Letter of Intent

METHODS – Methodology-oriented consortium for generic research data infrastructures for multiple disciplines
1 Binding letter of intent as advance notification or non-binding letter of intent

☐ Binding letter of intent (required as advance notification for proposals in 2020)
☒ Non-binding letter of intent (anticipated submission in 2021)

2 Formal details

Planned name of the consortium
Methodology-oriented consortium for generic research data infrastructures for multiple disciplines

Acronym of the planned consortium
METHODS

Applicant institution
GWDG, Am Faßberg 11, 37077 Göttingen
Geschäftsführer: Ramin Yahyapour, and
University of Erfurt, Nordhäuser Str. 63, 99089 Erfurt
Präsident: Walter Bauer-Wabnegg
serve jointly as interim applicant institutions. The applicant institution will be defined by the METHODS general assembly meeting in 2021.

Spokesperson
The METHODS general assembly will make the designation of the spokesperson in 2021¹. The joint e-mail address of the preliminary spokesperson-board is spokesperson@methods4nfdi.org
The temporary spokesperson and contact person for the NFDI (functioning as steering committee) will jointly be:
Claudia Biniossek, claudia.biniossek@methods4nfdi.org, University of Erfurt; University of Cologne: OVGU Magdeburg; GESIS – Leibniz Institute for the Social Sciences
Thomas Lauer, thomas.lauer@uni-erfurt.de, University of Erfurt
Philipp Wieder, philipp.wieder@gwdg.de, GWDG, Göttingen
Peter Wittenburg, peter.wittenburg@mpcdf.mpg.de, GWDG, Göttingen; MPCDF, Garching

Co-applicant institution
University of Cologne, Albertus Magnus Platz, 50923 Köln (Rektor: Axel Freimuth)
Max Planck Computing and Data Facility (MPCDF), Gießenbachstraße 2, 85748 Garching
(Acting director: Erwin Laure)
University of Konstanz, Universitätsstraße 10, 78457 Konstanz (Rektorin: Kerstin Krieglstein)
University of Leipzig Ritterstraße 26, 04109 Leipzig (Rektorin: Beate A. Schücking)

¹ According to the DFG FAQ page 5: „Wesentlich ist, dass die gewählte Organisationsform die Ziele des Konsortiums bestmöglich unterstützt und sie von allen mitantragstellenden Einrichtungen mitgetragen wird.”
Co-spokesperson
Roger Berger, University of Leipzig
Dirk Betz, University of Cologne; OVGU Magdeburg; GESIS
Urs Fischbacher, University of Konstanz
Martin Peuker, Charité – Universitätsmedizin Berlin
Raphael Ritz, Max Planck Computing and Data Facility (MPCDF), Garching
Susumu Shikano, University of Konstanz

Participant
Martin Jaekel, Zurich University of Applied Sciences, Switzerland
Bernhard Kittel, University of Vienna, Austria
Elisabeth Lex, Graz University of Technology, Austria
Brian Nosek, Center for Open Science, Charlottesville, VA, USA
Limor Peer, Institution for Social and Policy Studies, Yale University, New Haven, CT, USA
Jon Peirce, University of Nottingham, GB
Nicole Pfeiffer, Center for Open Science, Charlottesville, VA, USA
Owen Powell, Vienna University of Economics and Business, Austria
Elisabet Rutström, Center for Economic Analysis of Risk, Georgia State University, USA
Martin Schonger, Lucerne School of Business, Switzerland
Bülent Tezkan, Institute of Geophysics and Meteorology, University of Cologne
Martin Zünkeler, KAIROS, Bochum

