DFG form nfdi10 - 03/24 page 1 of 14

Letter of Intent for Renewal Proposals in 2024



DFG form nfdi10 - 03/24 page 2 of 14

1 Binding letter of intent as advance notification of a full renewal proposal

Binding letter of intent (required as advance notification for renewal proposals in 2024)

2 Formal details

Name of the consortium

NFDI4Chem - Fachkonsortium Chemie in der NFDI

Acronym of the consortium

NFDI4Chem

Applicant institution

Friedrich-Schiller-Universität (FSU), Fürstengraben 1, 07743 Jena

Interim director: Prof. Dr. Georg Pohnert

Spokesperson

Prof. Dr. Christoph Steinbeck, <u>christoph.steinbeck@uni-jena.de</u>, Friedrich-Schiller-Universität

Co-applicant institution TIB

Technische Informationsbibliothek (TIB), Welfengarten 1B, 30167 Hannover

Director: Prof. Dr. Sören Auer

Co-spokesperson:

Dr. Oliver Koepler, oliver.koepler@tib.eu, Technische Informationsbibliothek

Co-applicant institution RWTH

RWTH Aachen Universität (RWTH), Templergraben 55, 52062 Aachen

Rector: Prof. Dr. rer. nat. Dr. h.c. mult. Ulrich Rüdiger

Co-spokesperson:

Prof. Dr. Sonja Herres-Pawlis, <u>sonja.herres-pawlis@ac.rwth-aachen.de</u>, RWTH Aachen Universität

Co-applicant institution KIT

Karlsruher Institut für Technologie (KIT), Kaiserstraße 12, 76131 Karlsruhe

Acting president: Prof. Dr. Oliver Kraft



DFG form nfdi10 - 03/24 page 3 of 14

Co-spokesperson:

Dr. Nicole Jung, nicole.jung@kit.edu, Karlsruher Institut für Technologie

Co-applicant institution JGU

Johannes Gutenberg-Universität Mainz (JGU), Saarstraße 21, 55122 Mainz President: Prof. Dr. Georg Krausch

Co-spokesperson:

Dr. Johannes Liermann, <u>liermann@uni-mainz.de</u>, Johannes Gutenberg-Universität Mainz

Co-applicant institution FIZ

FIZ Karlsruhe - Leibniz Institute for Information Infrastructure (FIZ), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen

Director and managing director: Prof. Dr. Wolfram Horstmann

Co-spokespersons:

Dr. Felix Bach, felix.bach@fiz-karlsruhe.de, FIZ Karlsruhe - Leibniz Institute for Information Infrastructure

Co-applicant institution IPB

Leibniz Institute of Plant Biochemistry (IPB), Weinberg 3, 06120 Halle Managing director: Prof. Dr. Alain Tissier

Co-spokesperson:

PD Dr. habil. Steffen Neumann, <u>sneumann@ipb-halle.de</u>, Leibniz Institute of Plant Biochemistry

 Participant institution (new participants in **bold green font**, legally dependent organisational entities of co-applicant institutions in *italic font*)

Bundesanstalt für Materialforschung und -prüfung (BAM) Unter den Eichen 87, 12205 Berlin

Beilstein-Institut (BI)

Trakehner Str. 7-9, 60487 Frankfurt am Main

Center for Transformation in Chemistry (CTC)
Am Mühlenberg 1, 14476 Potsdam-Golm



DFG form nfdi10 - 03/24 page 4 of 14

Deutsche Pharmazeutische Gesellschaft e.V. (DPhG)

Varrentrappstr. 40-42, 60486 Frankfurt am Main

FIZ Karlsruhe, Bereich Immaterialgüterrechte (FIZ-IGR)

Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen

Freie Universität Berlin (FUBe)

Kaiserswerther Str. 16/18, 14195 Berlin

Georg-August-Universität Göttingen (GAU)

Wilhelmsplatz 1, 37073 Göttingen

Gesellschaft Deutscher Chemiker e.V. (GDCh)

