

Abstract Jupyter4NFDI (Initialisation Phase)

Submission Round 4 - 02/2024

Name of proposed Basic Service

A central JupyterHub providing access to various software stacks and computing resources across the NFDI consortia

Acronym of the proposed Basic Service

Jupyter4NFDI

Lead institution

Forschungszentrum Jülich GmbH, 52425 Jülich

Supported by Section Common Infrastructures

Summary of the proposal in English and German

The widespread usage of Jupyter notebooks across scientific disciplines highlights their fundamental importance in current research practices. However, their deployment across various NFDI consortia currently occurs through individual JupyterHubs, resulting in access barriers to computational and data resources. Only some of the services are widely available, and others are barricaded within VPNs, creating a fragmented landscape that complicates access. Our proposal aims to unify these efforts by offering a centralized service. We will simplify access, significantly improve the user experience, and extend the reach of Jupyter to a broader audience within the NFDI and beyond. We will integrate with IAM4NFDI to govern access to the central service and external resources, with future integrations planned with upcoming basic services such as Multi-Cloud or nfdi.software. Moreover, HPC resources from the GCS and the NHR Alliance will be accessible via the centralized Jupyter service, which lowers HPC entry barriers for a wide range of users. Importantly, the centralized system will not only simplify access but also support the import of projects along with their necessary dependencies, fostering an ecosystem conducive to creating reproducible FAIR Digital Objects (FDOs), possibly along with notebook identifiers supported by PID4NFDI. By ensuring a consistent user experience across all consortia, we will promote continued innovation. Our commitment is to a collaborative and centralized approach that brings Jupyter notebooks within everyone's reach and democratizes access to the infrastructures for reproducible computational research within the NFDI and beyond.

Die weit verbreitete Nutzung von Jupyter-Notebooks in allen wissenschaftlichen Disziplinen unterstreicht ihre grundlegende Bedeutung in der aktuellen Forschungspraxis. Allerdings erfolgt ihr Einsatz in verschiedenen NFDI-Konsortien derzeit über einzelne JupyterHubs, was zu Zugangsbarrieren zu Rechen- und Datenressourcen führt. Nur einige der Dienste sind weithin verfügbar, andere sind innerhalb von VPNs verbarrikadiert, wodurch eine fragmentierte Landschaft entsteht, die den Zugang erschwert. Unser Vorhaben zielt auf eine Vereinheitlichung der Bemühungen ab, indem wir einen zentralisierten Dienst anbieten. Wir werden den Zugang vereinfachen, die Benutzerfreundlichkeit deutlich verbessern und die Reichweite von Jupyter auf

ein breiteres Publikum innerhalb der NFDI und darüber hinaus ausdehnen. Wir werden IAM4NFDI integrieren, um den Zugriff auf den zentralen Dienst und externe Ressourcen zu regeln, wobei zukünftige Integrationen mit kommenden Basisdiensten wie Multi-Cloud oder nfdi.software geplant sind. Darüber hinaus werden HPC-Ressourcen des GCS und des NHR-Verbunds in den zentralen Jupyter-Dienst eingebunden, wodurch die HPC- Einstiegshürden für eine breite Nutzerschaft reduziert werden. Wichtig ist, dass das zentralisierte System nicht nur den Zugang vereinfacht, sondern auch den Import von Projekten mit ihren notwendigen Abhängigkeiten unterstützt und so ein Ökosystem fördert, das die Erstellung reproduzierbarer FAIR Digital Objects (FDOs) ermöglicht, möglicherweise zusammen mit Notebook-Identifikatoren, die von PID4NFDI unterstützt werden. Durch die Gewährleistung einer konsistenten Benutzererfahrung über alle Konsortien hinweg werden wir kontinuierliche Innovation fördern. Wir setzen uns für einen kollaborativen und zentralisierten Ansatz ein, der Jupyter-Notebooks zur rechnergestützten, reproduzierbaren Forschung für jedermann innerhalb der NFDI und darüber hinaus zugänglich macht.