3 Objectives, work programme and research environment

Research area of the proposed consortium (according to the DFG classification system)
104-04 Psycholinguistik (subject area)
106-05 Religionswissenschaft und Judaistik
107 Theologie
108-03 Praktische Philosophie
109 Erziehungswissenschaft und Bildungsforschung
110 Psychologie
111 Sozialwissenschaften
111-03 Publizistik und Kommunikationswissenschaft
111-04 Politikwissenschaft
112 Wirtschaftswissenschaften
113 Rechtswissenschaft
2 Sportwissenschaft (subject area) im Wissenschaftsbereich Geistes und Sozialwissenschaften in Verbindung mit dem Bundesinstitut für Sportwissenschaft (BISp)
205 Medizin
205-01 Medizinische Informatik (subject area)
205-02 Public Health, medizinische Versorgungsforschung, Sozialmedizin
206-04 Kognitive, systemische und Verhaltensneurobiologie
206-08 Kognitive und systemische Humanneurowissenschaften
206-09 Biologische Psychiatrie
206-10 Klinische Psychiatrie, Psychotherapie und Kinder- und Jugendpsychiatrie
Concise summary of the planned consortium’s main objectives and task areas

**Method oriented approach:** The main objective of METHODS is to converge and connect existing empirical research communities which share substantial methodologies. Each of these meaningful communities is characterized by significant similarity in their workflows despite the many differences in data types, standards, and tools. However, they also share the same trends towards increasing volume, heterogeneity of types and formats, and a massive increase in relationships between Digital Objects (DOs) encapsulating discipline knowledge. Most communities have developed a mixture of discipline-specific and method-specific tools and have also started to build repositories with idiosyncratic structures. This ongoing separation by discipline limits the potential of shared infrastructures. The necessary integration into commons-based workflows and best practices constitutes THE big challenge since traditions and discipline-specific characteristics and flavors need to be preserved as widely as possible.

METHODS is a growing, method-oriented open alliance that is determined to work jointly on a common and modular workflow-framework that

1. covers the entire data lifecycle in many disciplines
2. guides the researcher through seamless, semi-automatic research data management,
3. helps to put the FAIR principles into practice with focus on interoperability and reusability and not just findability and accessibility,
4. is thus based on a FAIR Digital Object (FDO) ecosystem as invented by RDA and FAIR core members,
5. helps to overcome the many inefficiencies in data-driven science that amount to a huge waste in time. METHODS will therefore be based on a common workflow framework that automates processes where possible.
6. It will be aligned with the legal and ethical regulations that are relevant for the engaged disciplines in a customized fashion.

An in-depth analysis of more than 50 research infrastructure plans that was carried out recently\(^2\) revealed that while, on the one hand, there are many commonalities with respect to data creation, management and processing across many disciplines from all scientific domains, on the other the researchers are sticking to the data-generation and -management procedures and methods that they are used to – practices that do not, in general, lead to FAIR data. In addition, the analysis uncovered that in general, the data collecting, managing, and processing practices in the Research Data System (RDS)\(^3\) are not yet adapted to fully reap the benefits of digitalisation. At the same time, publishers are pushing the communities to adopt standards and procedures. It is increasingly obvious that the local research environment, where most of the data is created, managed, and consumed, will only change when commonalities are addressed that lead to more efficiency and enable new types of research. The analysis confirmed largely

\(^2\) The analysis has not yet been published.
\(^3\) We are using the generic term Research Data System (RDS) to point to all institutions, departments, or laboratories that generate data with the help of observations, experiments, analyses and simulations, that manage data or that are processing and analyzing data.
what was discussed in the RDA Data Fabric Group and resulted among others in the notion of Digital Objects.

**Domains and scope of METHODS:** METHODS is strictly committed to respecting and supporting the specificities of the participating disciplines and communities. The core tasks are (1) to develop "the modular workflow-framework" as it has already been introduced and described in detail above based on commonalities with regard to facilitate cross-silo/discipline data-driven research, (2) to facilitate data management by reducing, as much as possible, the growing administrative and management load introduced by the increasing number of regulations, documentation profiles (DMP) and the general data trends (volume, complexity) and (3) to offer state-of-the-art and FAIR-compliant repositories for those who lack support. The FAIR DO approach ensures at the data management level that all digital objects are treated equally. Docking existing tools and repository systems via adapters to the FDO-System reduces the complexity from a N*N task, in which every system must be individually connected to every other one, to an 1*N task, where each system is connected to a common layer exactly once. Once generic frameworks are available, extended with discipline-based adapters, meaningful modular metadata sets and surfaces, researchers will not be encumbered in their scientific work, but their work will be promoted to new frontiers, and higher efficiencies will give them time to focus again on substantive questions and not waste ¾ of their time wrangling data as is currently the case⁴.

