NFDI4Chem
Fachkonsortium Chemie für die Nationale Forschungsdateninfrastruktur

*Sprecher: Prof. Dr. Christoph Steinbeck, christoph.steinbeck@uni-jena.de,
Institut für Anorganische und Analytische Chemie, Lessingstr. 8, 07743 Jena*

*Dr. Oliver Koepler, oliver.koepler@tib.eu, Technische Informationsbibliothek,
Welfengarten 1B, 30167 Hannover*

Consortium, Objectives, Data Types

The vision of NFDI4Chem is the digitalisation of all key steps in chemical research to support scientists in their efforts to collect, store, process, analyse, disclose and re-use research data. NFDI4Chem intends to represent all disciplines of chemistry in academia. In the initial phase, NFDI4Chem focuses on data related to molecules and reactions including data for their experimental and theoretical characterisation.

The consortium and in particular the members of the steering committee are experienced scientists working in the field of cheminformatics, chemistry-focussed research data management (RDM), information infrastructure and services. They represent know-how for domain-specific scientific aspects, digitalisation strategies and methods, e-infrastructures and digital libraries, data preservation and long term archiving of scientific data as well as standards and metadata. The members established open chemistry databases and repositories as well as cheminformatics infrastructure components as Open Source software in the past. They are part of several communities such as ELIXIR, GO FAIR, RDA, IUPAC working groups. Chemistry is one of the core natural sciences with interdisciplinary connections to many other fields such as biology, physics, medicine, or material sciences. Chemical research dealing with molecules, their reactions and properties can be described by data resulting fro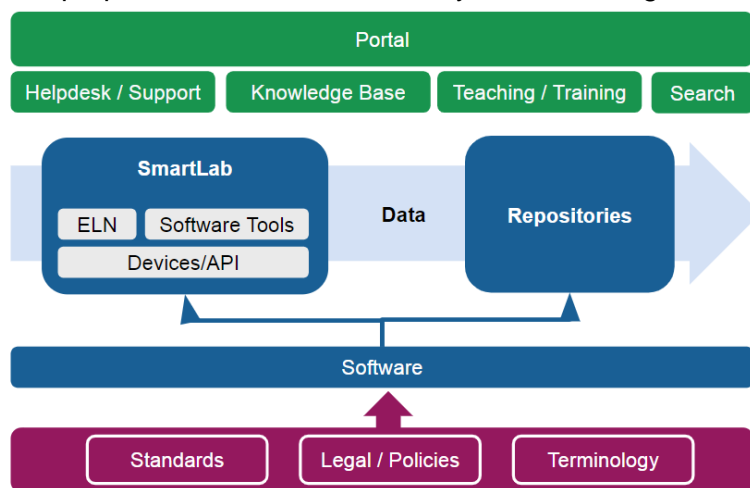m experimental procedures, observations, and theoretical models. The large diversity of experimental and theoretical methods results in a plethora of different data and metadata types and formats, most of them being proprietary. NFDI4Chem will support chemists in all steps of the research data management data life cycle. The workflow starts capturing data at the lab workbench, from spectroscopic / analytical methods, and from simulations. The data is then organised via virtual research environments and can be collected, shared and disclosed by repositories and curated databases.

Fig. 1: Proposed digital data lifecycle from SmartLab to repositories

At all levels of this workflow, we selected existing infrastructures and services to be integrated within NFDI4Chem.

Missing key elements for chemistry RDM life cycle support were identified and their development will complement the envisaged infrastructure. Parallel to this methodological progress, NFDI4Chem focuses on intense community interaction. We will promote increased and sustainable awareness of RDM in the minds of researchers towards a cultural change, provide

them with tutorials and workshops for RDM tools and mediate the communication between the community (e.g. collecting the requirements of all sub-domains) and RDM facilities in both ways.

Services

Our services will foster the digitisation of key steps in research workflows starting from experiment planning, to data acquisition and management to data publication and thus will promote a digital change towards FAIR data handling. The foundation of our services will be **standards**, **policies** and **terminologies** in close connection with NFDI cross-cutting topics. On that foundation, we will offer two service pillars of a) infrastructure and b) information and support services. Those measures are complemented by teaching and training activities. Documentation and support will be available via helpdesk and a NFDI4Chem Knowledge Base. Databases and repositories that cover the relevant data types used by the NFDI4Chem community will be included into the envisioned federation of repositories. Those **Repository Services** are complemented by a Search Service and a Terminology Service. The **Search Service** facilitates the access to all the resources available in the NFDI4Chem core and associated repositories. The NFDI4Chem portal realises the concept of a single point of entry to NFDI4Chem services and further information. The Search Service, the knowledge base or the helpdesk will be provided via the portal. The **Terminology Service** will provide machine-readable and human-readable descriptions of research data, thus enabling researchers and components of NFDI4Chem and NFDI to access, curate and update vocabularies for chemistry and related domains. The **Smart Lab** fosters data workflows from devices towards repositories in the federation. It integrates software components that are necessary requirements to build the NFDI4Chem services as well as an Electronic Laboratory Notebook (ELN) which supports scientists in collecting, managing, storing, analysing, and sharing data as a preparatory step to disclose data via a repository. A cultural change in data handling and management in chemistry, as envisaged by NFDI4Chem, is necessary to implement the FAIR data principles. NFDI4Chem will support the scientists in this process with a service for Teaching and Training and a Helpdesk to foster acceptance and use of the community. The **teaching and training** team is of high importance for the general acceptance of the infrastructure and to enable the scientists to use the components of the infrastructure in the right manner. NFDI4Chem will teach scientists about how to use the provided infrastructure for digitalised work according to the FAIR data principles and will raise awareness for its importance. The team will plan roadshows and will be available on request, offering not only theoretical advice but live demo sessions with the centrally hosted ELN and its functions.

