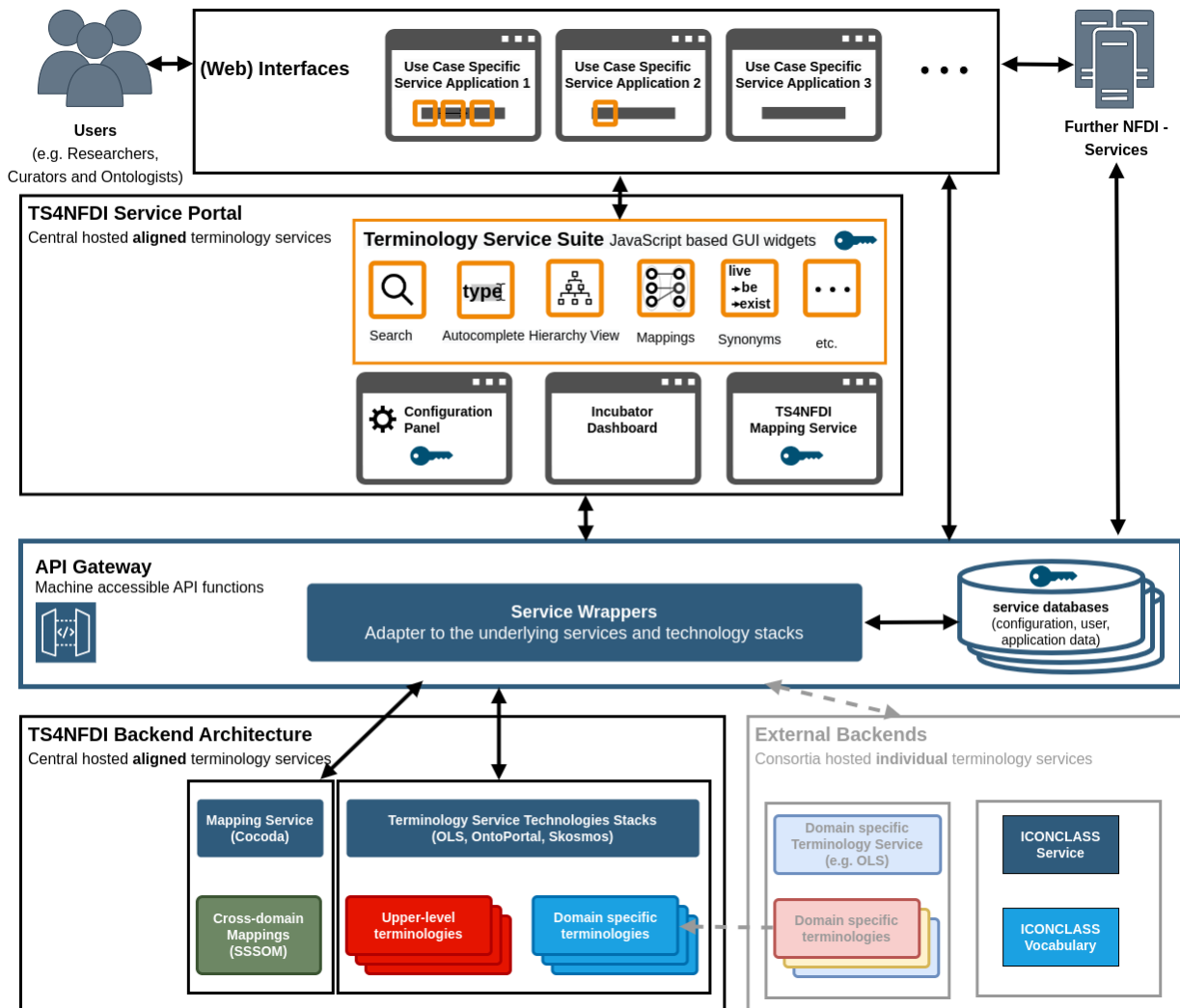# Proposal

Terminology Services 4 NFDI

TS4NFDI

A cross domain service for the provision, curation, development, harmonisation, and mapping of terminologies

Integration Phase

1. General Information

Name of proposed Basic Service (in English)

Terminology Services 4 NFDI

Acronym of the proposed Basic Service

TS4NFDI

Service "subtitle" explaining key functionality

Corresponding NFDI Section

Section (Meta)data, Terminologies, Provenance

Lead institution

TIB – Leibniz Information Centre for Science and Technology (lead)

Welfengarten 1 B, 30167 Hannover

ZB MED – Information Centre for Life Sciences (co-lead)

Gleueler Straße 60, 50931 Cologne

Name of lead institution principal investigator

Dr. Oliver Koepler (oliver.koepler@tib.eu) (lead)

Roman Baum (baum@zbmed.de) (co-lead)

Participating institutions

Principal Investigator	Institution, location	Contact E-mail	Member in [consortium][1]	Funding requested
Dr. Oliver Koepler, Kolja Bailly	TIB - Leibniz Information Centre for Science and Technology, Hannover	oliver.koepler@tib.eu, kolja.bailly@tib.eu	NFDI4Chem, NFDI4Culture	yes
Roman Baum	ZB MED - Information Centre for Life Sciences, Bonn	baum@zbmed.de	NFDI4Health, FAIRagro	yes
Dr. Naouel Karam, Prof. Adrian Paschke	Institut für Angewandte Informatik e.V. (InfAI), Leipzig	karam@infai.org, paschke@infai.org	NFDI4Biodiversity	yes
Dr. Jakob Voß	Verbundzentrale des GBV (VZG), Göttingen	jakob.voss@gbv.de	NFDI4Objects, NFDI4Memory	yes
Timo Mühlhaus	RPTU University of Kaiserslautern-Landau	timo.muehlhaus@rptu.de	DataPLANT	yes
Etienne Posthumus	FIZ Karlsruhe	etienne.posthumus@partners.fiz-karlsruhe.de	NFDI4Culture	no

Table 1: List of participating institutions

---

[1] Name one DFG-consortium the institution is or has a route to become a member of and through which funds should be appropriated if this proposal is approved.

Initialisation Phase

The initialisation phase started for ZB MED in 11/2023. InfAI and TIB started in 1/2024. The estimated end of the initialisation phase will be in 12/2024.

