

# ForumX

## Letter of Intent

### 1 Binding letter of intent as advance notification or non-binding letter of intent

<input checked="" type="checkbox"/>	Binding letter of intent (required as advance notification for proposals in 2019)
<input type="checkbox"/>	Non-binding letter of intent (anticipated submission in 2020)
<input type="checkbox"/>	Non-binding letter of intent (anticipated submission in 2021)

### 2 Formal details

#### Planned title of the consortium

ForumX: An open forum on experiments across the disciplines

#### Acronym of the planned consortium

ForumX

#### Applicant institution

Otto-von-Guericke University Magdeburg  
 Universitätsplatz 2  
 39106 Magdeburg  
 Germany

Jens Strackeljan - Rector of the Otto-von-Guericke University Magdeburg (OVGU)

#### Spokesperson

Prof. Dr. Joachim Weimann  
 Otto-von-Guericke University Magdeburg  
 Faculty of Economics and Management  
 Universitätsplatz 2  
 39106 Magdeburg  
 Germany  
 joachim.weimann@ovgu.de

#### Participant

Will be submitted by 16 August at the latest according to the DFG supplement.

### 3 Objectives, work programme and research environment

#### Research area of the proposed consortium (according to the DFG classification system)

110 Psychology  
 111 Social Sciences  
 111-02 Empirical Social Research  
 111-04 Political Science

112 Economics

205 Medicine

206 Neurosciences

409-06 Information Systems, Process and Knowledge Management

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

The main objective of ForumX is to create an institution addressing the experimental method that is transversal to the social, economic, and human sciences. ForumX pursues the goal of making experimental data in the social, economic, and human sciences accessible for re-use and processing and provides means for curation as well as permanent, reliable storage in accordance with FAIR standards. The successful implementation of the FAIR principles into practice is influenced by a number of factors such as scientific objectives (anti-fraud, replicability, transdisciplinary reuse with rich context information), discipline-specific requirements and the infrastructural choices within ForumX which will be based on the FAIR-DO concept as worked out in RDA (Research Data Alliance) and supported by EC's FAIR Implementation Expert Group.

Experimental data are generated decentrally, and a functioning data infrastructure can only be created if the individual actors are willing and able to store their data voluntarily in a structured form. To solve the associated incentive problem, ForumX makes use of the following organizational structure: The organizational structure of ForumX replicates that of the DFG to organize the overarching methodological approach in its disciplinary specificities. All decisions are made by the members of the association. Members can be laboratories, individual scientists (after an election), scientific societies and NFDI institutions (consortia and other forms of organization). The general assembly decides on the procedure for the realization of the core concerns of ForumX. In terms of task areas, in the initial phase of ForumX (the first five years) the tasks will be divided into the following sections, which will be taken over by units organized on a project-by-project basis.

#### **1. Task area experimental economic and social sciences research**

ForumX has its roots in experimental economic and social sciences research, the specific needs of which are not addressed by any other NFDI consortium. The first step is, therefore, to develop generic solutions covering the entire research data cycle (standards, curation, workflows, basic infrastructure) with low-cost and low-effort research data management for the economic and social sciences. This is to ensure that all laboratories and data-generating actors who are members of ForumX have the possibility to store the experimental data in an automated workflow in such a way that the standards developed by ForumX can be met, no matter in which pillar of the existing research data infrastructures a laboratory or a researcher decides to store her data. This includes: (a) the rapidly growing field of "Data Lakes" in universities and laboratories, which are (still) far from what can be called "FAIR compliant", (b) research data in international virtual research environments such as the Open Science Framework OSF or commercial providers such as Mendeley Data and (c) relevant and trustworthy (certified, i.e. FAIR compliant) repositories that form the backbone of the many well-known infrastructures. In a second, path dependent step, these exemplary solutions with lighthouse character will be actively made available to other experimental communities and

consortia. In a third step, local adaptation and modification of these generic solutions will be supported.

**Task area strategies and procedures:** To make data and services developed in the various laboratories and disciplines ready for sharing and re-use, ForumX will work out guidelines and procedures helping to adhere to FAIR Maturity Indicators and Core Trust Seal criteria. This includes simple registries for registering standard procedures as described above and for vocabularies making use of existing components. Furthermore, strategies will be developed and implemented to disseminate these ForumX standards beyond the member laboratories. Close interaction with other interested consortia and with RDA Working Groups will be used to adhere to emerging trends.