Cross-Cutting topics

The vision of the NFDI is the cross-linking of infrastructure components and services, as well as the creation of new, comprehensive services for the creation of an interdisciplinary national research data infrastructure. Joint strategies for cross-cutting topics have been discussed between consortia over the last months. Following the Berlin Declaration[1], the current status of these discussions is reflected in the Leipzig-Berlin Declaration[2]. NFDI4Chem actively contributes to the discussions and is co-author of both publications. Identified topics cover building of a cooperative framework, incentives for cultural change, development of a common data and services infrastructure and legal aspects of research data management. In addition to the definition of mutual, key issues in terms of content, the organizational shape of the consortia'

---

[1] Glöckner F. O., Diepenbroek M., Felden J., Overmann J., Bonn A., Gemeinholzer B., et al. Berlin Declaration on NFDI Cross-Cutting Topics. 2019. doi:10.5281/zenodo.3457213

[2] Bierwirth, M., Glöckner, F. O., Grimm, C., Schimmler, S., Boehm, F., Busse, C., Degkwitz, A., Koepler, O. & Neuroth, H. Leipzig-Berlin-Erklärung zu NFDI-Querschnittsthemen der Infrastrukturentwicklung. 2020. doi:10.5281/zenodo.3895209

cooperation (like the envisaged inter-consortia working groups, the role of the NFDI directorate and the embedding in the NFDI structure) must be clarified. NFDI4Chem has allocated resources to actively engage with other consortia within the NFDI (and beyond) on particular cross-cutting topics. NFDI4Chem will especially contribute to the development of a NFDI-wide Authentication and Authorisation Infrastructure (AAI), the implementation of a legally reliable framework of policies and guidelines for FAIR research data management, and interdisciplinary standards on vocabularies. Ontologies will be used as an integral aspect of the standards where possible, and missing terminological artifacts will be created, and available through the terminology services.
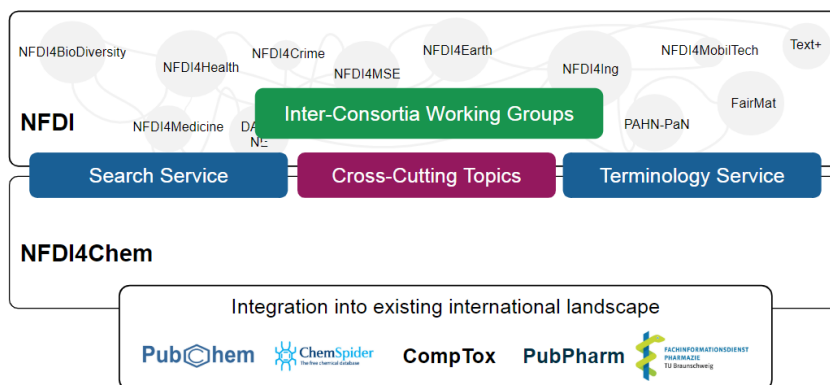

Fig. 2: Embedding of NFDI4Chem in the NFDI and community landscape

The NFDI4Chem Terminology Service will enable semantic data interoperability, discovery and exploitation across disciplines and thus supports user-centric scientific applications.

NFDI4Chem and the NFDI as a whole will not operate in isolation, hence an international embedding will be very crucial from the onset. Links to other European and International initiatives will be important to avoid reinventing the wheel and benefit from synergies. Key personnel of NFDI4Chem are active in a number of international efforts, such as GO FAIR, RDA interest groups, ELIXIR implementation networks, the European Open Science Cloud (EOSC) and more. We will help to leverage these connections for the benefit of the NFDI as a whole.

Conference expectations
Expecting the launch of our consortium in October 2020, we hope to intensify networking and coordination with other starting consortia and the NFDI directorate about NFDI governance and cross-cutting topics. We will lay out our planned first steps in establishing the chemistry consortium and the constitution of our advisory boards of chemistry stakeholders.