Planned duration of the project

2 years

Significance

An overarching research data management across all domains is built upon metadata describing research data and their generation. To make sure that disciplines can create and share a common understanding of the concepts and relations used to describe data, we need terminologies as a formal representation of domain knowledge. Terminology Services 4 NFDI (TS4NFDI) provides unified access to terminologies across all disciplines. TS4NFDI integrates or connects existing terminology services (TSs) and, most importantly, reduces the efforts to integrate terminologies in any service or tool to manage research data. It provides a central access point to all terminologies and a suite of widgets that can be added to existing applications of the users. TS4NFDI supports cross-domain application, harmonisation, and curation of terminologies.

Summary of the proposal in English and German

Terminologies are essential for semantically linking research data across disciplines, ensuring conceptual alignment even with different nomenclatures. Standardised terminologies promote interoperability and data integration, making data understandable and reusable across various sources and fields. Terminologies are used in services and tools for annotating and linking research, necessitating the development of several terminology services (TSs) for publishing, describing, and curating terminologies. Aligning, managing, and developing these TSs is challenging but crucial for consensus building and harmonisation across disciplines. Terminology Services 4 NFDI (TS4NFDI) aims to standardise and harmonise terminology management within NFDI, facilitating consensus and interoperability across disciplines to achieve a shared knowledge representation framework. During the integration phase, we will consolidate the service design of the API Gateway, Service Wrappers, Mapping Service, and reusable GUI widgets. We will extend the coverage of disciplines and integrate additional terminology sources and services to meet the requirements gathered during the initialization phase.

Terminologien sind entscheidend für die semantische Verknüpfung von Forschungsdaten über Disziplinen hinweg und sorgen für konzeptuelle Übereinstimmung, selbst bei unterschiedlicher Nomenklatur. Standardisierte Terminologien fördern die Interoperabilität und Datenintegration, wodurch Daten verständlich und wiederverwendbar über verschiedene Quellen und Fachgebiete hinweg werden. Terminologien werden in Diensten und Werkzeugen zur

Annotation und Verknüpfung von Forschung eingesetzt, was die Entwicklung mehrerer Terminologiedienste (TSs) zur Veröffentlichung, Beschreibung und Pflege von Terminologien erfordert. Die Ausrichtung, Verwaltung und Entwicklung dieser TS ist herausfordernd, aber entscheidend für den Konsens und die Harmonisierung über Disziplinen hinweg. Der Basisdienst Terminology Services 4 NFDI (TS4NFDI) zielt darauf ab, das Terminologiemanagement innerhalb der NFDI zu standardisieren und zu harmonisieren, um Konsens und Interoperabilität über Disziplinen hinweg zu erleichtern und eine gemeinsame Wissensrepräsentation zu erreichen. Während der Integrationsphase werden wir das Servicedesign des API Gateways, der Service Wrappers, des Mapping-Dienstes und wiederverwendbarer GUI-Widgets konsolidieren. Wir werden die Abdeckung der Disziplinen erweitern und zusätzliche Terminologiequellen und -dienste integrieren, um die während der Initialisierungsphase gesammelten Anforderungen zu erfüllen.

2. Summary of Initialisation Phase Results (max 3 pages)

2.1. Change in Background and Motivation since the start of the Initialisation Phase

Data harmonisation and integration efforts require standards for mapping between different representations of similar entities in different knowledge and data management systems. A strong community need for a mapping service has been identified through meetings with various consortia, section meetings, workshops, and hacking sessions. In the proposal, TS4NFDI has planned to integrate the EMBL-EBI Ontology Xref Service (OxO) [1, 2] as a promising candidate for a mapping service integrating Simple Standard for Sharing Ontological Mappings (SSSOM) [3]. However, despite initial announcements of collaboration between the OxO team and the SSSOM team, there has been no visible progress in extending the OxO service to OxO2 and to include SSSOM. We, therefore, chose to use another existing mapping service, the Cocoda mapping tool [4], which is used within the NFDI4Objects consortium and could be extended to support SSSOM.

2.2. Results of Initialisation Phase

2.2.1. Interim report on requirements for finalisation of the initialisation phase

A comprehensive summary can be found in the Open Project attachment to this proposal.

Documentation of requirements analysis (D1.4.1)	Due: **3** months after start date	Percent finished: **70%**
Status:	survey | **analysis** | report / planned | **running** | finished	
Overview of relevant outcomes:	D1.2 - First Results of the Requirements Analysis D1.3 - Usage of terminology services @NFDI	
Outlook:	The final analysis of the survey is currently underway and is expected to be completed in August 2024. The subsequent analysis report is expected to be completed and published in September 2024.	

Documentation of software evaluation (D1.4.2)	Due: **6** months after start date	Percent finished: **90%**
Status:	survey | analysis | **report** / planned | **running** | finished	
Overview of relevant outcomes:	D1.1 - Report on TSs and technologies D2.3 - API specification for a selected set of functions	
Outlook:	The last details of the report will be written and published in September 2024.	

Documentation of service design (D1.4.3)	Due: **6** months after start date	Percent finished: **100%**
Status:	survey | analysis | **report** / planned | running | **finished**	
Overview of relevant outcomes:	D2.1 - Technical documentation D2.3 - API specification; OpenAPI documentation	
Outlook:	The task is finished.	

Service prototype (D1.4.4)	Due: **12** months after start date	Percent finished: **70%**
Status:	**prototype** / planned | **running** | finished	
Overview of relevant outcomes:	D3.1 - API Gateway is available D3.3 - A TS user interface based on the TSS V1 is available	
Outlook:	The general architecture of the gateway is implemented for the complex search functionality. This architecture could be easily extended to encompass additional functionalities by the end of the initialization phase in Q4/24. A first TS user interface will be published in	

Service piloting and user testing (D1.4.5)	Due: **12** months after start date	Percent finished: **50%**
Status:	**piloting** / planned | **running** | finished	
Overview of relevant outcomes:	D4.1 - Demonstrator(s) with use cases implemented: Metadata Annotation Workbench; preVIEW	
Outlook:	The evaluation of use cases for the integration phase is running and the report will be available by the end of the initialisation phase Q4/2024 (D4.2) just as the recommendations for the integration phase (D4.3).	

2.2.2. Other

Contributions to CoRDI and UC4B2024 [5, 6]. Community workshops [7, 8, 9].