Varrentrappstr. 40-42, 60486 Frankfurt am Main

Leibniz-Institut für Photonische Technologien e.V. (IPHT)

Albert-Einstein-Str. 9, 07745 Jena

Johannes Gutenberg-Universität Mainz, Experimental Biophysical Chemistry (JGU-EBC)

Duesbergweg 10 - 14, 55128 Mainz

Julius-Maximilians-Universität Würzburg (JMU)

Sanderring 2, 97070 Würzburg

Karlsruher Institut für Technologie, Institute für Nanotechnologie (KIT-INT)

Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen

Karlsruher Institut für Technologie, Institut für Technische Chemie und

Polymerchemie (KIT-ITCP)

Engesserstr. 18 / 20, 76131 Karlsruhe

Karlsruher Institut für Technologie, Institut für Theoretische Informatik (KIT-ITI)

Kaiserstraße 12, 76131 Karlsruhe

Karlsruher Institut für Technologie, Scientific Computing Center (KIT-SCC)

Kaiserstraße 12, 76131 Karlsruhe

Ludwig-Maximilians-Universität München (LMU)

Geschwister-Scholl-Platz 1, 80539 München



DFG form nfdi10 - 03/24 page 5 of 14

Physikalisch- Technische Bundesanstalt (PTB)

Bundesallee 100, D-38116 Braunschweig

Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau (RPTU) Gottlieb-Daimler-Strasse, Building 47, D-67663 Kaiserslautern

RWTH Aachen Universität, Institut für Physikalische Chemie (RWTH-IPC) Landoltweg 2, 52074 Aachen

RWTH Aachen Universität, IT Center (RWTH-ITC) Seffenter Weg 23, 52074 Aachen

RWTH Aachen Universität, Universitätsbibliothek (RWTH-UB) Templergraben 61 RWTH Navigator, 52062 Aachen

Technische Universität Braunschweig (TUBr) Universitätsplatz 2, 38106 Braunschweig

Technische Universität Dortmund (TUDo)

August-Schmidt-Straße 4, 44227 Dortmund

Technische Universität Dresden (TUDr) 01062 Dresden

Helmholtz-Zentrum für Umweltforschung GmbH (UFZ) Permoserstraße 15, 04318 Leipzig

Universität Leipzig (ULe) Ritterstraße 26, 04109 Leipzig

Universität Münster (UM) Schlossplatz 2, 48149 Münster

Universität Stuttgart (USt) Keplerstraße 7, 70174 Stuttgart

Universität Ulm (UUlm) 89069 Ulm

- Participant individual (new participants in bold green font)
 - Dr. Ricardo Cuhna, Institut für Umwelt & Energie, Technik & Analytik e.V. (IUTA)

 Dr. Stefan Kuhn, Tartu University Estland (UTartu)



DFG form nfdi10 - 03/24 page 6 of 14

3 Objectives, work programme and research environment in the second funding period

 Research area of the proposed consortium (according to the DFG classification system: www.dfg.de/dfg profil/gremien/fachkollegien/faecher/index.jsp)

| - | 2.11-01 | Biochemie |
|---|---------|---|
| - | 2.22-08 | Pharmazie |
| - | 3.11 | Molekülchemie |
| - | 3.13 | Physikalische Chemie |
| - | 3.14 | Analytische Chemie |
| - | 3.15 | Biologische Chemie und Lebensmittelchemie |
| - | 3.16 | Polymerforschung |
| _ | 3.17 | Theoretische Chemie |

Concise summary of the consortium's main objectives and task areas

The National Research Data Infrastructure for Chemistry in Germany (NFDI4Chem) consortium aims to complete and consolidate a comprehensive research data management (RDM) infrastructure for the chemical sciences. This initiative is ensuring that chemical data are systematically managed, shared, and utilised efficiently, adhering to the FAIR (Findable, Accessible, Interoperable, and Reusable) principles. The consortium, involving numerous partners from universities, research institutions, information infrastructure institutions, and industry, focuses on creating an integrated and sustainable infrastructure that supports all stages of the chemical research data lifecycle. Addressing our vision that *All chemists publish FAIR data* we have formulated the following Key Objectives:

Key Objective 1: Evolve the federation of services and repositories for the storage, disclosure, search and reuse of research data. Integration of community standards for the seamless integration of distributed data sources and uniform access to data for innovative data re-use.

