先端科学技術研究科　修士論文要旨

所属研究室 (主指導教員)	ソフトウェア工学 (松本　健一　(教授))		
学籍番号	2311196	提出日	令和 7 年 1 月 21 日
学生氏名	中野　剛志		
論文題目	The Role of GitHub Biographies: Impact on Pull Request Acceptance and Recruitment Decisions		

要旨

This thesis examines the linguistic properties of GitHub biographies (bios) and their influence on pull request acceptance, recruitment outcomes, and social biases introduced by large language models (LLMs), aiming to uncover the roles of GitHub bios. In the digital age, self-presentation on platforms like GitHub has become a crucial factor influencing developer recruitment, collaboration, and recognition. However, there is limited work on studying how these social profiles signal competence, ability, and other social cues exhibited by GitHub users. Using a dataset of 45,017 GitHub bios, this study analyzed the contents of GitHub bio through textual and statistical methods to identify linguistic trends and conducted counterfactual experiments to uncover biases embedded in LLMs. The findings uncover that bios highlighting personal motivation and a willingness to contribute are more likely to secure pull request acceptance, while those emphasizing analytical and motivational language tend to perform better in recruitment scenarios. Moreover, this study demonstrates how LLMs, such as GPT-4, exhibit geographical bias when recruiting or assigning roles to developers based on their GitHub bios. These results underscore the influence of GitHub bios on professional outcomes and provide valuable insights into optimizing developers' self-presentation. The biases uncovered when LLMs recruit developers using GitHub bios offer a deeper understanding of LLM behavior and have important implications for mitigating social bias embedded in LLMs.