2.3. Update on the Technical Readiness Level (TRL) of the proposed Basic Service

The underlying technologies and open source solutions OLS, OntoPortal, and Skosmos have been in operation in several TSs for years (e.g. OLS EBI [10], BioPortal [11], AGROVOC [12] or TheSoz [13]) and are classified as TRL 9. The integrated NFDI terminology services of the TS4NFDI architecture are operational in the NFDI infrastructure and federation with other services for a period of up to 24 months and can be classified as TRL 8 with regards to OLS systems with separate backend and frontend components (NFDI4Chem, NFDI4Health, NFDI4Ing). The mapping application Cocoda and terminology service (TS) DANTE are both used in an operational environment outside of NFDI, so their current TRL is TRL 9 for general use and TRL 7 for use in TS4NFDI. The targeted TRL by the end of the integration phase is TRL 9. An early version of the terminology service widgets was released in October 2022 and has been in active development since then. They are being used and tested in NFDI4Health services in production and therefore have a TRL of 6. By the end of the initialisation phase, the Terminology Service Suite (TSS) widgets will have a TRL of 7. The first version of the API Gateway has been deployed on Month 6 of the initialisation phase and has at present a TRL of 4, by the end of the current phase we aim for a TRL 7. The annotation service ANTELOPE [14] is used in an operational environment inside and outside of NFDI and can be considered as TRL 7. Swate [15], with its distinct frontend and backend components, is highly adaptable to various technical environments. The system prototype has been demonstrated in operational

settings (TRL 7), particularly within NFDIs for the life science sector (DataPLANT, FAIRagro, NFDI4Bioimage), where it is primarily utilised for process annotation, though not exclusively so.

3. Working Concept for the development of the Basic Service

In the integration phase, we build upon the proof of concept of the architecture and the resulting pilot implementations to reach our goal of offering a basic service to provide, access, and curate terminologies from the various terminology sources of NFDI for all NFDI communities in one centralised place. We improve the maturity of the service architecture and strengthen the interplay of its components. We further extend the range of scientific disciplines, the number of terminology sources, and further tailor the services to user needs. The Mapping Service to be integrated will support curation and harmonisation workflows for the NFDI.

3.1.   Service Integration Concept

The integration phase will be driven by expanding the scope of TS4NFDI to incorporate new features, terminologies, and terminology sources to align with the needs of the scientific user communities. We will implement a **Service Portal** for users and communities to customise key aspects of the terminology services. We will integrate IAM4NFDI services to allow for personalised access in the portal. The portal will enable the configuration of entity sets and terminology collections as a subset of all available terminologies, tailored to the needs of the users. The Service Portal will further provide access to the Mapping Service and the Terminology Service Suite, including configuration options for the widgets. Mappings between terminologies are essential for achieving and enhancing interoperability within the NFDI consortia and across disciplines. Despite efforts to harmonise terminology development, mapping remains essential due to the inevitable emergence of domain terminologies, each with its own terms and nomenclature. This situation can result in overlapping or conflicting terminologies. We will provide a **Mapping Service** to create, curate, and access simple mappings between terms of terminologies based on Simple Standard for Sharing Ontology Mappings (SSSOM) via the integration of the Cocoda service. Addressing standardisation and harmonisation efforts for terminology curation, we will support **Terminology Curation Workflows** within the TS4NFDI components to foster the quality and sustainability of terminologies. These workflows will enable terminology users and curators to manage terminology suggestions, term requests, issue reports, and documentation of curation. We will combine curation workflows with the Swate data annotation tool to allow instant application of terms suggested. Extending the scope of TS4NFDI to cover more disciplines and their terms, as well as related existing terminology services or sources. This will **extend** the services in the backend architecture and **the API Gateway**, which will connect to the ICONCLASS service as

an example of a linked external terminology source. The Mapping Service will be a further addition to the Backend Architecture. We will **extend the initial version of the Terminology Service Suite (TSS)** which provides basic terminology capabilities integrated through widgets by additional functionality, and addressing needs identified together with the consortia, to ensure that it can support the broad range of tasks and use cases of the NFDI consortia. The Terminology Service Suite will be extended to provide functionalities of the ANTELOPE tool to enable seamless access to terminology search and visualisation, entity linking, and image recognition, especially for the cultural and humanities knowledge domains integrating ICONCLASS and lobid-gnd. Bridging to further knowledge management systems we aim to extend the TSS to provide widgets to be integrated in Wikibase supporting various NFDI consortia using Wikibase instances and also to connect to the new KGI4NFDI basic service. We aim to create an informed and **engaged community** around TS4NFDI, ensuring that the services are widely integrated and continually improved through community feedback and collaboration. We will introduce incubator projects to achieve the integration of TS4NFDI services in a use case-specific manner. Regular feedback from the incubators will ensure that integration phase developments are aligned with the needs of the different consortia and serve as models for other domains, demonstrating the applicability of TS4NFDI and encouraging further integration. An incubator dashboard linked to the service portal will facilitate transparency and showcase success stories. We will provide documentation and training material and continue training sessions, workshops, and hackathons to gather and discuss requirements and use cases. The **IAM4NFDI** basic service integration provides a unified AAI framework enabling us to track and acknowledge contributions to terminologies across the services, but also control access to terminologies with licence restrictions. Other NFDI services will depend on our service's availability and scalability. We tested Kubernetes [16] for container orchestration and will continue using it to handle high request volumes, ensuring uninterrupted service. Kubernetes enables horizontal and vertical scaling, supporting service duplication and resource allocation to prevent downtime. To monitor both usage and performance we will implement monitoring and analytics in close coordination with NFDI, Base4NFDI and other basic services to provide detailed KPIs like number of terminologies, number of integrated services, disciplines covered, availability of backend services, number of API Gateway calls.

