5-11-2023

Unleashing The Potential of Data Ecosystems: Establishing Digital Trust through Trust-Enhancing Technologies

Fabian Schäfer
University of St. Gallen, fabian.schaefer@unisg.ch

Jeremy Rosen
University of St. Gallen, jeremy.rosen@student.unisg.ch

Christian Zimmermann
Robert Bosch GmbH, christian.zimmermann3@de.bosch.com

Felix Wortmann
University of St. Gallen, felix.wortmann@unisg.ch

UNLEASHING THE POTENTIAL OF DATA ECOSYSTEMS: ESTABLISHING DIGITAL TRUST THROUGH TRUST-ENHANCING TECHNOLOGIES

Research Paper

Fabian Schäfer, University of St.Gallen, Switzerland, fabian.schaefer@unisg.ch

Jeremy Rosen, University of St.Gallen, Switzerland, jeremy.rosen@student.unisg.ch

Christian Zimmermann, Robert Bosch GmbH, Germany, christian.zimmermann3@de.bosch.com

Felix Wortmann, University of St.Gallen, Switzerland, felix.wortmann@unisg.ch

Abstract

Companies increasingly innovate data-driven business models, enabling them to create new products and services. Emerging data ecosystems provide these companies access to complementary data, offering them additional potential. This, however remains untapped, as a lack of digital trust prevents companies from sharing data within these ecosystems. By using trust-enhancing technologies, companies can establish trust; this can be explained through the theoretical lens of system trust. Using a design science research approach helped us to unlock the knowledge of 21 experts and identify five technologies with the potential to solve the trust challenge: self-sovereign identities, differential privacy, fully homomorphic encryption, trusted execution environments and secure multiparty computation. We integrated these technologies into the data sharing process in data ecosystems and elaborated on their limitations and maturity. Ultimately, we derived two principles that allow for adapting our results to future technological developments: complementarity and customization.

Keywords: Data ecosystems, Digital trust, Trust-enhancing technologies, System trust.

1 Introduction

In the realm of the networked economy, companies are increasingly sourcing data from smart connected products and digital applications and creating new data-driven business models (Gassmann and Ferrandina, 2021; Fadler and Legner, 2022). In these business models, data is considered a key resource, allowing companies to innovate their products, processes and services (Hartmann et al., 2016; Porter and Heppelmann, 2014). Such transformation requires companies to rethink their value creation structures and leads to the emergence of so-called data ecosystems (Gelhaar and Otto, 2020). In these ecosystems, companies can get access to complementary data from data providers (Chen et al., 2011; Russo and Albert, 2018). Hence, data becomes a shared resource thus offering significant potential for revenue growth for participating companies (Jiang et al., 2021; Oliveira et al., 2019).

However, the potential of data sharing is not yet being tapped in company ecosystems. In the EU Data Strategy, the European Commission states “in spite of the economic potential, data sharing between companies has not taken off at sufficient scale” (European Commission, 2020, p. 7). One major reason for this is a lack of trust between participating actors (European Commission, 2020; Gelhaar and Otto, 2020; Sahut et al., 2022). Although certain instruments exist that facilitate trust between these actors (e.g., contracts, certifications, guarantees), the degree of trust is often insufficient, thus causing every second ecosystem to fail (Aguiar et al., 2021). One ground for the lack of trust is that ecosystem actors fear other parties will not use shared data in line with contractual agreements, regulations and ethical norms (Aguiar et al., 2021; Culnan, 2019; European Commission, 2020). Furthermore, the increasing usage of Internet of Things (IoT) technologies and cloud solutions for data sharing exposes companies and their data to new cybersecurity risks (Aguiar et al., 2021; European Commission, 2020).

In addition, the emergence of new digital technologies has changed the way in which trust is established between companies in today's digital economy. Trust between companies is increasingly replaced by trust in a system based on digital technologies (e.g., secure multiparty computation, blockchain technologies) (Agahari et al., 2022; Lumineau et al., 2023; Mubarak and Petraite, 2020). While there are several reasons for lack of trust in data sharing, it is unclear which digital technologies have the highest potential to enhance trust in data sharing processes in data ecosystems and how they can be combined in a system (Mubarak and Petraite, 2020; Sahut et al., 2022). Accordingly, we conducted an exploratory study to shed light on the potential of digital technologies for creating trust in data sharing processes. We addressed the following two research questions:

RQ1: What trust-enhancing potential for data sharing in data ecosystems do the most promising digital technologies provide, and what are their limitations and levels of maturity?