Key Objective 2: Lead international community processes to establish minimum information (MI) standards for data and machine-readable metadata to create semantically rich linked chemistry data.

Key Objective 3: Develop and foster the application of Smart Laboratory Environments by promoting the use of digital tools in all stages of research and create seamless digital data workflows.

Key Objective 4: Foster cultural and digital change together with the chemistry community in Germany to create awareness for, and foster a community-agreed, legally and technically reliable RDM infrastructure for all levels of academia, beginning in undergraduate studies curricula.



DFG form nfdi10 - 03/24 page 7 of 14

Key Objective 5: Collaborate with other consortia and promote cross-cutting developments to extend synergies within OneNFDI and enable interdisciplinary data harmonisation and integration.

Key Objective 6: Facilitate data re-use and enable AI in chemistry by providing AI-ready data and data infrastructures.

The consortium's work is divided into six primary task areas (TAs), each focusing on specific aspects of the infrastructure development and implementation.

TA1 Management and Coordination oversees the administrative, financial, and technical management of the consortium. It coordinates consortium activities, manages funds, ensures compliance with legal and ethical standards, and facilitates communication among partners.

TA2 Smart Laboratory develops and implements digital tools and environments for efficient data capture and management in laboratories necessary to capture data early in the life cycle. It creates and enhances electronic lab notebooks (ELNs), integrates laboratory instruments, and develops workflows that facilitate seamless data transfer and interoperability within the infrastructure.

TA3 Repositories develops and evolves a federated system of repositories for the storage, sharing, and preservation of chemical data. This includes raw data in diverse formats as well as curated datasets. It develops and maintains core repositories like RADAR4Chem, Chemotion, and nmrXiv, ensuring they meet the needs of various sub-disciplines in chemistry. It facilitates data deposition, retrieval through standardised protocols, and re-using research data across distributed data sources.

TA4 Metadata, Data Standards, and Publication Standards defines and implements standards for metadata and data formats to ensure consistency and interoperability, together with reference implementations and data validation. Ontologies are used where possible, and missing terminological artefacts are added. It develops minimum information standards, harmonises metadata practices, and collaborates with international organisations like IUPAC to adopt and foster global standards.

TA5 Community Involvement and Training fosters a culture of effective data management practices within the chemical research community. It conducts training workshops, develops educational materials, and engages with researchers to promote the adoption of RDM practices and offers incentives for innovations. It is present at chemical conferences with booths and talks and maintains multiple active communication channels to disseminate information and collect the community's requirements, needs and feedback.

TA6 Synergies and Cross-Cutting Topics ensures the integration and interoperability of various components of the NFDI4Chem infrastructure, including the search service. It promotes the use of ontologies and semantic technologies to enhance data interoperability, Al-readiness and



DFG form nfdi10 - 03/24 page 8 of 14

machine-actionability. It coordinates with other NFDI consortia, sections and working groups on cross-domain metadata standards, cross-domain mapping of ontologies for semantic data annotation, provision of terminology services and addresses legal and ethical issues related to RDM. Collaborate with international bodies such as the Research Data Alliance (RDA) or the International Union of Pure and Applied Chemistry (IUPAC).

The NFDI4Chem consortium is making significant progress in building a national research data management infrastructure for chemistry in Germany. These efforts are already advancing chemical research and ensuring that data are managed efficiently, shared openly, and utilized to their full potential.