3.2. Future Development and Ramp-up Outlook

The Ramping-Up-for-Operation phase will focus on the stability, scalability, sustainability, and expandability of the TS4NFDI service, complemented by continuous development and integration strategies to meet community needs and connect to related EOSC services such as EOSC Data Type Registry (DTR) and Metadata Schema and Crosswalk Registry (MSCR). TIB

and ZB MED are committed to long-term maintenance and hosting of their OLS-based terminology services. Terminology Registry, Services, and Mapping Applications run by VZG are part of their established, permanent infrastructure. InfAI as an active member of the OntoPortal alliance and the GFBio e.V. association is committed to the long-term maintenance and hosting of the OntoPortal based GFBio terminology service BiodivPortal [17]. We recognize the need to develop additional business models for the long-term operation of TS4NFDI as part of the overall NFDI strategy and will actively contribute to this. Along with institutional commitment, additional staff will be needed for high operational availability and prompt incident response. In the Ramping-Up-for-Operation Phase, we will consolidate and extend monitoring and analytics to provide NFDI-wide harmonised KPIs on service performance and usage. Success measures include terminology coverage and improved overall terminology quality through TS4NFDI curation workflows. We expect increasing integration of terminologies in downstream services like data annotation tools and will develop caching strategies to manage traffic. Continuous communication with user communities will promote awareness and acceptance of the basic service TS4NFDI. We will participate in consortia events and organise workshops and hackathons to integrate and use our service, collecting user feedback to inform development. Providing open educational resources and implementing a helpdesk are also priorities. To ensure integration and interoperability with future EOSC services, TS4NFDI will demonstrate connectivity between NFDI and EOSC. The OntoPortal technology, part of the TS4NFDI backend, is central to a task in the EOSC FAIR-IMPACT project. Strategically, EOSC integration is driven by Oliver Koepler's role in the EOSC Technical and Semantic Interoperability Task Force. TS4NFDI members will attend the EOSC Symposium in Berlin 2024 and the EOSCxNFDI side event to explore intersections with EOSC activities. TS4NFDI will engage in the EOSC Node discussion and define its role in an EOSC national node.

3.3. Risks and Challenges

Risk	Involved WP	Mitigation Measures
TS4NFDI does not upscale fast enough to respond to increased usage of the service. *Likelihood: medium*	WP2, WP3 WP4, WP5	We implement a monitoring system for TS4NFDI. We use a service orchestration system to ensure service availability and scalability alongside constant performance monitoring.
NFDI consortia implement their own terminology services for their communities. *Likelihood: medium*	WP1, WP5 WP6	We establish communication and reporting workflow to include the consortia and their communities in the progress of TS4NFDI, collecting their requirements and demonstrating use cases and their adaption.
The underlying TS technology of the TS4NFDI architecture shifts in a direction that jeopardises the integration into TS4NFDI. *Likelihood: medium*	WP2, WP3 WP4	We establish continuous exchange with the open-source projects of OLS, OntoPortal, or Skosmos and service providers to identify and address upcoming conflicts.
There are not enough resources available to support the NFDI consortia to integrate TS4NFDI in their services. *Likelihood: medium*	WP1, WP3, WP4, WP5, WP6	We will collaborate with the Base4NFDI service stewards to coordinate the requirements and support for the consortia.
TS4NFDI does not have enough resources to integrate future external terminology services into TS4NFDI. *Likelihood: medium*	WP4, WP6	Together with the communities we will identify potential candidates as early as possible and evaluate efforts and benefits in a transparent manner.

Table 2: Summarised the risks and defines measures to minimise the probability of occurrence

4. Support Actions from Base4NFDI / NFDI Sections, and Integrating NFDI Consortia / Efforts

Support from	Work package(s)	Contact Person Basic Service
Service Stewards (Base4NFDI)	WP6	Roman Baum
WG Terminology Services + WG Ontology Harmonisation and Mapping (Section Metadata)	WP1	Oliver Koepler, Roman Baum
WG Ontology Harmonisation and Mapping (Section Metadata)	WP3	Oliver Koepler
IAM4NFDI (basic service)	WP1, WP2, WP4, WP5	Naouel Karam, Oliver Koepler, Roman Baum, Jakob Voss

Table 3: Support needs from Base4NFDI / Service Stewards

Description of contribution	Involved effort	Consortium (contact)
Integration Mapping Service widgets in NFDI4Chem Metadata Schema Service	Participation in an Incubator Project	Steffen Neumann; NFDI4Chem
Searching over entities in the knowledge graphs registry; Reuse of mappings	Participation in an Incubator Project	Konrad Förstner; Base4NFDI/KGI4NFDI
Integration of TSS widgets to the NFDI4Health Study Hub	Participation in an Incubator Project	Johannes Darms; NFDI4Health Juliane Fluck; NFDI4Health

Table 4: Contributions required from the integrating consortia

5. Work Programme

5.1. Overview of Work Package

Work package	Deliverables (D) and milestones (M)	Responsible partner
WP 1 TS4NFDI Service Portal	**D1.1** TS4NFDI Service Portal (Month 12, iterations until Month 24) **D1.2** Authentication and authorization is implemented for access-restricted area (Month 18) **M1.1** Stable TS4NFDI Service Portal pilot providing basic functionalities (Month 12) **M1.2** Extended TS4NFDI Service Portal and application available (Month 24)	**TIB** ZB MED InfAI
WP 2 Mapping Service	**D2.1** Mapping registry and corresponding API (Month 12, iterations until Month 24) **D2.2** IAM integration to connect external applications (Month 18) **M2.1** Stable mapping service pilot with basic functionality (Month 12) **M2.2** Extended mapping service and application available (Month 24)	**VZG** TIB ZB MED InfAI
WP 3 Workflows for terminology curation	**D3.1** Implementation terminology suggestion workflow (Month 12) **D3.2** Implementation issues and term requests workflows (Month 18) **D3.3** Implementation terminology annotation workflow (Month 24) **M3.1** MVP terminology suggest feature (Month 6) **M3.2** MVP annotation workflow feature (Month 12) **M3.3** MVP term request workflows (Month 18)	**TIB** RPTU InfAI
WP 4 Extension of the API Gateway	**D4.1** API specification for additional functionalities (Month 6 first specifications ready, iterative extensions until Month 24) **D4.2** Service wrappers for mapping services (Month 18) **D4.3** Updated API gateway technical documentation (Month 24) **M4.1** Successful integration and support for additional TS technologies (Month 18) **M4.2** Advanced gateway functionalities up and running (Month 24)	**InfAI** TIB VZG ZB MED
WP 5 Extension of the Terminology Service Suite	**D5.1** Readiness to deploy Terminology Service Suite V2 (Month 9 first version, iterative extensions until Month 24) **D5.2** The Terminology Service Suite is a MediaWiki extension (Month 18 first version, iterative extensions until Month 24) **M5.1** IAM functionality is integrated to the TSS (Month 12) **M5.2** Terminology mappings are integrated to the TSS (Month 18) **M5.3** Terminology entity sets are integrated to the TSS (Month 18) **M5.4** ANTELOPE functionalities are integrated to the TSS (Month 12) **M5.5** First integration of TSS in a Wikibase instance (Month 18)	**ZB MED** TIB VZG RPTU
WP 6 Community engagement, communication, and training	**D6.1** Incubator Projects (Month 6 first cycle, iterative cycles till Month 24) **D6.2** Community workshops (Month 3 first, iterative extensions till Month 15) **D6.3** Training sessions (Month 1, iterative extensions till Month 24) **D6.4** Training material (Month 6, iterative extensions till Month 24) **M6.1** End of first Incubator Cycle (Month 6) **M6.2** First Training Material is published (Month 6)	**ZB MED** TIB VZG InfAI RPTU