RQ2: How can these technologies be integrated into the data sharing process in data ecosystems to enhance trust?

Our study followed a design science research process and can be summarized in three steps. First, we elicited design requirements that address trust challenges in the data sharing process in today's company data ecosystems. Secondly, we assessed the potential of digital technologies to fulfill these requirements, their limitations and maturity with regard to creating trust based on insights from experts in both data ecosystems and data sharing digital technology. Thirdly, we focused on the most promising technologies and proposed a design for integrating them into the data sharing process.

Our answers should increase awareness of technological opportunities for improving trust in systems, support company executives in their digital technology investment decisions and help them to gain a competitive advantage (Boehm et al., 2022; Kluiters et al., 2023).

2 Theoretical Background

2.1 Data Ecosystems and Trust as a Major Challenge for Data Sharing

Company data ecosystems are networks of actors, including companies and individuals, for whom data exchange is increasingly enabled by interconnectivity through the IoT, cloud computing and digital platforms (Beverungen et al., 2022; Curry & Sheth, 2018; Oliveira et al., 2018). According to Oliveira and Lóscio (2018) and Gelhaar and Otto (2020), these actors can be categorized into at least two parties: data providers and data consumers. The party that collects data and shares its data in a data ecosystem being the *data provider* and the party that obtains and generates value with this data (e.g., by offering a digital service) the *data consumer* (Badewitz et al., 2020; Gelhaar and Otto, 2020). Data sharing is the process of providing a specific data set for usage under defined conditions (Dalmolen et al., 2019; Jarke et al., 2019; Jussen et al., 2023). Recent IS literature describes the detailed process phases and related activities that are performed by one or both parties when data is shared. For instance, based on a literature review, Jussen et al. (2023) derived a five-phase process for data sharing (preparation of a data set, establish a data sharing agreement, planning of the data trading process, process of data sharing transaction and feedback to the actors). Dalmolen et al. (2019) focus on four phases that are required for the sharing of metadata (e.g., agreement, policies). These are defining and publishing a data set; making a data sharing agreement; performing a data sharing transaction; and logging, provenance and reporting. Krasikov et al. (2022) took a data consumer perspective decribing a process for sourcing and managing external data including six phases: start; screen; assess; integrate; manage and use; and retire. Ultimately, depending on the perspective (e.g., from data consumer, focus on meta data), the process has a different starting point or is structured differently. However, between the process models, there is a large overlap in terms of covered data sharing activities.

A lack of trust in the data consumers within this data sharing process is key to companies' inability to leverage data sharing potential in data ecosystems (Gelhaar and Otto, 2020). This lack of trust is specified in existing literature that either focuses on specific trust challenges in the realms of data evaluation (Azkan et al., 2020; Song et al., 2021); securing data in use or the sharing of personal data (Culnan, 2019; Jansen, 2011; Weiss, 2018); or gives a broad overview of these challenges (Garrido et al., 2022; Otto and Jarke, 2019). It predominantly results from data providers' fear their sensitive and confidential data may be misused by data consumers or shared further with third parties that misuse the data (Agahari et al., 2022; European Commission, 2020; Gelhaar and Otto, 2020). Such misuse can threaten business and put a data provider's competitive advantage at risk and can be rooted in intended or unintended violations of contractual agreements, regulations or ethical norms (Agahari et al., 2022; Culnan, 2019; Guo et al., 2018; Joshi and Wade, 2020). The Cambridge Analytica (CA) scandal shows how shared data from 87 million Facebook users was misused by CA, violating contracts and resulting in a major loss of trust for Facebook (Casadesus-Masanell and Hervas-Drane, 2020). While intended violations are often driven by opportunism and bounded rationality (Lumineau et al., 2023; Simon, 1957; Williamson, 1985), unintended violations can be caused by the legal uncertainty that comes with new or upcoming regulations. As recent studies show, this is particularly the case around data privacy regulations (e.g., General Data Protection Regulation) (Bitkom, 2022).

