



高恺

中共党员

电话: (+86)130-5157-3625

邮箱: gaokai19@pku.edu.cn

出生年月: 1999年7月

主页: <https://gaokai320.github.io/>

GitHub: <https://github.com/gaokai320>



北京大学
PEKING UNIVERSITY

教育经历

北京大学 软件工程	博士	软件与微电子学院, 导师: 周明辉、谢冰	2019年09月 – 2024年06月
北京大学 经济学	双学位	国家发展研究院	2017年09月 – 2019年06月
北京大学 计算机科学与技术	本科	信息科学技术学院	2015年09月 – 2019年06月

研究方向

开源软件数据分析: 开源软件供应链、开源生态、软件仓库挖掘、软件演化与维护

荣誉奖励

- 奖学金: 国家励志奖学金、北京大学方正奖学金、中国石油奖学金、新鸿基地产郭氏基金-北京大学双学位奖学金; 北京大学实践公益奖、北京大学学习优秀奖*3
- 学生工作: 北京大学研究生代表大会常务代表委员会立法监察部部门负责人
- 其他: 中国软件大会2020, 2021, 2022嘉宾; 2023 ICSE (CCF A会议) AE track PC; JSEP (CCF B期刊)审稿人

发表论文

1. On the Variability of Software Engineering Needs for Deep Learning: Stages, Trends, and Application Types (TSE 2022, CCF A, 第一作者)
2. 开源软件供应链研究综述——一个整体性的视角 (软件学报 2023, CCF 中文A, 第一作者)
3. An Exploratory Study of Deep Learning Supply Chain (ICSE 2022, CCF A, 第二作者)
4. Understanding and Remediating Open-Source License Incompatibilities in the PyPI Ecosystem (ASE 2023, CCF A, 第三作者)
5. Demystifying Software Release Note Issues on GitHub (ICPC 2022, CCF B, 第四作者, 杰出论文奖)
6. Characterize Software Release Notes of GitHub Projects: Structure, Writing Style, and Content (SANER 2023, CCF B, 第三作者)
7. Characterizing Deep Learning Package Supply Chains in PyPI: Domains, Clusters, and Disengagement (TOSEM 2023, CCF A, 一作在投)
8. PyRadar: Towards Automatically Retrieving and Validating Source Code Repository Information for PyPI Packages (FSE 2024, CCF A, 一作在投)
9. Demystifying Software Release Note Issues and Practices on GitHub (EMSE 2023, CCF B, 三作在投)

项目经历

开源软件开发 (国家杰青项目) 2019年1月 – 至今
作为项目参与者, 领导开源软件供应链的可持续性、低可见性和可追溯性等课题的研究, 参与开源软件供应链中版本发布纪要和许可证冲突等课题的研究。

云计算和大数据开源社区生态系统 (国家重点研发计划项目) 2019年06月 – 2020年5月
作为项目主要参与者, 与田纳西大学Audris Mockus教授团队合作构建World of Code开源大数据基础设施 (300TB数据量), 并作为唯一负责人在国内维护其镜像。

可信开源社区项目 (与华为合作项目) 2020年11月 – 2021年11月
作为项目主要课题负责人, 构建深度学习软件供应链, 并分析其结构、应用类型和演化因素, 支持国内深度学习框架社区的发展, 在华为内部评审优秀。

专业技能

- 编程: Python, Java, Shell; Linux, Git, Docker; Pandas, Matplotlib, Numpy; MongoDB, MySQL; Scikit-Learn, PyTorch; attention, transformer
- 办公: PowerPoint, Excel, Word; 英语六级