Progress Report

National Research Data Infrastructure for and with Computer Science (NFDI×CS)

NFDIxCS Progress Report Part1

National Research Data Infrastructure for and with Computer Science

1.   General Information

Name of the consortium

NFDIxCS National Research Data Infrastructure for and with Computer Science

Research domains or research methods addressed by the consortium

The research domain addressed by NFDIxCS is the scientific discipline Computer Science (in German Informatik). Thus the range of research methods and related research data types is quite broad. A good overview and categorisation can be found in the DFG classification (DFG Fachkollegium[1] *4.43 Informatik*) which defines a certain breakdown of the field which is, however, changing - even growing - over time due to needs of basic research funding management. A similar but slightly different view is the breakdown of the discipline in terms of special interest groups (in German Fachbereiche) of the Gesellschaft für Informatik (GI) which can be found here[2]. This better reflects the structure of the community addressed by NFDIxCS. Classic empirical research to semi structural research (e.g. software engineering or theory of computing) to natural language / video audio are topics of the scientific processes NFDIxCS supports.

URL of the consortium website and repositories used for publishing output

https://nfdixcs.org/

2.   Summary

The ambitious goal of theNFDIxCS consortium is to build a common research data management infrastructure for the Computer Science (CS) community. The starting point for this activity was around 2022 and can be characterized as different ways and different maturity levels in handling research data and software across the range of CS subdisciplines. In an abstract way, CS research data is available across the board, but the systematic management especially with regards to the FAIR principles and for the envisaged archival period of 10 years was hardly to be observed. In addition, the publication culture required the publication of the respective research data only in a few CS areas. Thus, for the consortium this entailed a quite basic startup position for a broad range of activities. Although this was a huge task, the related activities were already well prepared due to the efforts during the proposal preparation and the subsequent activities.

The start of the project in March 2023 saw a well prepared team of key persons and a process to hire new staff as well as building the necessary structural elements for communication and legal aspects of the collaboration. Moreover, initial connection to sections, taskforces, other

---

[1] https://www.dfg.de/de/ueber-uns/gremien/fachkollegien/fachsystematik/ingenieurwissenschaften-4-43
[2] https://gi.de/netzwerk/fachbereiche Fachbereiche (SIGs) der Gesellschaft für Informatik

committees and BASE4NFDI within the general setting of the NFDI association is in full flow already. As usual this ramp up phase took some time.

The technical work addressed the following key areas: (1) to build an abstraction for data and software artifacts in the form of the *research data management container (RDMC)* for integrating the various types of CS research data along with the infrastructure to handle it, including all necessary (meta) information, execution context and documentation. The other key area (2) was to reach out to the *community of CS researchers* to introduce potential users to the new challenges in RDM as well as to define processes and standards, quality criteria and other community-related aspects. These included also the aforementioned publication culture for conferences and journals. In addition the area of outreach was addressed to include the international partners and the related communities abroad including EOSC.

RDMC Research Data Management Container

The necessary abstractions to build repositories for the variety of types of research data were analysed from the perspectives of CS sub-disciplines. The initial ideas and additional progress were discussed with leading roles in the computer science community in Germany and at international events. The key idea is to create a RDMC specification which implements the requirements for FAIR Digital Objects[3]. The conceptualization phase hammering out the details of these requirements included a series of prototypes leading to minimal viable products which are published at the consortium's website.They are now available for testing by the community. This phase included a series of open and internal hackathons which helped a lot to either discuss the properties, use cases and details on one hand and included interested parties outside NFDIxCS as well.

The RDMC serves as a timecapsule for research data, the related software and further context to allow the reuse of the containerized research data years later. The challenges to long term archival of research software are addressed by the so-called *Reusable Execution Environment* REE which includes physically or virtually[4] all necessary software components. The containerization is the basis to cryptographically store the research data and protect it against unwanted access for privacy and/or intellectual property reasons. The initial set of use cases to deal with the tasks to create, store, assess, search for and reuse a RDMCs will be provided in our portal ndfixcs.org and are implemented in the form of a number of services run by the consortium of NFDIxCS partners. This is also the starting point for community engagement and related processes and workflow.

---

[3] https://www.go-fair.org/today/fair-digital-framework/
[4] This is provided by the Software Heritage Foundation Archive which stores all available open source software over time c.f. https://www.softwareheritage.org/

Community Building Including International Outreach

The community processes are organized in cooperation with the Gesellschaft für Informatik (GI[5]). This is facilitated by a clear structure of so-called Fachbereiche (special interest groups , SIGs) covering the entire area of computer science including a special cross cutting SIG on modelling in computer science. In addition a series of web based and in person events to reach out to the community at large are being organized to communicate and discuss all aspects of RDM within computer science. In addition, the outreach to European partners and beyond is organized via EOSC and two associations Informatics Europe[6] and IFIP[7].