To mitigate these uncertainties, data providers demand privacy and security standards from partners in ecosystems that match or better their own (Garrido et al., 2022; Weiss, 2018). In addition, companies seek to retain control over the data they share, including the decisions with whom, and when to share what data (Otto, 2019; Pigni et al., 2016). This enhancement of privacy, cybersecurity and self-determination of data use provides a basis for increased trust (Casadesus-Masanell and Hervas-Drane, 2020; Lumineau et al., 2023). Although self-determination of data use in the case of personal data is often associated with data privacy (Westin, 1967), in a more general context it meets the definition of data sovereignty: self-determination regarding the use of data (Jarke et al., 2019). In contrast, cybersecurity focuses on protecting digital information from unauthorized access and maintaining confidentiality, integrity, and availability of digital assets (von Solms and von Solms, 2018; ISO 27032, 2012). In summary, the digital trust dimensions data privacy, data sovereignty and cybersecurity are central for the success of data ecosystems as they form the basis for trustworthy data sharing. Recent research identified these dimensions as the pillars of trust for the digital economy and grouped them under the umbrella term digital trust (Abraham et al., 2019; Boehm et al., 2022; Mubarak and Petraite, 2020; Sahut et al., 2022; Wang et al., 2020).

2.2 Digital Trust and Trust-Enhancing Technologies

In establishing digital trust, emerging digital technologies such as secure multiparty computation or blockchain-powered smart contracts are valuable instruments for strengthening data protection, data security and data sovereignty (Agahari et al., 2022; Lumineau et al., 2023). We use the term trust-enhancing technologies to denote digital technologies that address the digital trust dimensions for data sharing. To take these technologies and their trust-enhancing potential into account, in our study we applied the theory of *system trust*. This theory enables a better understanding of how trust is built in the digital economy, as it not only accounts for the data provider and the data consumer, but also for the regulations, standards and norms that span the system boundaries and digital technologies that can play a major role in building trust (Agahari et al., 2022; Lumineau et al., 2023; Sumpf, 2019).

While the most widely adapted inter-personal (or inter-organizational trust) theory developed by Mayer et al. (1995) requires a direct bond between two interacting parties, system trust describes a form of trust placed in the functioning and reliability of a socio-technical system (Bachmann, 2003; Giddens, 1990; Luhmann, 1979; Lumineau et al., 2023). In such socio-technical systems, a mix of different non-digital and digital instruments engenders trust amoung system actors (Aguiar et al., 2021). As shown in Figure 1, digital technologies that fall under the second category of instruments can enhance trust by mediating data sharing processes between data providers and data consumers (Lumineau et al., 2023).

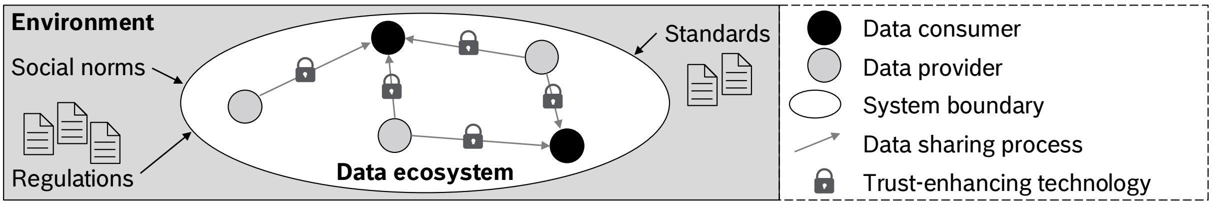


Figure 1. Data sharing and trust-enhancing technologies in data ecosystems.

There are several streams of literature that investigate these trust-enhancing technologies; one major stream is the research on privacy-enhancing technologies (PETs). Many of these technologies go beyond tackling pure privacy issues and contribute to enhancing data security and sovereignty. Studies on PET in data sharing either give a broad overview of the PET landscape (Garrido et al., 2022; Goldberg et al., 1997; Van Blarkom et al., 2003) or focus on adoption barriers (Zöll et al., 2021). Other studies analyze the use of a specific technology or a combination of these. For instance, Agahari et al. (2022) explored the use of secure multiparty computation in the automotive industry and Mo et al. (2021) combined privacy-preserving federated learning with trusted execution environments.

