Manifesto of the Economic and Social Sciences goINg FAIR Implementation Network [EcoSoc-IN]

Introduction:

The German Data Forum (Rat für Sozial- und Wirtschaftsdaten, RatSWD) proposes to establish EcoSoc-IN, an implementation network for the social, behavioural, and economic sciences. Established in 2004, the German Data Forum (RatSWD) is an independent council. It advises the German Federal Government and the federal states (Länder) in matters concerning the research data infrastructure for the empirical social, behavioural, and economic sciences. The German Data Forum (RatSWD) has accredited 31 research data centres (RDCs) according to set standards and fosters their interaction and cooperation. The RDC Committee supports the German Data Forum (RatSWD) and coordinates cooperation among RDCs with the long-term aim of safeguarding and continuously improving the availability of research data in the social, behavioural, and economic sciences. These RDCs offer flexible and secure access to a broad range of data (more than 3,500 datasets with more than 21,000 users in 2017) to the international scientific community. The data – often personal or institution-related – are provided by state agencies, research institutes, and private-sector actors and are consequently subject to special legal requirements, such as general data protection regulation or confidentiality under statistics and tax laws.
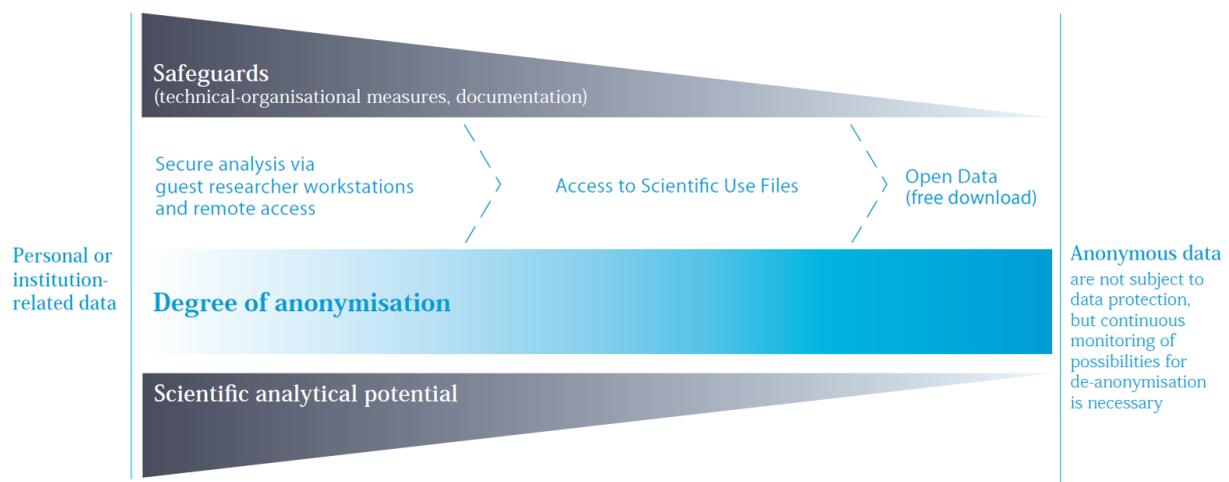
Purpose of the Implementation Network:

The German Data Forum (RatSWD) and all participating RDCs are committed to the continuous improvement of the research data infrastructure and the implementation of the FAIR principles. Since the German Data Forum's (RatSWD) foundation, its guiding principles and accreditation criteria for RDCs have promoted the findability, accessibility, and reusability of high-quality data. The coordination, monitoring, and strong cooperation between RDCs safeguard the continuous improvement of the research data infrastructure, for example, by enhancing the accessibility and interdisciplinarity of RDC services. The systematic implementation of persistent identifiers in most RDCs was a milestone for increasing data findability.

A key component of the RDCs' mission is to comply with the legal requirements of privacy protection. They do so using the Five Safes framework, which aims to preserve the analytical potential of the data for researchers in the best possible way.[1] By working closely together, the RDCs have developed a range of data access paths depending on the required degree of data anonymisation (see Fig. 1 below).

The EcoSoc-IN intends to share its expertise on all aspects of accessing sensitive data from the social sciences, economics, and behavioural sciences. Moreover, we aim to develop solutions to enhance the interoperability of available datasets.

Fig. 1: Anonymisation as risk assessment

Solutions addressing the tradeoff between anonymity and analytic potential of research data.

In doing so we commit to the following guiding principles:

[**Data culture**] We are committed to raising awareness for the significance of data sharing, data provision, and data curation among researchers and funding organisations. We are convinced that even highly sensitive data can and should be made available for research purposes with appropriate safeguards of data protection in place.

[**Free and equal data sharing**] We are committed to ensuring equal treatment of all scientific data users by means of transparent and standardised application and access policies.