Vorgesehene Mitglieder des Konsortiums (Co-Sprecherinnen/Co-Sprecher und die weiteren, beteiligten Institutionen):

Co-Sprecher/in	Zugehörige Institution
Prof. Dr. Christoph Steinbeck *Professor for Analytical Chemistry, Cheminformatics and Chemometrics* *christoph.steinbeck@uni-jena.de*	Friedrich-Schiller-University, Jena Fürstengraben 1, 07743 Jena
Dr. Felix Bach *Group Lead Research Data Management* *felix.bach@kit.edu*	Karlsruhe Institute of Technology (KIT) Steinbuch Centre for Computing (SCC) Kaiserstraße 12, 76131 Karlsruhe
Prof. Dr. Sonja Herres-Pawlis *Chair of Bioinorganic Chemistry* *sonja.herres-pawlis@ac.rwth-aachen.de*	RWTH Aachen University Institut für Anorganische Chemie, Landoltweg 1, 52074 Aachen
Dr. Nicole Jung *Group Leader Compound Platform* *nicole.jung@kit.edu*	Karlsruhe Institute of Technology (KIT) Institute of Organic Chemistry Kaiserstraße 12, 76131 Karlsruhe
Dr. Oliver Koepler *Head of Lab Linked Scientific Knowledge* *oliver.koepler@tib.eu*	Technische Informationsbibliothek, Welfengarten 1B, 30167 Hannover
Dr. Johannes Liermann *Head of NMR spectroscopy service unit* *liermann@uni-mainz.de*	Johannes Gutenberg University Mainz, Department of Chemistry, 55099 Mainz
Dr. Steffen Neumann *Group leader Bioinformatics and Scientific Data* *sneumann@ipb-halle.de*	Leibniz Institute of Plant Biochemistry Weinberg 3, 06120 Halle
Matthias Razum *Head e-Research* *matthias.razum@fiz-karlsruhe.de*	FIZ Karlsruhe - Leibniz Institute for Information Infrastructure Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen

Participants

Ansprechpartner/in	Zugehörige Institution
Dr. Martin G. Hicks, Dr. Carsten Kettner	Beilstein-Institut, Trakehner Str. 7-9, 60487 Frankfurt am Main
Prof. Dr. Stefan Laufer, Prof. Dr. Andreas Link	Deutsche Pharmazeutische Gesellschaft e.V., Varrentrappstr. 40-42, 60486 Frankfurt am Main
PD Dr. Carsten Baldauf, Prof. Dr.	Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin
Prof. Dr. Franziska Boehm	FIZ Karlsruhe - Leibniz Institute for Information Infrastructure, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen
PD Dr. Thomas Bocklitz	Friedrich Schiller Universität Jena, Institute of Physical Chemistry and Abbe Center of Photonics (IPC), Helmholtzweg 4, 07743 Jena, Germany
Prof. Dr. Ricardo Mata	Georg-August-Universität Göttingen, Institut für Physikalische Chemie Tammannstr. 6, 37077 Göttingen

Ansprechpartner/in	Zugehörige Institution
Prof. Wolfram Koch, Dr. Hans-Georg Weinig	Gesellschaft Deutscher Chemiker e. V., Varrentrappstr. 40-42, 60486 Frankfurt am Main
Prof. Ludger Wessjohann, Dr. Silke Pienkny, Dr. Frank Broda	Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle
Prof. Dr. Alfred Forchel, Prof. Dr. Ulrich Schatzschneider	Fakultät für Chemie und Pharmazie, Julius-Maximilians-Universität Würzburg, Sanderring 2, 97070 Würzburg
Prof. Dr. Wolfgang Wenzel, Dr. Frank Biedermann	Karlsruher Institut für Technologie, Institut für Nanotechnologie Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen
Prof. Dr. Patrick Théato	Karlsruher Institut für Technologie, Institute for Chemical Technology and Polymer Chemistry Engesserstrasse 18, 76128 Karlsruhe
Dr. Thomas Engel	Ludwig-Maximilians-Universität München, Department Chemie, Butenandtstr. 5-13, 81377 München
Dr. Giacomo Lanza	Physikalisch-Technische Bundesanstalt, Bundesallee 100, D-38116 Braunschweig
Prof. Dr. Matthias Müller	IT Center, RWTH Aachen University, Templergraben 55, 52062 Aachen
Katrin Stump, Stefan Wulle	Universitätsbibliothek der TU Braunschweig, Fachinformationsdienst Pharmazie, Universitätsplatz 1, 38106 Braunschweig
Prof. Dr. Stefan M. Kast, Prof. Dr. Paul Czodrowski	Technische Universität Dortmund, Fakultät für Chemie und Chemische Biologie Otto-Hahn-Str. 4a, 44227 Dortmund
Prof. Dr. Wolfgang E. Nagel, Dr. Ralph Müller-Pfefferkorn, Dr. Richard Grunzke	Technische Universität Dresden (TUD), Center for Information Services and High Performance Computing (ZIH), 01062 Dresden
PD Dr. Werner Brack, Dr. Tobias Schulze	Helmholtz-Zentrum für Umweltforschung GmbH - UFZ, Permoserstraße 15, 04318 Leipzig
Dr. Nils Schlörer	Universität zu Köln, Department für Chemie, Greinstr. 4, 50939 Köln