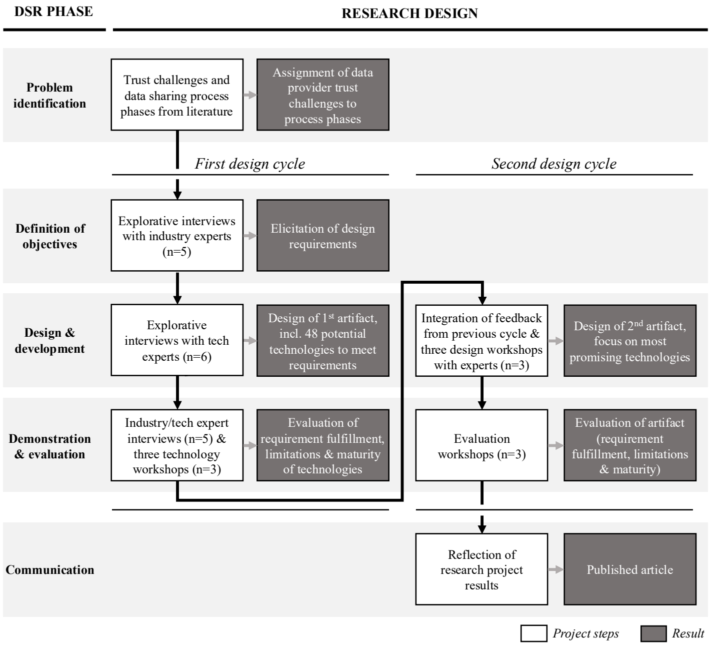
In the context of cybersecurity, some studies deal with implementing data sharing security mechanisms (e.g., attribute-based encryption, proxy reencryption) (Koo et al., 2013, Ali et al., 2015; Liu et al., 2014), other studies evaluate the potential of emerging technologies (e.g., blockchain and smart contracts) for secure data storage and sharing (Huang et al., 2018; Kang et al., 2018).

With regard to data sovereignty, research is ongoing particularly in the context of data spaces; this research is accompanied by the European Commission and its legislative initiatives (e.g., Data Act). Theoretical and legislative efforts have been translated into practical initiatives such as International Data Spaces and Gaia-X. In this context, publications focus on the creation of entire trust infrastructures combining digital technologies that ensure data sovereignty (Braud et al., 2021; Dalmolen et al., 2019; Otto, 2022; Otto and Jarke, 2019). Another literature strand focuses on the governance of such infrastructure (cental vs. decentral). For instance, Lumineau et al. (2021) and Ballatore et al. (2022) focused on blockchain technologies for decentralized data sharing governance systems. In conclusion, we see trust-enhancing technologies have been studied across different facets of data sharing, however, an elaboration of the most promising technologies from a data provider perspective and their application within a data sharing process has yet to be conducted. This article attempts to close this research gap.

3 Methodology

To address this research gap and to ensure the practical relevance of our study, we conducted design science research (DSR) (Hevner et al., 2004), following the DSR approach layed out by Peffers et al. (2007). Responding to the existing trust challenge for data providers in data ecosystems, we opted for an objective-centered solution entry (Peffers et al., 2007), seeking to design an artifact that addressed industry needs by fulfilling requirements for building system trust (Hevner et al., 2004). Thus, we designed an artifact that combines a set of trust-enhancing technologies in the context of a data ecosystem (Benbasat and Zmud, 2003; Hevner et al., 2004). In order to make these digital technologies work together, we integrated them in a data sharing process (Orlikowski and Iacono, 2001).

Our DSR process (cf. Figure 2) starts with upstream problem identification, which is followed by two design cycles encompassing objective definition and two iterations of design and development as well as demonstration and evaluation, and ends with post-analysis communication. It combines several data collection methods including 16 interviews and nine workshops with 21 experts (P1–P21; cf. Figure 2) resulting in about 24 hours of conversation time. We applied purposive sampling to gain access to experts from a broad range of industries (automotive, financial services, information and communication technology, manufacturing) (Saunders and Townsend, 2018). Only interviewees with three or more years of theoretical or practical experience in various data sharing constellations in data ecosystems were considered as experts. We selected experts from Europe due to its current data space and legislative initiatives that are highly relevant for establishing digital trust (European Commission, 2020).


