



Juni 2024

Study accompanying pilot projects developing and testing data trustee models in science and business



Paper – Executive Summary



Study accompanying pilot projects developing and testing data trustee models in science and business

Paper – Executive Summary

Technopolis Group: Stephan Kreutzer, Prof Dr Thomas Heimer, Fiona Bauer, Lea Rabe

Fraunhofer ISI: Prof Dr Knut Blind, Dr Nicholas Martin

Law & Innovation: Prof Dr Max von Grafenstein, LL.M.

GRI GmbH, RWTH Aachen: Prof Dr Rita Streblov, Junsong Du, Joel Schölzel

This study has been commissioned by the German Federal Ministry of Research and Education (co-financed by the programme NextGenerationEU of the European Union)



Executive Summary

Shared data is considered **a key driver of value creation**. Shared data is valuable for research, innovation and value creation and helps to solve social challenges such as the development of sustainable mobility concepts, research into complex climate models or the fight against rare diseases. At the same time, data is often fragmented, both within industry and the scientific community, and in between them. **Data trustees** can act as intermediaries between data providers and potential users, overcoming information asymmetries and thus contributing to data sharing.

Since 2021, the **German Federal Ministry of Education and Research (BMBF)** has been **funding 20 projects for the development and practical testing of data trustee models** in research and industry testing out data trustee models in various domains ranging from healthcare to wind energy.

A consortium consisting of Technopolis Deutschland GmbH, Fraunhofer ISI, the Berlin University of the Arts (Universität der Künste, UDK) and RWTH Aachen is conducting a **study** accompanying the implementation of the pilot projects. This paper presents preliminary results of the ongoing study.

The study identifies implementation challenges and solutions for data trustee models across four horizontal topics: **acceptance, legal framework conditions, business model development, technical infrastructure, standardisation and certification**.

First, though, a number of **challenges can be identified when establishing data trustee models:**

The central task of a data trustee is to increase cost efficiency, general usefulness and trust in a data ecosystem to such an extent (i.e. to reduce costs for data users and providers and increase benefits to such an extent) that data sharing becomes attractive for all participants. The **acceptance of data trustee models**, as in the willingness of (potential) data providers and users to interact via a data trustee, depends on the costs and risks on the one hand, and the benefits and trust, as well as possible alternatives, on the other. Ensuring **data protection, data security and data sovereignty** via technical, organizational and regulatory measures can reduce risks and establish trust. **The neutrality of the data trustee model** or the operating organisation - that its interests do not conflict with those of the data providers and users - is generally seen **as a prerequisite for the acceptance of data trustee models** in the data ecosystem. However, how best to achieve such neutrality is an open question. Another prerequisite is trust in the security of data trustee models. Most importantly, data providers and users need to be confident that data-sharing will create value for both of them. This is made more difficult by the fact that the actual value creation will only be ascertained once the data has already been shared. Data trustees must thus act as **brokers and/or matchmakers**, i.e. identify possible use cases, and proactively bring together potential data providers and users.

Against this backdrop, the accompanying study identifies **legal uncertainty** as a key obstacle to the establishment of data trustee models. Issues relating to data protection, the handling of business secrets, competition law, legally compliant technical and organisational implementation, the distribution of roles in the data ecosystem and unresolved liability issues are slowing down the development of the technical infrastructure and potential business models. One way to classify legal compliance risks is to distinguish between **three models of intermediated data-sharing**. The first is an **open data model** in which legal compliance risks for

data users are relatively low and data intermediary activity by a trustee only required for technical and organizational elements. The second is a **"shared data" model**, in which the compliance risks for users are higher because data may include business secrets or personal data. Here, the data trustee may need to provide additional protection mechanisms such as secure data processing spaces. In the third, **"shared analysis results" model**, users do not get access to the data provided directly, but rather can make enquiries to data providers via the data trustee. This third model is particularly interesting for data trustees, not least because of its high scalability potential: Enquiries frequently made by data users can be offered as a standardised service by the data trustee.

Defining the functions of a data trustee model is of central importance when it comes to designing **sustainable business models**. The preliminary results of this study indicate that such functions are seen on the one hand in the **guarantee of control options, participation, self-determination** and the **involvement of data providers** in accordance with data protection regulations, and on the other hand in increased **economic data usability** to promote innovation, competition and value creation. The **value of data trustee models** then arises when the collection, access and use of data not only provides an advantage to the individual data providers and recipients, but also creates benefits for a broader society from the use of the data, despite all the risks.

The **funding of data trustee models poses** a challenge. There is a tension between primarily profit-orientated business models, neutrality requirements and ethical aspects of data sharing. Based on the evidence assessed for this study, it is generally recommended that business models for Data Trustees should not be based on the pricing of the data itself, but rather on fees or "flat rate models". A distinction can be made here between subscription models, fixed-price models, pay-per-use models, package price models, memberships, transaction fees and fees for services. Pricing based on the respective data pools could provide an answer to the risk of **monopolisation** of data trustee model markets, which has been mentioned in numerous discussions in the literature, as a variety of pricing approaches can serve as an instrument for creating organisational diversity in the data trustee models on offer.

As regards the **technical implementation** of data trustee models, **general technical requirements** relate to guaranteeing data transfer and interoperability, ensuring data security and sovereignty, and creating added value through quality management and effective metadata management. In order to achieve these goals, a distinction is made between centralised and decentralised architectures. In **centralised architectures**, the data trustees act as service providers who establish a business relationship between data providers and data users. A central database is usually set up for this purpose, which collects the data from the data providers and makes it available to the data users. In **decentralised architectures**, a data ecosystem can be set up and managed by the data trustees. Participants can publish their data description and metadata and search for suitable data exchange, whereby the data sets do not have to leave their data holders. The participants often communicate via a uniform interface, such as standardised application programming interfaces (API) or connectors. Based on the empirical work, the accompanying study identifies **three central challenges and barriers in the area of technical infrastructure**: high development effort, diversity of technical implementations and balancing data security and usability.

Standardisation (data trustee model-specific standards, data standards and established cross-industry international standards) and **certification** can play an important role in the potential **scaling of solutions** for data trustee models in various application areas. At the same time,



empirical evidence shows that there are still few concrete efforts by relevant players and institutions in this area.

At present, it is **not possible to** identify a single **data trustee model that works equally well for all domains**. Rather, **a toolbox of services and governance models** appears to be in the making enabling solutions that can be applied to various fields. In the further course of the accompanying study, emerging solutions for the successful establishment of data trustee models will be systematically analysed.¹ A final report is expected to be available by mid-2025.

¹ In a further funding guideline, the BMBF is supporting the research and development of practical solutions for data trustees ([link](#))



technopolis
group 

www.technopolis-group.com