

# Development of Sustainable Aggregating Services for an Improved FAIR Research Software Ecosystem

## Authors:

Maxence Azzouz-Thuderoz<sup>1\*</sup>, Madhurima Deb<sup>2</sup> and Moritz Schubotz<sup>2</sup>

<sup>1</sup>maxence.azzouz-thuderoz@fiz-karlsruhe.de, FIZ Karlsruhe Leibniz Institute for Information Infrastructure, Berlin, Germany.

<sup>2</sup>FIZ Karlsruhe Leibniz Institute for Information Infrastructure, Berlin, Germany.

## Abstract:

Version control systems (VCS), among the various components shaping the research software landscape, have emerged as essential tools in the software development process. Platforms like GitHub and GitLab have gained exponential popularity among scholars by facilitating collaborative development and contributing significantly to the adoption of software engineering practices in academia. These platforms have also made research software more compliant with the FAIR principles (findability, accessibility, interoperability, and reusability). A notable trend is the substantial reliance of academic researchers on VCS platforms, as evidenced by 127,529 Uniform Resource Identifiers (URIs) cited in 385,817 arXiv publications referencing software hosted on platforms like GitHub and GitLab. Similarly, 16,690 of the 44,632 software tools indexed by the mathematical software database swMATH link back to these platforms. However, despite their widespread use, these platforms lack suitable recommender systems for related work, limiting the discoverability of research software.

Moreover, concerns have arisen regarding the inability to host code in desired locations, prompting universities and research institutions to establish their own repository services. This has driven the development and adoption of on-premise solutions. The diverse use of repository platforms—ranging from public GitHub repositories to on-premise GitLab deployments and standalone Git servers—presents additional challenges in monitoring the research software landscape. This diversity hinders the discoverability and reproducibility of research outputs, complicating the curation of metadata and long-term adherence to the FAIR principles.

In response to these challenges, aggregator platforms like swMATH play a crucial role in ensuring the sustainability and accessibility of research software. By curating and organizing software metadata, these platforms enhance the reliability and discoverability of software source code. However, their focus on specific communities often limits their broader adoption, restricting their potential to reach users beyond niche groups.

A promising solution is the establishment of an NFDI Research Software Marketplace, envisioned as an interoperable platform integrated with standardized software metadata catalogs like swMATH. This marketplace could serve as a centralized hub for discovering, accessing, and collaborating on research software projects, regardless of the repository services used. Crucially, it would also support advanced recommender system services tailored to the specific needs of research software developers, closing the gap in discoverability addressed by the FAIR criteria.

**Keywords:** Research software, FAIR, aggregator, NFDI, recommender systems