DSR PHASE
RESEARCH DESIGN
Problem identification
Trust challenges and data sharing process phases from literature
Assignment of data provider trust challenges to process phases
First design cycle
Second design cycle
Definition of objectives
Explorative interviews with industry experts (n=5)
Elicitation of design requirements
Design & development
Explorative interviews with tech experts (n=6)
Design of 1st artifact, incl. 48 potential technologies to meet requirements
Integration of feedback from previous cycle & three design workshops with experts (n=3)
Design of 2nd artifact, focus on most promising technologies
Demonstration & evaluation
Industry/tech expert interviews (n=5) & three technology workshops (n=3)
Evaluation of requirement fulfillment, limitations & maturity of technologies
Evaluation workshops (n=3)
Evaluation of artifact (requirement fulfillment, limitations & maturity)
Communication
Reflection of research project results
Published article
Project steps
Result

INTERVIEWEES

ID	*Organization*	*Position*	*Experience*
P01	Global universal bank	Head of emerging technologies	10+ yrs.
P02	Information technology service provider	Country general manager	10+ yrs.
P03	Smart home IoT data platform	CEO	10+ yrs.
P04	Data marketplace	CEO	10+ yrs.
P05	Web3 - decentralized finance provider	Founder	3 yrs.
P06	Research institution	Doctoral researcher (DLT, smart contracts)	5 yrs.
P07	Multinational technology company	Research engineer (privacy engineering)	7 yrs.
P08	Research institution	Doctoral researcher (blockchain)	3 yrs.
P09	Multinational political institution	Head of sector IoT	10+ yrs.
P10	Data marketplace	Head of sales	10+ yrs.
P11	Big data software provider	Director of privacy and civil liberties	10+ yrs.
P12	Multinational political institution	Cabinet expert	10+ yrs.
P13	Research institution, data space initiative	Researcher data space, data ecosystems	5 yrs.
P14	Multinational technology company	Software engineer (industrial IoT)	10+ yrs.
P15	Decentral data marketplace	Head of infrastructure development	10+ yrs.
P16	IT service company (data privacy)	Founder	10+ yrs.
P17	Multinational technology company	Researcher (web3 strategy, IoT)	4 yrs.
P18	Telecommunication provider	Head of digital trust	10+ yrs.
P19	Multinational technology company	Research engineer – distributed systems, DLT	5 yrs.
P20	Emerging technologies & dataspace provider	CEO	10+ yrs.
P21	Emerging technologies & dataspace provider	CTO	10+ yrs.

Figure 2. Design science research process and overview of experts.

In the *problem identification phase* we built on the existing knowledge base as provided by the literature on data sharing challenges (see Section 2.1). In doing so, to pre-structure the challenges, we allocated them to data sharing process phases (see Section 2.1) where a data provider and a data consumer interact and thereby need to trust each other to successfully share data.

In the first design cycle, in the *definition of objectives* phase, the data sharing process phases and the assigned trust challenges were discussed in five interviews with experts to ensure their consistency with practice. In the second step of the first phase, we analyzed the challenges with these experts in more detail and elicited design requirements for system trust.

In the *design & development phase* of the first design cycle, we aimed at identifying technologies that enhance trust in data sharing scenarios. Building on the literature regarding trust-enhancing technologies (see Section 2.2), we conducted explorative interviews with six experts in PET, distributed ledger technologies, and cryptography and created an initial set of 48 technologies.