**Main objectives and task areas until autumn 2021 (development and community building phase):** METHODS sees itself as a facilitator in letting heterogeneous groups of researchers and research communities participate in the emerging FAIR-based data space and thus make them part of the NFDI evolution. METHODS is based around an integrative strategy for the next decades and thus believes in a duty to add a methodology-oriented dimension to NFDI. It, therefore, fits into the long-term intentions of NFDI. The first stage will be community building and the dialogue within the NFDI based on the existing networks and partners. From February 2021 to May 2021, METHODS will enter its second stage to shape an inter-consortium alliance and submit a joint grant proposal in autumn 2021. A preliminary steering board has been set up, which will incrementally be replaced by a governance structure that optimally supports the aforementioned goals. In parallel, METHODS will participate in ongoing deep discussions in RDA (Data Fabric, FAIR Maturity), GEDE (global discussions about FDO details), and GOFAIR Implementation Networks (Driving FAIR implementation). In addition, it will start a new RDA Working group about a common generic workflow approach.

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

METHODS will demonstrate its flexibility by pushing FAIR compliance in its core, by also adapting to well-known open information systems such as OSF and existing infrastructures, software, and tools covering the entire research data lifecycle: Recruiting and conducting tools

(e.g., PsychoPy, z-Tree unleashed, J-Tree, oTree, lab.js, classEx, Lioness), workflows (e.g., YARD, CentraXX), and archiving solutions (e.g., Pavlovia, x-science).

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

METHODS is highly open to any form of organizational or content-related cooperation within the NFDI, other consortia, and national and international stakeholders. The idea is to reconcile the interests of all initiatives in terms of autonomy and financial resources with the interests of the NFDI community as a whole in terms of synergies, generic solutions, and actual cooperation. For this concern, we do welcome the coordinating aspects of the NFDI directorate to direct future developments. We have spoken to several already existing and planned consortia and have identified great mutual interest in collaboration. This will be worked out in the next phase.

### 4 Cross-cutting topics

**Please identify cross-cutting topics that are relevant for your consortium and that need to be designed and developed by several or all NFDI consortia.**

**Initial situation:** After proposing our method-oriented approach in the first round, we learned that our proposed overarching and integrative approach with FAIR-DOs and the adaptation of existing repositories, metadata frameworks, experimental tools, and specialized infrastructures such as OSF were evaluated as excellent both by the referees as well as the NFDI experts committee. They agreed that this approach has relevance for NFDI as a whole and should be implemented on the broadest possible basis of participating communities to cover Germany's research data entirely without unreasonable exclusions. As indicated above, we see two complementary paths to meet these requirements: (1) Forming an infrastructure with a thematic domain focus and expanding it from such a nucleus or (2) forming an infrastructure based on common methodologies being neutral to discipline cultures. Both have their advantages for NFDI, and both approaches are necessary in our view as described above.

A large fraction of research data in the humanities, medicine, life science, environmental and social sciences are not addressed by any existing or proposed NFDI consortium. This is especially true for observational research data that exist in the local and often inaccessible Research Data System and which is far from being published in the "classical sense" but which will need to be shared to implement Open Science and to make Germany a first-rate location for advanced data-driven research (see Figure 1). History of funding has shown that it would be wrong to expect that exclusively discipline-specific approaches will have a massive impact on other disciplines. Through its methodological focus, METHODS is able to address research data, researchers, and actors in the German research landscape which are not covered by any other existing or planned consortia. METHODS explicitly addresses data at all stages of the life cycle, and its methodology will support the lifecycle optimally to enable a FAIR compliant data landscape. The support will start from planning an observation, survey, simulation, or

---

5 Experts agree that the "classical publishing" procedure will not be the primary option for sharing and reusing data and software. Much shorter cycles are required to be at the cutting edge.
experiment based on theoretical considerations and end with the registration of FDOs in trustworthy repositories, including their metadata and Persistent Identifiers (PIDs). These can include sensitive data for which special protection and tracking mechanisms need to be implemented, which is facilitated effectively by the systematic application of the FDO approach.