Table 5: Overall work program with work packages, deliverables, milestones, and responsible partners.

The corresponding lead of a work package is marked in bold.

5.2. Detailed Work Programme

5.2.1. WP1 TS4NFDI Service Portal (Lead: TIB)

TS4NFDI aims to support different scientific communities and their various use cases and downstream services. We therefore want to enable consortia and users to put together customised and individually tailored configurations of the basic service via a graphical user interface. To facilitate the integration and customisation of the features of TS4NFDI into the services of the NFDI consortia, TS4NFDI will provide a central TS4NFDI Service Portal. The TS4NFDI Service Portal is designed to facilitate the creation and management of entity sets and terminology collections. This enables domain experts to configure the response of the API Gateway or the TSS following their specific requirements. To enable customisation, an administrator user interface will be provided within the configuration panel, utilising the widgets provided by the TSS. This will ensure simple access and enhance usability for administrators. The configuration panel will display a comprehensive list of all available terminologies from the various terminology services in the backend. Furthermore, it is possible to list terminologies that are subject to licence restrictions. Should a user meet the licence conditions, these terminologies can also be utilised via the widgets. To facilitate the decision-making process for administrators concerning the selection of terminologies, the OntoPortal FAIRness assessment tool O'FAIRe [18] will also be adapted. Additionally, the configuration panel will facilitate the transfer of novel terminologies to one of the backend systems. The tool SHACL Play! [19] is used to assess the quality of a proposed terminology. Next to this, the Mapping Service from WP2 will be also accessible via the TS4NFDI Service Portal. To implement these access-restricted areas or licenced terminologies the features of the basic service IAM4NFDI will be used. The TS4NFDI Service Portal will also integrate the TSS from WP5, and will provide documentation of the Incubator Projects from WP6 similar to the IAM incubator project dashboard [20]. The utilisation of an incubator dashboard within this system will facilitate transparency, ensure that NFDI communities remain informed and engaged, and provide a platform for showcasing success stories. Furthermore, the dashboard can be used as a key performance indicator (KPI).

5.2.2. WP2 Mapping Service (Lead: VZG)

Mappings between terminologies are crucial for interoperability within NFDI consortia. This work package aims to create and establish a mapping service, building upon the infrastructure from the DFG-funded coli-conc project, including the Cocoda application and related standards and APIs [21]. A mapping registry will be developed to store all relevant NFDI terminology

mappings, accessible via APIs and open data formats (JSKOS, SSSOM, RDF, CSV) to adhere to FAIR principles. The registry will enable identifier lookups and alignment of research data using mappings and inference rules, and will import mappings in new formats like SSSOM and from other sources like Ontology Xref Service (OxO) and Metadata Schema and Crosswalk Registry (MSCR). A web-based mapping application will facilitate the inspection, creation, and evaluation of mappings, building on Cocoda [22], already in use in NFDI4Objects. This TS4NFDI instance will include integration of terminologies from connected NFDI services via API Gateway, enhanced support for mapping terminologies, integration of mapping suggestions from other tools, support for personal terminology collections, notification mechanisms for mapping reviews and changes, and import/export functionality via web interface. Terminology mappings will be applied in diverse data curation and processing environments. While the mapping registry APIs may suffice, authenticated access is necessary for storing selections of mappings. We will implement the NFDI basic service IAM for authentication and authorization in the mapping application and API access, enabling tasks like mapping creation and review in external applications and widgets (see WP5).

5.2.3.   **WP3 Workflows for terminology curation (Lead: TIB)**

Terminologies evolve to reflect advancements in the domains they describe. This work package aims to implement core functionalities identified during the requirement analysis phase into TS4NFDI, facilitating regular curation tasks for high-quality terminologies. The service backend and widgets will enable community users, administrators, curators, and terminology maintainers to collaboratively manage terminologies. Terminology developers can inform users about usage policies and recommendations, while users can contribute by reporting issues and suggesting improvements. Interactions will be organised into four categories. **Term Request**: Users often notice missing terms during curation or data annotation. Two workflows will be implemented for filing term requests. Using the TSS, API Gateway, and user-friendly templates, a process will be integrated to ingest term requests into original terminology repositories on GitHub, utilising GitHub templates [23]. Additionally, following DataPLANT's broker ontology approach, users can directly contribute terms to a version-controlled terminology text file from tools like Swate. These terms will enter an evaluation pipeline where terminology experts review and incorporate them, linking them to suitable international terminologies. **Issue Report**: Users frequently encounter issues with existing terms and terminologies, such as mistakes in descriptions and synonyms. A facilitated issue reporting process will connect users directly with terminology maintainers, accommodating various reporting policies. Users can file issues on a curation page for each terminology, even if not hosted on public repositories. Maintainers will be notified of new issues. **Terminology Suggestion**: As new terminologies are created by different NFDIs and

communities, a channel for suggesting new terminologies for integration will be established. The TSS and API Gateway will allow users to submit a terminology configuration file, which will be reviewed by backend service curators. Users will be informed of the outcome through profile status updates or notifications. **Annotations**: Annotations will support collaborative curation, adding detailed insights, contextual information, and collaborative comments to terminologies. Terminology developers can add announcements and share curation policies. Annotation notes will be attached to terms via unique resolvable identifiers and stored in the API Gateway service databases, ensuring portability across systems.