## **2. Task area common ForumX infrastructures: Support, supervisory and control functions, training and teaching**

The independent projects within all task areas are actively supported by ForumX by defining a common infrastructure, common guidelines for service and support structures, a set of supervisory instruments, and control functions where this is necessary. This means that all strategic decisions with relevance to the functioning of the association are made by the association members, leaving enough space to do competitive data science in the disciplines and to develop the required federation infrastructure compliant with the general trends (RDA, EOSC, GOFAIR, CODATA). This will help to build generic infrastructure components and to avoid the creation of isolated solutions and duplicate structures. This includes knowledge transfer and consultancy on key competencies on data management strategies, data quality, standards, workflows, and the FAIR principles. In this task area, the closest cooperation partners within the NFDI are 2linkNFDI, NFDI-Neuro, and CompeNDI.

## **3. Task area community specific ForumX projects 1: Extension**

Experiments always consist of similar things: Testable hypotheses, a specific experimental design, a new set of primary data collected in a laboratory, online or in the field, and a structured data analysis. This is always the same structure. Therefore, ForumX also addresses experimental epidemiological studies (health economics, medicine), clinical studies with the need for platform-based data management and custodians to meet the specific legal (e.g. dual use) and incentive requirements (medicine, pharmacology), experiments in neuroscience, field- and laboratory experiments (social psychology, ethnology, psychology, marketing, etc.), experimental philosophy (X-Phi), and archeological experiments. In this funding program, the special conditions of the respective disciplines are dealt with and supported in cooperation with the respective communities. This funding program includes projects that develop solutions for one or more experimental disciplines. This ForumX funding program will make the expertise in FAIR data handling accessible across disciplines and scientific communities.

## **4. Task area community specific ForumX projects 2: Combination of data**

The second task area strives to develop procedures that help to combine data from experiments with data from other sources like survey data or unstructured online data. This combination of different data types is high on the agenda of many researchers because in this way, the external validity, especially of laboratory findings, can be increased. Thus, this also enables to account for the heterogeneity of participants preferences and personality traits. Possible targeted data are, for example, survey data, but also unstructured and Internet data.

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

ForumX builds on a repository for a comparatively large and established scientific community (experimental economics) and is deeply rooted in the methodological standards of the discipline. This repository has been deliberately designed as an open solution. ForumX has also developed a social sciences umbrella repository covering different types of experiments (laboratory, field experiments, survey experiments, etc.). Both existing modular infrastructures are essential to open up the possibility of connecting disciplines in which experimental methodology is only just developing and of extensions to include more advanced research data management functionalities. The infrastructures, workflows, and routines already developed by ForumX have proved their value so far. They enable a simple, low-cost entry into systematic data preparation and storage according to FAIR standards. Therefore, they have an example function and should be well suited to be adapted for other experimental approaches and to promote the integration of further disciplines.

ForumX will use and extend its existing infrastructures, tools, and services developed by its members and partners for collecting data (hroot, classEx, Lioness), data curation and friction free workflows (YARD, CentraXX), archiving (x-econ, x-science) and making data interdisciplinary available. ForumX intends to find infrastructure solutions with GWDG and MPCDF using their existing infrastructures.

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

- To maximize data sharing and re-use the core objective is to implement a generic and extendable infrastructure of key components that is actively provided to all ForumX members and that could be used by other NFDI consortia as well. Discussions with other consortia (Material Science, Environmental Science, Engineering Science, Linguistics & Humanities) revealed that there is a great interest to join forces and share labor to integrate existing components and develop yet missing components that are needed to implement the FAIR-DO concept and thus realize the NFDI networking plans. For the producers of more memory intensive data, e.g., from the astro- and geo-sciences, exchange and mutual support to the respective standards are currently intended.
- ForumX works together with the joint consortium NFDI4Medicine (DZG and MII) in the area of methodical data processing, on metadata standards, metadata mapping and harmonization, workflow standards, the further development of Snomed CT, and handling sensitive personal data from participants and patients in laboratory experiments and clinical studies.
- ForumX has the goal to develop common infrastructure for experimental research data. Furthermore, its concept foresees a strong role of the scientists in shaping the NFDI process, similar to the concept of NFDI-Neuro. We intend to collaborate on developing standards and services to implement the FAIR principles, and to interact on matters of governance and coordination with NFDI-Neuro.
- ForumX will join the data access network of KonsortSWD when it has established its research data centre services. It will benefit from the existing repository experience in the group of RDCs. Furthermore, the synergies with the research data centre PsychData can lead to harmonized metadata for experimental data. ForumX will close a gap within the

repository framework of KonsortSWD and profit from access to the broader data infrastructure as coordinated within KonsortSWD.