Outlook

Currently the development of the genuine NFDIxCS services is in full swing and the first versions are available for testing(!) at nfdixcs.org/services. As planned we start with a selection of sub-disciplines which are software engineering, human computer interaction, knowledge engineering and modelling. This will be quickly extended while the processes of the community building and renovating the related publication processes will pick speed in the coming months.

---

[5] https://gi.de/
[6] https://www.informatics-europe.org/
[7] https://www.ifip.org/

3.   Composition of the consortium

Applicant institution

Applicant institution	Location	Duration
University of Duisburg-Essen	Essen and Duisburg	01.03.2023 - today

Spokesperson

Spokesperson	Institution, location	Duration
Prof Michael Goedicke	Fak Informatik / paluno	01.03.2023 - today

Co-applicant institutions

Co-applicant institutions	Location	Duration
Gesellschaft für Informatik e.V. (GI)	Berlin/Bonn	01.03.2023 - today
Universität Potsdam (UP)	Potsdam	01.03.2023 - today
Karlsruher Institut für Technologie (KIT)	Karlsruhe	01.03.2023 - today
Universität Hamburg (UHH)	Hamburg	01.03.2023 - today
Ludwig-Maximilians Universität München (LMU)	München	01.03.2023 - today
Technische Universität München (TUM)	München	01.03.2023 - today
Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)	Göttingen	01.03.2023 - today
Johannes Gutenberg Universität Mainz (JGUM)	Mainz	01.03.2023 - today
Technische Universität Dresden (TUD)	Dresden	01.03.2023 - today
Forschungszentrum Jülich (FZJ)	Jülich	01.03.2023 - today
Uni Kiel	Kiel	Until project start
Universität Bayreuth (UBT)	Bayreuth	01.03.2023 - today
Universität Paderborn (UPB)	Paderborn	01.03.2023 - today
Rheinisch-Westfälische Technische Hochschule Aachen (RWTH)	Aachen	01.03.2023 - today
Technische Universität Darmstadt (TUDA)	Darmstadt	01.03.2023 - today
Schloss Dagstuhl – Leibniz Center for Informatics (LZI)	Wadern	01.03.2023 - today

| FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur | Karlsruhe | | 01.03.2023 - today |

Co-spokespersons

Co-spokespersons	Institution, location	Task area(s)	Duration
Daniel Krupka	Gesellschaft für Informatik e.V. (GI), Berlin	C3, C1, M3, M1	01.03.2023 - today
Prof Dr. Ulrike Lucke (Deputy Spokesperson)	Uni Potsdam	C1, A1, M2	01.03.2023 - today
Prof. Dr. Anne Koziolek	KIT Karlsruhe	M2, D3, A2	01.03.2023 - today
Prof. Dr. Ralf Reussner	KIT Karlsruhe	D1, D3, C2	01.03.2023 - today
Prof. Dr. Hannes Federrath	Uni Hamburg	I4, C4	01.03.2023 - today
Prof. Dr. Albrecht Schmidt	LMU München	C3, D2	01.03.2023 - today
Prof. Dr. Tobias Nipkow	TU München	I1	01.03.2023 - 31.03.2025
Prof. Dr. Martin Schulz	TU München	D2, I1, C3	01.03.2023 - today
Prof. Dr. Ramin Yahyapour	GWDG Göttingen	I2, A2, D1, D2, D4, I1, A3, M3	01.03.2023 - today
Prof. Dr.-Ing. André Brinkmann	JGU Mainz	I1, D2, D4, I3	01.03.2023 - today
Prof. Dr. Wolfgang E. Nagel	TU Dresden	I3, D4	01.03.2023 - today
Prof. Dr. Dr. Thomas Lippert	FZ Jülich	D4, I3	01.03.2023 - today
Prof. Dr. Agnes Koschmider	Uni Kiel Uni Bayreuth	D3 D2	Until project start 01.03.2023 - today
Prof. Dr. Christian Plessl	Uni Paderborn	D4, I3	01.03.2023 - today
Prof. Dr. Matthias Müller	RWTH Aachen	D4, I3	01.03.2023 - today

Prof. Dr. Christian Bischof	TU Darmstadt	D4	01.03.2023 - 31.3.2023
Prof. Dr. Felix Wolf	TU Darmstadt	D4	01.04.2023 - today
Prof. Raimund Seidel, Ph.D.	Schloss Dagstuhl – Leibniz Center for Informatics	A3, C2	01.03.2023 - 30.04.2025
Prof. Holger Hermanns	Schloss Dagstuhl – Leibniz Center for Informatics	A3, C2	01.05.2025 - today
Dr. Marcel R. Ackermann	Schloss Dagstuhl – Leibniz Center for Informatics	A3, D3, C2	01.03.2023 - today
Dr. Michael Wagner	Schloss Dagstuhl – Leibniz Center for Informatics	C2, A3	01.03.2023 - today
Prof. Dr. Franziska Boehm	FIZ Karlsruhe	C4, I4	01.03.2023 - today

