Contribution ID: **15**   Type: **Talk**

Exploring Computational Reproducibility in Jupyter Notebooks: Insights and Challenges

Tuesday, March 12, 2024 12:00 PM (20 minutes)

Reproducible research emphasizes the importance of documenting and publishing scientific results in a manner that enables others to verify and extend them. In this talk, we explore computational reproducibility within the context of Jupyter notebooks, presenting insights and challenges from our study. We will present the key steps of the pipeline we used for assessing the reproducibility of Jupyter Notebooks. In our study, we analyzed the notebooks extracted from GitHub repositories associated with publications indexed in the biomedical literature repository PubMed Central. Our process involved identifying the notebooks by mining the full text of publications, locating them on GitHub, and attempting to rerun them in an environment closely resembling the original. We documented reproduction success and exceptions and explored relationships between notebook reproducibility and variables related to the notebooks or publications, including results related to programming languages, notebook structure, naming conventions, modules, dependencies, etc. Furthermore, we will discuss the common issues and practices, identify emerging trends, and explore potential enhancements to Jupyter-centric workflows. Through this comprehensive examination, we aim to provide actionable insights and practical strategies for researchers striving to enhance the reproducibility of their work within the Jupyter notebook ecosystem and contribute to the ongoing dialogue surrounding reproducibility and computational methodologies in scientific research.

Type of Poster

Presenter: SAMUEL, Sheeba (Friedrich Schiller University)

Session Classification: Talks

Track Classification: Vortrag