- ForumX and BERD@NFDI will strive for close collaboration. While BERD@NFDI deals with unstructured data from business, economics, and related fields, ForumX focusses on the experimental method. In fields of intersection (conjoint analysis in marketing – factorial surveys in sociology; unstructured finance data and experimental finance data), both consortia will work closely together. Fields of collaboration are (1) metadata standards and models of metadata harmonization, (2) transdisciplinary research regarding methods and algorithms (AI), (3) the combination of similar experimental datatypes in different communities (factorial surveys in sociology vs. conjoint analysis in marketing).
- Collaboration with 2linkNFDI and PNet, if participating in the NFDI, as strong partners in their respective core competencies, are desirable and discussed.
- Cooperation on common NFDI standards in experimental data with NFDI4Culture, NFDI4Objects, and Text+ are intended.
- ForumX will liaise with the "Forum geisteswissenschaftliche NFDI" as a way of coordinating infrastructure and subject-specific requirements between the Social Sciences and the experimentally oriented Humanities.

#### 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.**

ForumX is developed by experimental researchers and thus takes a strict researcher and research-oriented perspective. ForumX addresses all experimental researchers, all experimental disciplines, and all data types across the entire research data cycle. It is able to integrate all existing repositories and data storage solutions. Cross-cutting topics that are relevant for ForumX can be assigned to two categories: (I) Optimal services and structures addressed by existing high-quality research data infrastructures. Currently, only a marginal proportion of the experimental economic and social science research data is archived so far in these infrastructures. (II) Most of the experimental data addressed in ForumX is stored, for example, on personal devices, in data lakes, or university archives. Discussions with other consortia have revealed that this is a cross-cutting topic to consortia within the NFDI and data generating communities, as well as disciplines so far not part of the NFDI (sports, law, etc.) with a relatively young data management tradition. Moreover, from an overarching perspective, the reality of research data management may be even more dramatic. With the market entry of commercial players, more and more data is stored in data silos. ForumX strives to address this cross-cutting issue by including the forgotten research data and make it available to scientific communities. This topic is not only relevant for ForumX, but to all NFDI consortia. It is not overarching the quality pillars, but a vertical development and a common task. ForumX intends, therefore, to establish a FAIR DO ecosystem.

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

Specific cross-cutting topics to be addressed by ForumX are

- An evaluation of a variety of tools being used in local setups would allow participants in ForumX to see which tools are available and which of them can be seen as best practice solutions for example from medicine (Kairos: CentraXX) or political science (ISPS Yale:

YARD). These can be easily implemented into other disciplines and thus common, generic solutions can be created. Here ForumX can take a role in improving coherence and reducing efforts.

- In particular, guidelines and reference implementations will be developed to set up FAIR and CTS compliant repositories and research data centers (FDZ) that are ready to be adapted and connected to the cross-disciplinary infrastructure. FAIR and CTS compliant repositories (research data centers) supporting the FAIR-DO concept have already indicated to be transdisciplinary, since they can handle, for example, different metadata sets and offer them via well-known protocols such as OAI-PMH. A harmonization of core metadata elements to enable semantic mapping would be useful also to support a generic catalog. Existing approaches (DataCite, etc.) and mapping techniques (EUDAT B2FIND) in this area need to be evaluated and eventually adopted.
- The generic and extendible ForumX data and service infrastructure will be based on the FAIR-DO concept which per definition is cross-disciplinary and therefore highly suitable to implement federated infrastructures. It thus allows to connect hitherto isolated solutions and data silos (laboratories, computer centers, universities, virtual research environments). Essential components of this FAIR-DO concept are
  - The universally used Handle System to associate persistent, unique, and globally resolvable IDs with every digital object be it data, metadata, software, assertions, etc. This system has been proven functioning for 20 years and is now widely used in the form of ePIC Handles (provided by GWDG) or DOIs (provided by DataCite). Eventually, ForumX will run its own Handle service as part of the redundant ePIC federation to have maximal degrees of freedom in using PIDs and associating crucial attributes with it. Using such PIDs is already implementing the first FAIR principles.
  - PIDs are being resolved to information as specified by the community. These can be information such as checksums to prove authenticity, version indications, but also pointers to different types of metadata such as descriptive, deep scientific, provenance, rights, blockchains to store transactions, etc. Using these attributes satisfies additional FAIR criteria.
  - PIDs allow users to assign "types" with each object and using type registries as provided by GWDG would allow to associate operations with types and thus offer a way towards automated processing.
  - DOIP as a generic interface to interact with Digital Objects of any type stored in repositories.
- Adaptors to DOIP and the Handle System that can be easily used by the various repositories and also allowing to integrate typical systems such as Fedora, D-Space, OSF, Github, etc.
- For all these measures it is important for ForumX to work together with the existing broad international networks working in this area, adopting components where possible, making use of existing services and collaborate with others including several NFDI initiatives applying the same approach.
- Transdisciplinary research as well as further development and implementation of the approaches developed by the group at the Department of Economic Sociology at the University of Vienna.

- With NFDI4Medicine and NFDI-Neuro, ForumX wants to support and execute the implementation of Snomed CT as standard and wants to use its methodical approach for faster data mapping within medicine and its methods for other research fields.