[**Data protection by design**] We are committed to ensuring that data usage complies with European and member state data protection policies and, if applicable, with policies specific

---

[1] Compare page 12 in Ritchie, Felix (2017): Spontaneous recognition: an unnecessary control on data access? European Central Bank Statistics Paper Series (24). DOI: http://dx.doi.org/10.2866/430525 (accessed on 01.11.2018).

to individual research areas by taking appropriate technical and organisational measures. Depending on the level of anonymisation, RDCs provide different data access paths.

[Research data repositories] We promote the model of independent and accredited RDCs, which provides researchers with user-friendly, transparent, and trusted access to high-quality microdata. The RDCs play an important mediating role between data users and data producers and foster cooperation between them.

[Accreditation] We promote the organisational principle of accreditation. The German Data Forum (RatSWD) developed minimum standards and transparent criteria to accredit RDCs. To receive full accreditation, RDCs must provide at least one data access path, provide sufficient data documentation, and ensure the long-term availability of their research data. These standards and criteria foster coordination and cooperation among the RDCs which work together in the RDC Committee of the German Data Forum (RatSWD). The exchange of knowledge and the communication of best practice models within the committee ensure the quality of the research data infrastructure. Regular monitoring serves the continuous evaluation and enhancement of the research data infrastructure.

[Skills] To gain better recognition for data sharing and the RDC model, it is essential that RDCs provide trustworthy, reliable, and high-quality services. They support data users in various ways and build the competencies required to use their specialised data.

[Data stewardship] Most academic staff in RDCs conduct independent research using their data holdings, which ensures that the RDC staff possess unique expertise regarding the content, quality, and potential of its data. This is a key factor in ensuring the creation and provision of user-friendly data products, which are tailored to the scientific needs of specific disciplines or subjects and are not limited to quantitative data but may also comprise digital object types like text, audio, and film. Furthermore, RDCs provide high-quality standardised metadata, comprehensive documentation, training materials, and support for data analysis.

[Persistent Identifiers] We promote the registration of datasets with providers of persistent identifiers like da|ra in order to make them findable and identifiable through digital object identifiers (DOI) and rich meta-data.

[Interoperability] We are committed to optimising the interoperability of the datasets we offer, for example, by means of common metadata schemes**.** The RDCs are working on achieving this while taking domain-specific requirements into account.

[Distributed and federated solutions for data access] We are committed to further developing legal and technical concepts to facilitate the use of secure environments within RDCs to access the data of other RDCs.

[Reputation and incentives] We promote data production, data archiving, and data sharing as essential services for the scientific community, which should receive greater recognition within the academic reputation system.

Overarching Principle of Operation:
We commit to comply with the Rules of Engagement of GO FAIR Implementation Networks.

Targeted Objectives for the Internet of FAIR Data and Services (IFDS):
1. Fostering implementation of FAIR principles in social science and economics data centres. Based on a careful analysis of the status quo in the participating RDCs, we will identify possible gaps in the implementation and develop targeted solutions to close said gaps. A major target is to improve the quality of metadata and to continue to harmonise the standards used for data documentation. Our goal is to systematically provide machine-readable, detailed metadata along with the datasets provided by the RDCs. The implementation will take the special requirements of sensitive data into account.

2. Securing and expanding access for social, behavioural, and economic researchers to high-quality research data, for example, by supporting and advising new data centres. The experts in the EcoSoc-IN will offer advice on issues ranging from contract design to solutions for secure access to sensitive data.

3. Continuously developing data access paths and services by exchanging knowledge about best practice models (e.g., automated output control to guarantee data protection) and by working together in thematic working groups.

4. Monitoring of the research data infrastructure to assess its impact and to regularly evaluate compliance with accreditation criteria.

5. Fostering reciprocal networks of access to sensitive data and to provide multiple access points for research data on a European and an international level.

6. Exchanging knowledge between academic disciplines and initiatives through regular meetings, workshops, and publications, and fostering access to highly sensitive data by promoting RDCs as a best practice model. Access paths include, for example, secure remote access, guest researcher workstations, and scientific use files (see Fig. 1 above).

7. Engaging with other GO FAIR implementation networks to exchange best practice solutions and increase interoperability.

8. Continuously developing cross-disciplinary and international links of the network.

Membership list:
We consider this Manifesto to be one way by which the undersigned stakeholders can **speak**

with one voice on critical issues that are of generic importance to the objectives of FAIR, and on which we feel we have reached consensus.

- German Data Forum (RatSWD)

The Network is represented by the chairs of the German Data Forum (RatSWD)

- Prof. Dr. Monika Jungbauer-Gans
- Prof. Dr. Kerstin Schneider

and the chairs of the RDC Committee

- Dr. Daniel Fuß
- Tatjana Mika

The Network is coordinated via KonsortSWD by

- Daniel Buck
- Ute Hoffstätter