In the *demonstration and evaluation phase*, following five expert interviews and three workshops with technology experts, the capabilities of the initial set of technologies were demonstrated in illustrative data sharing use cases. The experts evaluated the technologies based on their potential to fulfill the defined requirements, as well as on their limitations. Furthermore, they were asked to elaborate on the current maturity of the technologies. The maturity levels were specified according to the methodology for PETs maturity assessment developed from the European Union Agency for Network and Information Security (ENISA) which define PETs in a broad sense considering "all kinds of technologies […] that [protect data] privacy" (Hansen et al., 2015, p. 11). Accordingly, we differentiated between four maturity levels (from low to high): technology at research level; technology at proof-of-concept level; technology at pilot level; technology at product level (Hansen et al., 2015). As suggested by Hansen et al. (2015), to make the expert assessment more robust, a measurable indicator for the assessment was defined; this was the average readiness level of a technology for different use cases within the context of data sharing. This indicator was deemed suitable as our study focused on a broad application of trust-enhancing technologies in a data ecosystem.

In the second *design & development phase*, we created a set of technologies considered most promising, with the goal of allowing for deeper analysis in order to integrate them into the data sharing process. To be able to select the most promising technologies, we first assessed each technology's potential to meet one or more requirements, as well as its limitations and its current maturity level. Then, starting with the most promising one, we successively added technologies to create the set; each time we added a technology, we assessed whether the set of selected technologies collectively fulfilled the set of requirements for building system trust. When this was achieved, we stopped adding further technologies, so as to limit the number of selected technologies. The selection was based on insights from the evaluation phase of the first design cycle and supplemented by three design workshops with experts.

In the second *demonstration & evaluation phase*, the specifications of this technology selection and its potentials and limitations for enhancing system trust were finally demonstrated and evaluated. As suggested by Peffers et al. (2012), we used two illustrative data sharing scenarios (direct data sharing between suppliers and manufacturers; a central data marketplace) in three expert workshops to illustrate the suitability of our artifact.

4 Result

4.1 Trust Challenges and Requirements for Data Sharing

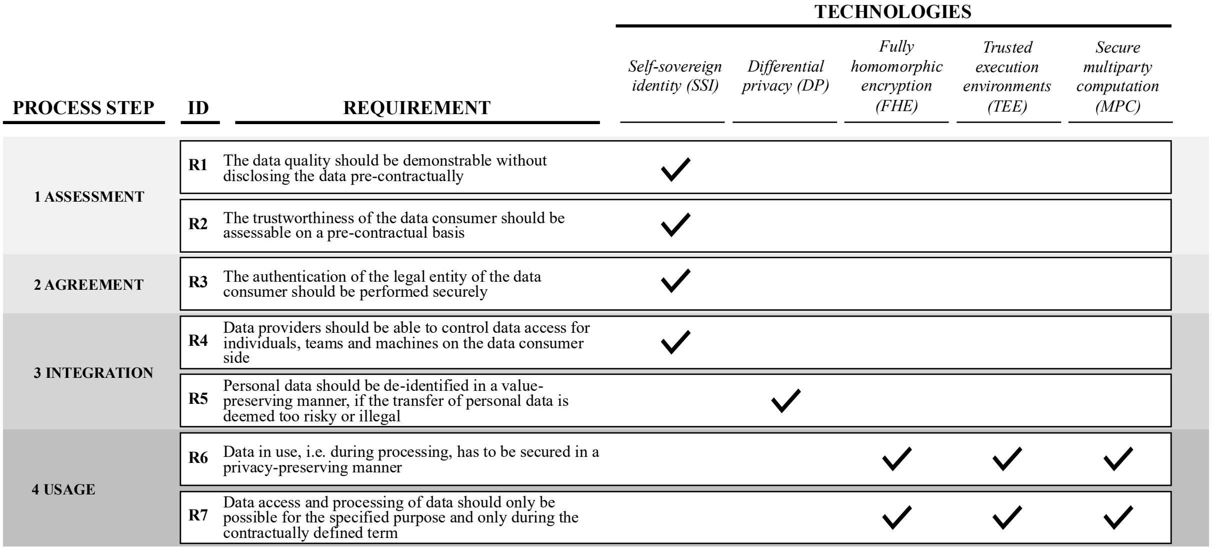
In this section, the relevant data sharing process phases, trust challenges and corresponding requirements will be explained. The trust challenges were structured along a four-phase data sharing process derived from the data sharing process model literature (see Section 2.1), which the interviewed experts also agreed on: assessment, agreement, integration, usage. These four phases comprise all activities of a data sharing process where a data provider faces trust challenges in regard to data sharing. In the *assessment phase*, the data provider and the data consumer are matched. The data consumer needs to identify a suitable data set for its particular use case. Therefore, an assessment of the quality of data sets offered by different data providers is required (Azkan et al., 2020; Krasikov et al., 2022; Oliveira et al., 2019). In this phase, data providers are usually not willing to share any data (Song et al., 2021), therefore, it is critical that the data consumer can assess the data sets without the data providers having to expose any data (R1). Moreover, to enter the agreement phase, the data provider should be able to evaluate if the data consumer is trustworthy (R2).