5.2.4.   WP4 Extension of the API Gateway (Lead: InfAI)

The architecture of the centralised system (see figure 1) now includes a central API Gateway that serves as a single access point for API requests across various NFDI TSs. A service wrapper is specified for each connected TS backend technology, acting as an adapter to the underlying services. The gateway invokes the appropriate service wrappers for incoming requests and combines their responses, managing format translations and service-specific security requirements like authentication tokens or API keys. During the integration phase, the list of supported TS technologies will be extended to include DANTE [24], alongside OLS, OntoPortal, and Skosmos, and several non-standard APIs for widely used terminologies like ICONCLASS. We will coordinate with national efforts, such as aligning with the Bundesinstitut für Arzneimittel und Medizinprodukte's central terminology service for Germany [25], and consider connecting additional resources like GND, NERC, GeoNames, CENtree, or the Catalogue of Life for NFDI4Biodiversity [26]. Essential functionalities required by WP1, WP2, and WP5 will be implemented. Enhancements include extending the API Gateway with a configuration API that encompasses Single Sign-On (SSO) capabilities and comprehensive user management based on IAM4NFDI, enabling storage of user-specific configurations like collections, entity sets, and multi-language preferences. In order to enhance the reuse of collections and entity sets while simultaneously reducing the necessity for maintenance, the API Gateway will be extended, potentially through the adaptation of the HL7® FHIR® terminology service API [27] for entity sets. Service wrappers for mapping services like Cocoda and OxO will be developed to ensure seamless integration with our Terminology Service Suite. A dedicated service wrapper will facilitate access to older versions of terminologies via the API Gateway or TSS if available from backend systems. Extending the API Gateway is also necessary to support workflows for terminology curation as outlined in WP3. All required API extensions and gateway adaptations will be performed to support these functionalities, with active contributions from project partners to address additional requirements of TS4NFDI.
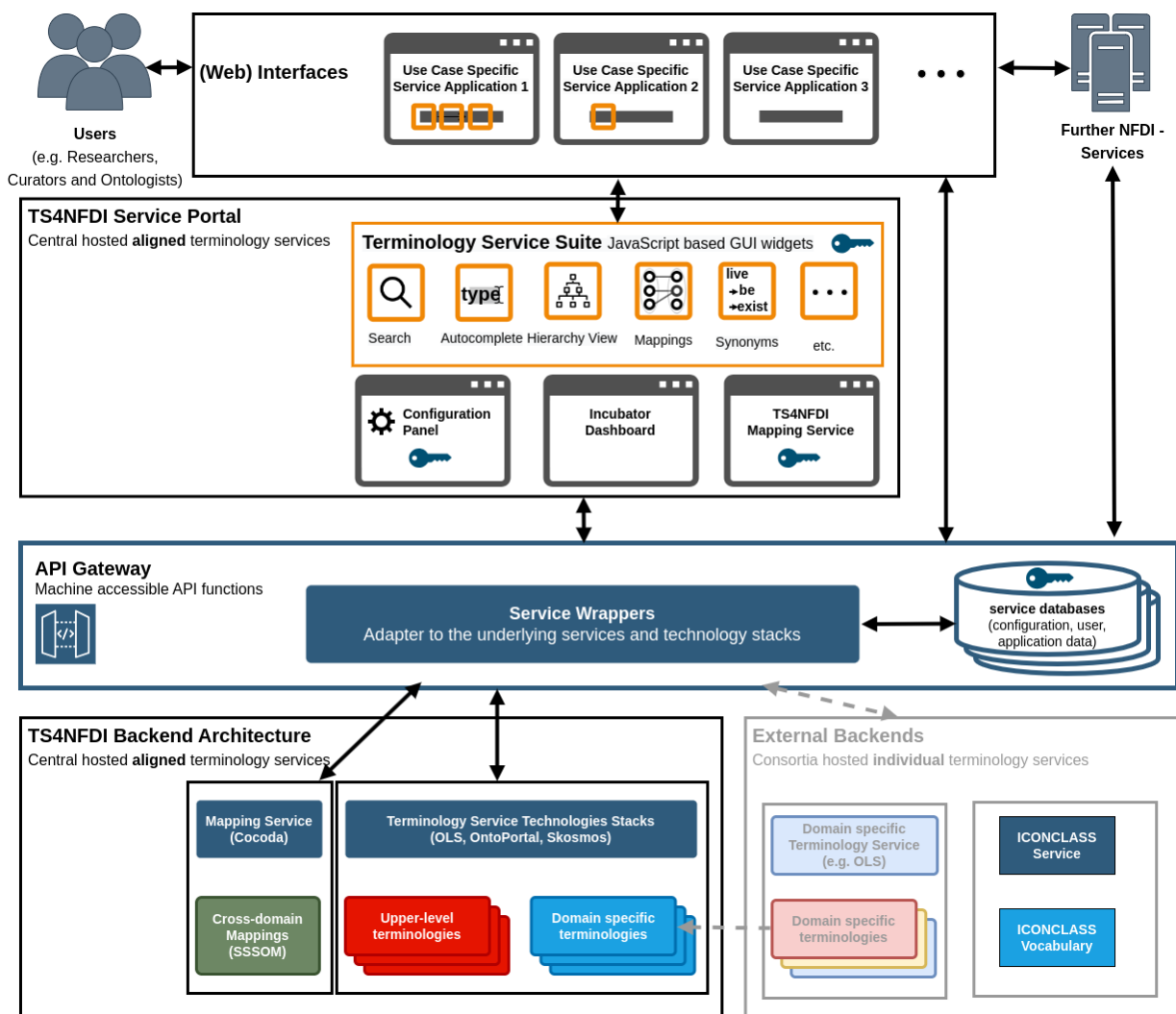
Figure 1: Overview of the TS4NFDI Architecture.

The aim is to improve system efficiency and robustness through gateway consolidation and optimisation. Advanced caching mechanisms will be implemented to enhance response times and reduce server load. The system will conform to the latest interoperability and data-sharing guidelines by adopting standards such as the EOSC MOD-API recommendation [28].