 Brief description of the proposed use of existing infrastructures, tools and services that are essential in order to fulfil the planned consortium's objectives

To achieve the planned objectives, NFDI4Chem will use mainly those tools and services that were built, established or extended during the first funding period. They are hosted and maintained by selected institutions and their existing infrastructures of which continuous support for e.g. hosting is required. In the future, basic services like IAM4NFDI, TS4NFDI, PID4NFDI or KG4NFDI need to be integrated or connected to the NFDI4Chem service federation in order to pave the way for and strengthen our long-term service strategy.

Tools: To enable and promote digitalization, NFDI4Chem develops an **ELN platform** supporting chemistry specific functions. Important for the integration into this ELN but also in the form of stand-alone software are other tools such as NMRium, ChemSpectra, ChemScanner, ChemConverter and Ketcher, to be supported by different means in the next funding period.

ELN-specific services: NFDI4Chem will continue to offer services based on the chemistry ELN (docker container service for IT admins/staff and Chemotion ELN as SaaS for individual researchers and small groups) and other ELNs of interest in certain subdisciplines of chemistry (hosting of relevant ELNs as test instances by KIT).

Repository services and databases: The federation of core repositories comprises five repositories and two databases hosted in Germany, each covering essential content in key subdisciplines of the chemical community. They were developed and/or provided as individual services that will be continued as community driven core infrastructure component of NFDI4chem: RADAR4Chem is a cross-domain repository (hosted by FIZ), the Chemotion Repository is provided for chemical reactions, chemical substances, and characterization data (hosted by KIT), the nmrXiv for all types of NMR spectroscopy data (hosted at FSU), VibSpecDB for vibrational spectroscopy data (hosted at FSU), and MassBank EU for mass spectrometry data (hosted at the UFZ). The database Suprabank offers data on intermolecular and supramolecular interactions (hosted by KIT) and STRENDA DB for enzymatic data (hosted by Beilstein Institute).



DFG form nfdi10 - 03/24 page 9 of 14

Overarching services: As a central hub for searching the federated repositories of NFDI4Chem, the <u>Search Service</u> (hosted by TIB) will be continued to include all datasets from all NFDI4Chem hosted repositories and further related ones. The NFDI4Chem <u>Helpdesk</u> (hosted by TIB) serves and will serve as a central hub for supporting the community's requests for all NFDI4Chem services and RDM topics. The <u>Knowledge Base</u> (hosted by JGU) already covers a wide range of topics, from basic RDM concepts to more in-depth articles about the use of NFDI4Chem's services, tools and infrastructures and will be extended with the growing community of NFDI4Chem. The <u>Terminology Service</u> (hosted by TIB) provides access to ontologies, terminologies and vocabularies in chemistry and related disciplines enabling their reuse and curation.

 Interfaces to other NFDI consortia: brief description of existing agreements for collaboration and/or plans for future collaboration

NFDI4Chem members have been active in shaping the NFDI and its initiatives build upon the cross-cutting topics from the very beginning. Many members of NFDI4Chem contribute to the sections (including two section leads) and their working groups in NFDI association. This broadbased commitment is complemented by additional individual collaborations with individual consortia on specific topics.

NFDI4Chem is part of Physical Sciences in NFDI collaboration. In the Joint Colloquium leading scientists showcase good data practices in physics, chemistry, mathematics and informatics to an international, interdisciplinary audience. Further joint activities include the discussion of Persistent Identifiers in Physical Sciences (Schörner et al. 2023) and ELNs in Experimental Sciences. NFDI4Chem and NFD4Ing have signed a memorandum of understanding to cooperate on the adoption of the RDMO tool to promote data management plans for their communities. With NFDI4Cat we have organised bilateral NFDI4C* workshops on synergy and cooperation, which have resulted into a close cooperation concerning ontology development. In our annual Ontologies4Chem workshop we discuss and coordinate ontology development for chemistry and the neighbouring sciences with NFDI4Cat, NFDI4Health, DataPlant and NFDI4Ing. NFDI4Chem coordinates with several consortia in the discussion of legal aspects of RDM in the section ELSA. As the latest example we contributed to the FAIRagro Workshop "Legal Status Quo" on RDM and licensing challenges. NFDI4Chem, NFDI4Cat, DataPlant, NFDI4Ing and NFDI4Culture closely collaborate on the development of domain-specified Terminology Services to support ontology management tasks and the integration into further services. The NFDI4Chem Terminology Service is part of TS4NFDI project, where we cooperate with NFDI4Health and NFDI4Biodiversity to provide Basic Terminology Service for the NFDI. NFDI4Chem closely collaborates with different



