

# Graduate School of Science and Technology Master's Thesis Abstract

Laboratory name (Supervisor)	Software Engineering (Ken-ichi Matsumoto (Professor ))		
Student ID	1911423	Submission date	2021 / 7 / 15
Name	WATTANAKRIENGKRAI SUPATSARA		
Thesis title	An Exploratory Study into the Linkages Between GitHub Repositories and Academic Papers GitHubリポジトリと学術論文との関連性に関する調査		
Abstract			
<p>Traceability between published scientific breakthroughs and their implementation is essential, especially in the case of open-source scientific software which implements bleeding-edge science in its code. However, aligning the link between GitHub repositories and academic papers can prove difficult, and the current practice of establishing and maintaining such links remains unknown. This thesis investigates the role of academic paper references contained in these repositories. I conduct a large-scale study of 20 thousand GitHub repositories that make references to academic papers. I use a mixed-methods approach to identify public access, traceability, and evolutionary aspects of the links. Although referencing a paper is not typical, I find that a vast majority of referenced academic papers are public access. These repositories tend to be affiliated with academic communities. More than half of the papers do not link back to any repository. I find that academic papers from top-tier SE venues are not likely to reference a repository, but when they do, they usually link to a GitHub software repository. In a network of arXiv papers and referenced repositories, I find that the most referenced papers are (i) highly-cited in academia and (ii) are referenced by repositories written in different programming languages. From the evolutionary aspect, I find very few changes of papers being referenced and links to them.</p>			