengpeng Zhao, Lei Zhao: CTextEM: Using Consolidated Textual Data for Entity Matching. DASFAA (1) : 117-132 (2016)

ACADEMIC  
SERVICE

PC memeber of CCKS 2019 and CCKS 2020, External Reviewer of CCF Big Data 2015, DSC 2016, CIKM 2016, CCF Big Data 2016, and ADC 2017, ACM TURC 2019.