5.2.5. WP5 Extension of the Terminology Service Suite (Lead: ZB MED)

The first version of the TSS was developed during the initialisation phase of TS4NFDI to facilitate integration via widgets into other services. The TSS will be extended with additional functionality, new components, and improved accessibility and usability. Most terminologies can be shared in line with FAIR principles so TSS did not require authentication so far. This will be changed to also support services with write-access such as terminology curation and mapping creation and to also access terminologies with licence restrictions (WP1). TSS will be extended with authentication features based on the NFDI basic service IAM. The TSS widgets will allow for SSO to obtain tokens that can be used to prove the identity and membership of users at the

service backend. Authenticated access will also allow for acknowledgment and incentivization of contributions to terminology extension, thus improving collaboration and contributions. Functionalities such as viewing, creating, and annotating mappings based on the mapping service developed in WP2 will be made available via widgets. This will allow easy access to terminology mapping data and integration of mapping functionalities into existing consortia services. This task will be carried out in close collaboration with WP2 to leverage established infrastructure and ensure seamless integration and functionality. Furthermore, the functionality of the TSS will be enhanced by the provision of widgets for entity sets. As a preliminary step, the Swate templates [29] could be adopted. The Swate templates have graphical user interfaces that already handle entity sets. These entity sets could be used to generate complex (spreadsheet-based) annotation schemas. These entity-set-based widgets will facilitate community efforts by enabling custom templates to more easily adhere to domain-specific minimal information standards. Moreover, these widgets will be integrated into the TS4NFDI Service Portal from WP1. The functionality of the ANTELOPE [14] service is currently available for use via API and a standalone web frontend. It is intended that ANTELOPE will be integrated into the TSS. The planned integration will extend the existing functionality of TSS concerning semantic annotations, such as entity linking based on textual and visual content. Furthermore, the TSS functionality should be provided as a separate MediaWiki extension to Wikibase. This will facilitate the seamless utilisation of terminology services directly within Wikibase. As a proof of concept, the Wikibase instance of the TIB Wikibase4Research [30] will be used as an incubator.

5.2.6.   **WP6 Community engagement, communication, and training (Lead: ZB MED)**

WP 6 will foster community interaction and integration with the TS4NFDI service, supporting user-driven development of new and existing features. Following the IAM4NFDI incubator approach, incubator projects will be conducted to integrate TS4NFDI into consortia services by exploring and testing new features with the TS4NFDI team. These projects will aim to i) establish a Minimum Viable Product (MVP), ii) align service needs with consortia resources, and iii) develop new ideas terminology applications. User-driven development cycles will include requirements analysis, service integration, testing and user feedback. The first incubator project, identified during the initialization phase, could involve NFDI4Health and the NFDI4Health German Central Health Study Hub [31]. This project would address specific medical terminology needs, such as using FHIR interfaces, and could integrate TSS widgets to develop further requirements, including entity sets or collections via the API Gateway. Training sessions on terminologies, mappings, and services will be provided for participants of varying experience levels, enhancing the training's efficacy. Training materials, including documents and recorded

sessions, will be published on the TS4NFDI Service Portal (WP1). Community workshops will be organised for outreach, and the project team will participate in consortia meetings and NFDI section working group meetings to foster community contact. To increase reusability and understanding of TS4NFDI services, the source code will be published with an open-source licence on GitHub, including demonstration projects and integration how-to descriptions.

III Appendix

a) Bibliography and list of references

[1] S. Jupp, T. Liener, S. Sarntivijai, O. Vrousgou, T. Burdett, and H. Parkinson, 'OxO - A Gravy of Ontology Mapping Extracts', presented at the International Conference on Biomedical Ontology, 2017. Accessed: Jul. 31, 2024. [Online]. Available: https://www.semanticscholar.org/paper/OxO-A-Gravy-of-Ontology-Mapping-Extracts-Jupp-Liener/98ff d29ab7b3d454c97dea76692bd4f6f0610623

[2] EMBL-EBI, 'OxO Ontology Mapping Service'. Accessed: Jul. 31, 2024. [Online]. Available: https://www.ebi.ac.uk/spot/oxo/

[3] N. Matentzoglu *et al.*, 'A Simple Standard for Sharing Ontological Mappings (SSSOM)', *Database*, vol. 2022, p. baac035, Jan. 2022, doi: 10.1093/database/baac035.

[4] U. Balakrishnan, 'DFG-Projekt: Coli-conc. Das Mapping Tool "Cocoda"', *O-Bib Offene Bibl.*, vol. 3, Mar. 2016, doi: 10.5282/o-bib/2016H1S11-16.

[5] **R. Baum and O. Koepler, 'Leveraging Terminology Services for FAIR Semantic Data Integration Across NFDI Domains: How to Integrate Terminology Services Into Other Service Applications', *Proc. Conf. Res. Data Infrastruct.*, vol. 1, Sep. 2023, doi: 10.52825/cordi.v1i.356.**

[6] **'1st Base4NFDI User Conference (UC4B2024)', Events (Indico). Accessed: Jul. 31, 2024. [Online]. Available: https://events.gwdg.de/event/658/contributions/**

[7] **'TS4NFDI Community Hands-On Workshop', NFDI4Chem. Accessed: Jul. 31, 2024. [Online]. Available: https://www.nfdi4chem.de/event/ts4nfdi-community-hands-on-workshop/**

[8] **T. Bender, 'Erfolgreicher Workshop über FAIRe Terminologiedienste', NFDI4Chem. Accessed: Jul. 31, 2024. [Online]. Available: https://www.nfdi4chem.de/de/erfolgreicher-workshop-uber-faire-terminologiedienste/**

[9] **'TS4NFDI Online Community Workshop Feb 2024', Google Docs. Accessed: Jul. 31, 2024. [Online]. Available: https://docs.google.com/document/d/17mcburo58K9gyP3UlE_2A82l571PrjSKn0w61qb9a_4/ed it**

[10] S. Jupp, T. Burdett, C. Leroy, and H. Parkinson, 'A new Ontology Lookup Service at EMBL-EBI', presented at the Workshop on Semantic Web Applications and Tools for Life Sciences, 2015. Accessed: Jan. 11, 2023. [Online]. Available: https://www.semanticscholar.org/paper/A-new-Ontology-Lookup-Service-at-EMBL-EBI-Jupp-Burdett/ b83bfbfc1f2f08e5b88af5ef65ef2a8687ac4112

[11] P. L. Whetzel *et al.*, 'BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications', *Nucleic Acids Res.*, vol. 39, no. Web Server issue, pp. W541–W545, Jul. 2011, doi: 10.1093/nar/gkr469.