DFG form nfdi10 - 03/24 page 10 of 14

consortia - especially with NFDI4Cat and FAIRMat - with respect to the development and adaptation of ELNs.

A comprehensive overview of collaborations of NFDI consortia and past joint events, publications and more can be found in (Creators Amelung, Lisa Bodenschatz, E...). NFDI4Chem is involved in more than ¼ of all cross-consortium collaborations within its domain and beyond listed there. We seek to continue all ongoing collaborations and extend our collaborative network within the NFDI.

4 International and national networking

Establishing community standards in chemistry, including both data and metadata, cannot be accomplished solely on a national scale but needs **international collaborative efforts**. NFDI4Chem leadership and participants have successfully initiated or joined such efforts and are engaging in a wide range of activities. Together, we have already achieved significant results with regards to minimum information standards, data formats, metadata and ontologies. We are committed to continuing this strategy, and strengthen our cooperation with the International Union of Pure and Applied Chemistry (IUPAC) on digital standards development, the IUPAC Gold Book as a canonical source for textual definitions in chemical ontologies. and collaboration in the WorldFAIR chemistry project.

Sonja Herres-Pawlis is Chair of the Organometallics/Inorganics Working Group of the InChI Trust and elected member of the InChI Trust, Oliver Koepler is a newly appointed member of the EOSC Technical and Semantic Interoperability Task Force. Nicole Jung is member of a consortium of ELN vendors to define common specifications that allow the interchange of data and metadata between different ELNs.

We collaborate with the Physical Sciences Data Infrastructure PSDI in the UK conducting joint workshops to interlink the individual working groups to join forces and work on standards for data, metadata or ELN. Members of NFDI4Chem are active in the ELIXIR community on topics like terminologies, terminologies services or metadata standards such as Bioschemas.org and its application for chemistry. Several key members of NFDI4Chem are engaged in the Chemistry Research Data Interest Group (CRDIG) and four more working groups of the Research Data Alliance RDA. Since 2020 we have organised and contributed to several sessions of CRDIG and other working groups during RDA Plenaries. As part of our ontology development and harmonisation efforts we organise the regular Int. Ontologies4Chem Workshop with participants from all major OBO ontology projects, NFDI4Cat, NFDI4Health, DataPlant, NFDI4Ing. The development of the VIBSO ontology is done in collaboration with the EU Charisma project, the National Physical Laboratory (NPL), and the NEST Laboratory, Scuola Normale Superiore. For Deutsche Forschungsgemeinschaft

Kennedyallee 40 · 53175 Bonn, Germany · Postal address: 53170 Bonn, Germany Tel.: + 49 228 885-1 · Fax: + 49 228 885-2777 · postmaster@dfg.de · www.dfg.de

DFG form nfdi10 - 03/24 page 11 of 14

future ontology development we plan to collaborate and network with Pistoia Alliance and Allotrope Foundation as part of our industry engagement strategy.

On a **national level** we closely operate together with the learned societies Deutsche Pharmazeutische Gesellschaft (DPhG), Bunsen Gesellschaft, Gesellschaft Deutscher Chemiker (GDCh). The FAIR4Chem Award is sponsored by the Fonds der Chemischen Industrie (FCI) and awarded at the annual meeting of the JungChemikerForum (JCF). Cross-discipline we collaborate within state-level or local RDM initiatives, and IPB is part of the de.NBI network.