Participants

Participating institutions	Location	Duration
Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI)	Kaiserslautern	01.03.2023 - today
Verein zur Förderung eines Deutschen Forschungsnetzes (DFN-Verein)	Berlin	01.03.2023 - today
Deutsches Klimarechenzentrum (DKRZ)	Hamburg	01.03.2023 - today
Universität Stuttgart, Höchstleistungsrechenzentrum Stuttgart (HLRS)	Stuttgart	01.03.2023 - today
Bayerische Akademie der Wissenschaften, Leibniz Supercomputing Centre (LRZ)	Garching near Munich	01.03.2023 - today

Participating individuals	Institution, location	Duration
Prof. Dr. Florian Alt	Universität der Bundeswehr München, Forschungsinstitut CODE	01.03.2023 - today 01.03.2023 - today

Prof. Dr. Christoph Benzmüller	Freie Universität Berlin, Dep. of Math; Confederation of Laboratories for Artificial Intelligence Research in Europe (CLAIRE)	01.03.2023 - today
Dr. Daniel Demmler	Universität Hamburg, Security in Distributed Systems	left before 01.03.2023
Prof. Dr. Gregor Engels	Universität Paderborn, Institut für Informatik, Datenbank-und Informationssysteme	01.03.2023 - today
Dr. Bernhard Fechner	Fernuniversität Hagen, Zentrum für Medien und IT	01.03.2023 - today
Prof. Dr. Martin Frank	Steinbuch Centre for Computing Karlsruher Institut für Technologie	01.03.2023 - today
Dr. Christian Grimm	DFN Verein zur Förderung eines Deutschen Forschungsnetzes e.V. Berlin	01.03.2023 - today
Prof. Dr. Wilhelm Hasselbring	Christian-Albrecht-Universität zu Kiel, AG Software Engineering	01.03.2023 - today
Prof. Dr. Jens Heidrich	Fraunhofer IESE, Division Smart Digital Solutions, Kaiserslautern Hochschule Mainz, Fachbereich Technik	01.03.2023 - 30.09.2023 01.10.2023 - today
Prof. Dr. Timo Kehrer	Universität Bern, Institute für Informatik	01.03.2023 - today
Prof. Dr. Ulf Leser	Humboldt Universität zu Berlin, Institut für Informatik	01.03.2023 - today
Dr. Jan Linxweiler	TU Braunschweig, Universitätsbibliothek	01.03.2023 - today
apl. Prof. Dr. Thomas Mandl	Universität Hildesheim, Institut für Informationswissenschaft & Sprachtechnologie	01.03.2023 - today
Prof. Dr. Vera Meister	TH Brandenburg, Fachbereich Wirtschaft	01.03.2023 - today
Prof. Dr. Andreas Oberweis	KIT, RG Business Information Systems, Karlsruhe	01.03.2023 - today
Prof. Dr. Klaus Pohl	Universität Duisburg-Essen, Software Systems Engineering, Paluno	01.03.2023 - today
Prof. Dr. Tilmann Rabl	Hasso Plattner Institute, Data Engineering Systems, Potsdam	01.03.2023 - today

Prof. Dr. Enrico Rukzio	Universität Ulm, Media Informatics	01.03.2023 - today
Prof. Dr. Bernhard Rumpe	RWTH Aachen, Software Engineering, Software Engineering	01.03.2023 - today
Prof. Dr. Philipp Schaer	Technische Hochschule Köln, Institut für Informationswissenschaft	01.03.2023 - today
Dr.-Ing. Gunther Schiefer	KIT, Betriebliche Informationssysteme, Institut für Angewandte Informatik und Formale Beschreibungsverfahren, Karlsruhe	01.03.2023 - today
Prof. Dr. Bernt Schiele	Max Planck Institute for Informatics	01.03.2023 - today
Prof. Dr.-Ing. Horst Schirmeier	TU Dresden, Institut für Systemarchitektur, Betriebssysteme	01.03.2023 - today
Prof. Dr. Ute Schmid	Universität Bamberg, Information Systems and Applied Computer Sciences / Confederation of Laboratories for Artificial Intelligence Research in Europe (CLAIRE)	01.03.2023 - today
Prof. Dr. Nicole Schweikardt	Humboldt Universität zu Berlin, Institut für Informatik	01.03.2023 - today
Prof. Dr. Peter Sobe	HTW Dresden, Faculty of Informatics/Mathematics, Computer Science Fundamentals and Programming	01.03.2023 - today
Prof. Dr. Alexander Steen	Universität Greifswald, Institut für Mathematik und Informatik	01.03.2023 - today
Prof. Dr. Sven Strickroth	LMU München, Institut für Informatik	01.03.2023 - today
Dr. Michael Striewe	Uni Duisburg-Essen, paluno	01.03.2023 - 30.09.2024
	Hochschule Trier, Fakultät Informatik	01.10.2024 - today
Prof. Dr. Matthias Weidlich	Humboldt Universität zu Berlin, Institut für Informatik	01.03.2023 - today