In the *agreement phase*, the data sharing agreement is negotiated and signed (Dalmolen et al., 2019; Krasikov et al., 2022). In this phase, it is extremely important that the data provider defines the purpose for which the data can be used by the data consumer. The contract may also include specific restrictions and obligations regarding data use, access, and deletion, as well as payment details (Dalmolen et al., 2019; Jussen et al., 2023). A key trust issue for the data provider at this point is the possibility of identity fraud by the data consumer, with or without fraudulent authentication through an unreliable third-party identity management system (Garrido et al., 2022). With growing ecosystems in a networked economy where data is increasingly shared with previously unknown parties (Jiang et al., 2021), this type of threat is growing. To prevent exposure of data to illegitimate consumers, the data provider and consumer require trustworthy identity authentication (R3), which assures that the contract is signed by two legitimate entities, also ensuring legal coverage.

In the *integration phase*, data usage by the data consumer is facilitated by transferring the data to the system in which data processing takes places (Dalmolen et al., 2019; Krasikov et al., 2022). During this step, the data provider has to address two challenges. Firstly, the possibility of unauthorized individuals, teams or machines associated with the data consumer gaining access to the transferred data. Secondly, data consumers may not apply the same privacy standards as the provider with respect to transferred personal data (Garrido et al., 2022). Any violation of privacy standards, regulations and norms related to personal data on the part of the data consumer may result in the provider being subject to legal proceedings and may result in damage to the reputation of the data provider (Culnan, 2019; Weiss, 2018). To address the former of the two challenges, access control on the data consumer side is necessary (R4). To address the latter, should the data provider deem the transfer of personal data too risky or illegal, it must be possible to ensure that data subjects are unidentifiable in the transferred data, while retaining as much of the value of the data as possible (R5).

In the *usage phase*, the shared data is processed to fulfill the purpose of data processing (Krasikov et al., 2022). Compared to securing data at rest or in transit, securing data in use, i.e., during computation in a privacy-preserving manner is highly complex, but required for a trusted exchange of data (R6) (Jansen, 2011). Finally, when the data processing purpose defined in the data sharing agreement has been fulfilled, data access or processing should be terminated (R7). If personal data has been processed and the data is no longer necessary for the original purpose, (in most cases) data deletion is required to comply with data protection laws such as the General Data Protection Regulation (GDPR).

The following sections will focus on the trust-enhancing potential, limitations and current maturity level of the five most promising trust-enhancing technologies that are collectively able to address the aforementioned requirements: self-sovereign identity (SSI), differential privacy (DP), fully homomorphic encryption (FHE), trusted execution environment (TEE) and secure multiparty computation (MPC) (cf. Figure 3). The experts mentioned SSI and DP for the fulfillment of specific requirements (R1–R5) and discussed them mainly in isolation. In contrast, FHE, TEE and MPC were selected for requirements R6 and R7 and were often juxtaposed by the experts. Section 4.4–4.6 will reflect this, allowing for an informed choice of these technologies in specific data sharing scenarios.