Figure 1: Types of DOs within the data life-cycle

METHODS will implement an open methodology-oriented approach which is flexible in different ways:
(1) It will submit an independent grant proposal in 2021, addressing a variety of relevant communities from the onset.
(2) It will be open and reserve funds to integrate further communities.
(3) It will be open to interact and adapt to already existing initiatives. All this will be decided by the METHODS steering board in close interaction with the NFDI directorate. Due to this approach, METHODS needs to adhere to strict neutrality with respect to the discipline-approaches and traditions and respect the diversity of community methods and ethical standards. The diversity of methodological approaches implies within METHODS a universe of heterogeneous ethical standards. METHODS and its partners, having roots in high-quality data from randomized controlled trials (RCTs), respect the opposing views on deception (psychology vs. experimental economics), incentives for participants, dual-use issues in medicine, or anonymity in data publications in education research and qualitative, hermeneutic approaches.

Please indicate which of these cross-cutting topics your consortium could contribute to and how

(1) An integrated national data infrastructure is not equal to the set of all services offered by the variety of discipline/domain-based approaches. An integrated infrastructure needs to be based on a few key standards and components everyone adheres to. Wittenburg and Strawn (2018) have shown that successful large infrastructures such as the electric grid, the Internet and the
Web are based on a few simple standards that not only have an integrative function but the capability of picking up momentum. We believe that such a common infrastructure will be needed to make NFDI a success at the end. With FAIR DOs, such a core standard has been suggested, and it now needs to be implemented broadly to demonstrate its capabilities on the one hand and fine-tune its approach on the other hand. METHODS would achieve exactly this by being open to including other infrastructure initiatives in the developments and evolving testbed. Such a project in NFDI would also confirm the leading role of German data professionals in global discussions and implementations.

(2) NFDI needs to finally include all scientific disciplines and local research environments working with data-driven approaches. NFDI started funding consortia that are driven by domain-competence and coherence. However, this will only cover a limited range of disciplines and the overall impact on further communities could be limited as they need to address their challenges and remain widely within their traditions. To include the many other disciplines, one has to turn to a method-oriented approach that enables the easy integration of different solutions without requiring researchers to change their habits abruptly, but which offers the chance to adopt FAIR-compliant practices incrementally. For sports psychologists, for example, the relevant, meaningful community is not their discipline "sports science," but the methodological approaches of their discipline of origin, psychology.

(3) It is agreed that large infrastructures such as NFDI and EOSC should put the FAIR principles into practice, which, according to their creators, needs to go beyond a sole focus on findability and accessibility. Anticipating the challenges of the coming decades, turning FAIR principles into practice actually means to be prepared to make all data science artifacts (data, code, metadata, do-files, software, etc.) interoperable and reusable in the sense that machines know what to do. This is an ambitious goal and can only be realized in steps. Currently, FAIR DOs are the only known suggestion that makes it possible to implement such a landscape if openness is of relevance. The NFDI evaluators and the NFDI expert committee assessed in 2020 that in the implementation of the FAIR principles and the high-quality and connectable metadata standards for FAIR Digital Objects, current stakeholders of METHODS have impressive expertise and competence that can be invaluable to the NFDI as a whole. The NFDI needs the FDO-technology to not only be prepared for the coming challenges but to become one of its global drivers. FAIR DOs would be a core technology to connecting meaningful method-oriented communities across the disciplines – more important for the researchers would be, however, the implementation of a cross-disciplinary workflow framework based on FDOs and other technologies and common metadata frameworks such as CERIF (the Common European Research Information Format), DDI / DDI-CDI (DDI - Cross Domain Integration), Dublin Core, etc. and globally available PID solutions.

(4) Up to now, many universities have fairly general approaches to RDM and data silos that are not connected to specific structures. METHODS will implement generic toolboxes, and shops that can help to overcome this actual inefficient situation.