[12] I. Subirats-Coll *et al.*, 'AGROVOC: The linked data concept hub for food and agriculture', *Comput. Electron. Agric.*, vol. 196, p. 105965, May 2022, doi: 10.1016/j.compag.2020.105965.

[13] B. Zapilko, J. Schaible, P. Mayr, and B. Mathiak, 'TheSoz: A SKOS representation of the thesaurus for the social sciences', *Semantic Web*, vol. 4, no. 3, pp. 257–263, Jan. 2013, doi: 10.3233/SW-2012-0081.

[14] **'ANTELOPE | NFDI4Culture'. Accessed: Jul. 31, 2024. [Online]. Available: https://service.tib.eu/annotation/**

[15] **DataPLANT Team, 'Swate'. Accessed: Jul. 31, 2024. [Online]. Available: https://github.com/nfdi4plants/Swate**

[16] 'Production-Grade Container Orchestration'. Accessed: Jul. 31, 2024. [Online]. Available: https://kubernetes.io/

[17] **'Welcome to the NFDI4Biodiversity BiodivPortal | NFDI4Biodiversity BiodivPortal'. Accessed: Jul. 31, 2024. [Online]. Available: https://biodivportal.gfbio.org/**

[18] E. Amdouni, S. Bouazzouni, and C. Jonquet, 'O'FAIRe: Ontology FAIRness Evaluator in the AgroPortal Semantic Resource Repository', in *The Semantic Web: ESWC 2022 Satellite Events*, P. Groth, A. Rula, J. Schneider, I. Tiddi, E. Simperl, P. Alexopoulos, R. Hoekstra, M. Alam, A. Dimou,

and M. Tamper, Eds., Cham: Springer International Publishing, 2022, pp. 89–94. doi: 10.1007/978-3-031-11609-4_17.

[19] 'SHACL Play!' Accessed: Jul. 31, 2024. [Online]. Available: https://shacl-play.sparna.fr/play/

[20] 'Incubator Dashboard'. Accessed: Jul. 31, 2024. [Online]. Available: https://incubators.nfdi-aai.de/

[21] **'coli-conc - Welcome to coli-conc'. Accessed: Jul. 31, 2024. [Online]. Available: https://coli-conc.gbv.de/**

[22] **J. Voß, 'Normdaten-Mapping mit Cocoda', 2020, *Technische Informationsbibliothek (TIB),Berufsverband Information Bibliothek e. V. (BIB)*. doi: 10.5446/36465.**

[23] **'VibrationalSpectroscopyOntology/.github/ISSUE_TEMPLATE/new-term-request-issue-template.md at main · NFDI4Chem/VibrationalSpectroscopyOntology', GitHub. Accessed: Jul. 31, 2024. [Online]. Available: https://github.com/NFDI4Chem/VibrationalSpectroscopyOntology/blob/main/.github/ISSUE_TEMPLATE/new-term-request-issue-template.md**

[24] **'Dokumentation der Schnittstellen zu DANTE'. Accessed: May 15, 2023. [Online]. Available: https://api.dante.gbv.de/**

[25] 'BfArM - Terminologyserver'. Accessed: Jul. 31, 2024. [Online]. Available: https://www.bfarm.de/EN/Code-systems/Services/Terminologyserver/_node.html

[26] 'The Catalogue of Life', COL - The Catalogue of Life. Accessed: Jul. 31, 2024. [Online]. Available: https://www.catalogueoflife.org/

[27] 'ValueSet - FHIR v5.0.0'. Accessed: Jul. 31, 2024. [Online]. Available: https://hl7.org/fhir/valueset.html

[28] 'MOD API'. Accessed: Jul. 31, 2024. [Online]. Available: https://fair-impact.github.io/MOD-API/

[29] ***nfdi4plants/Swate-templates*. (Jul. 17, 2024). F#. DataPLANT. Accessed: Jul. 31, 2024. [Online]. Available: https://github.com/nfdi4plants/Swate-templates**

[30] 'Service - Wikibase4Research - NFDI4Culture'. Accessed: Jul. 31, 2024. [Online]. Available: https://nfdi4culture.de/de/dienste/details/wikibase4research.html

[31] NFDI4Health, 'German Central Health Study Hub'. Accessed: Feb. 16, 2024. [Online]. Available: https://csh.nfdi4health.de/

b) Gantt chart

Deliverables/Milestones	2025												2026											
	Jan.	Feb.	Mar.	Apr.	Mai	June	July	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	Mai	June	July	Aug.	Sep.	Oct.	Nov.	Dec.
TS4NFDI Service Portal																		D1.1						
Authentication and authorization is implemented for access-restricted area																		D1.2						
Stable TS4NFDI Service Portal pilot providing basic functionalities												M1.1												
Extended TS4NFDI Service Portal and application available																								M1.2
Mapping registry and corresponding API																		D2.1						
IAM integration to connect external applications																		D2.2						
Stable mapping service pilot with basic functionality												M2.2												
Extended mapping service and application available																								M2.2
Implementation terminology suggestion workflow												D3.1												
Implementation issues and term requests workflows																		D3.2						
Implementation terminology annotation workflow																								D3.3
MVP terminology suggest feature						M3.1																		
MVP annotation workflow feature												M3.2												
MVP term request workflows																		M3.3						
API specification for additional functionalities														D4.1										
Service wrappers for mapping services																		D4.2						
Updated API gateway technical documentation																								D4.3
Successful integration and support for additional TS technologies																		M4.1						
Advanced gateway functionalities up and running																								M4.2
Readiness to deploy Terminology Service Suite V2																D5.1								
The Terminology Service Suite is a MediaWiki extension																			D5.2					
IAM functionality is integrated to the TSS												M5.1												
Terminology mappings are integrated to the TSS																		M5.2						
Terminology entity sets are integrated to the TSS																		M5.3						
ANTELOPE functionalities are integrated to the TSS												M5.4												
First integration of TSS in a Wikibase instance																		M5.5						
Incubator Projects														D6.1										
Community Workshops									D6.2															
Training sessions												D6.3												
Training material															D6.4									
End of first Incubator Cycle						M6.1																		
First Training Material is published						M6.2																		