			TECHNOLOGIES				
PROCESS STEP	ID	REQUIREMENT	*Self-sovereign identity (SSI)*	*Differential privacy (DP)*	*Fully homomorphic encryption (FHE)*	*Trusted execution environments (TEE)*	*Secure multiparty computation (MPC)*
1 ASSESSMENT	R1	The data quality should be demonstrable without disclosing the data pre-contractually	✓				
	R2	The trustworthiness of the data consumer should be assessable on a pre-contractual basis	✓				
2 AGREEMENT	R3	The authentication of the legal entity of the data consumer should be performed securely	✓				
3 INTEGRATION	R4	Data providers should be able to control data access for individuals, teams and machines on the data consumer side	✓				
	R5	Personal data should be de-identified in a value-preserving manner, if the transfer of personal data is deemed too risky or illegal		✓			
4 USAGE	R6	Data in use, i.e. during processing, has to be secured in a privacy-preserving manner			✓	✓	✓
	R7	Data access and processing of data should only be possible for the specified purpose and only during the contractually defined term			✓	✓	✓

Figure 3. Requirement fulfillment potential of the technologies along the data sharing process.

4.2 Self Sovereign Identity (SSI)

Self sovereign identities (SSIs) enable individuals or organizations to decentrally manage their own digital identities. These identities are stored as decentralized identifiers (DIDs) in a decentralized registry and are cryptographically verifiable (Mühle et al., 2018). Therefore, SSI users do not have to rely on a third-party that centrally hosts and controls identity data. SSIs also offer the capability to issue *verifiable credentials* of users (Wang and De Filippi, 2020). These are verifiable digital proofs linked to identity records issued by a trusted source, such as the state or any other trusted institutions. These verifiable claims are managed by the user and can be shared for a wide range of application scenarios (Mühle et al., 2018). In the context of data sharing, SSIs can be used to enhance trust in three data sharing process phases: assessment, agreement and integration.

In the assessment phase, according to an expert (P17), SSIs can provide a solution for the requirements R1 and R2 by establishing a rating system through verifiable credentials for data providers and data consumers in a data ecosystem. A trusted rating system based on past data sharing transactions would both attest to the quality of data providers and endorse the overall trustworthiness of data consumers. The use of verifiable credentials in this context yields the benefit of offering a rating that is instantly verifiable, tamper-proof through decentralized architecture and even privacy-preserving through its selective disclosure approach (Wang and De Filippi, 2020). *“In Amazon reviews, for example, you don't* *know if [...] it is a fake review used to promote the product. If you were to implement an SSI based rating you could cryptographically guarantee that only parties who have purchased the product can rate it and the consumer identity would not have to be disclosed"* (P17).

In the agreement phase, the issue of identity authentication (R3) can be solved using DIDs in the SSI concept (P07, P13, P14, P17, P18). Compared to centralized authentication points, SSIs, due to their distributed nature, have the advantage that they do not have to involve third parties (P13). The decentralized structure therefore prevents potential malicious attacks on a centralized critical element of the ecosystem, making identity authentication more trustworthy. According to one expert (P14) a data ecosystem is only as trustworthy as its least trusted component. Therefore, such centralized vulnerabilities should be avoided.

In the integration phase, SSIs can address the requirement (R4) to have cross-company access control mechanism (P07, P14, P17). They can build a granular, access management system on the data consumer side for individuals and machines, whose connections can be represented in parent-child relationships (P17). Due to the inter-organizational standardization of identities through SSIs, the data provider is able to configure the access rights for identities on the data consumer side (e.g., person, teams, machines) (P14). SSIs also have the potential to be combined with *sticky policies* by "sticking" authorized subjects' DIDs to the transferred data in order to carry the access rights to the data consumer (P14). Sticky policies are machine-readable usage policies attached to data (Pearson and Cassasa-Mont, 2011). Their execution in a *trusted execution environment* (compare Section 4.5) could potentially shift the enforceability of these policies from the legal to the technological realm, resulting in enormous trust-enhancing potentials. Why SSIs are suitable to implement access rights through sticky policies was described as follows: *"The usage policy must be carried on into the external systems. To achieve this, I anticipate technologies that come from the SSI environment, as this will require a cross-company approach, which means that identity silos will no longer make sense"* (P14).

Limitations: According to the experts, the implementation of SSIs remains less advanced for two reasons: technological maturity and lack of network effects. First, many of the SSI technologies are not technologically mature enough to reach a critical mass of users for productive scaling (P08, P17, P18, P19, P20). The introduction of a technologically enforceable access management system in conjunction with sticky policies and trusted execution environments considerably increases technological complexity (P14). According to an expert, such an implementation is still a long way off: *"The technological implementation of SSIs in combination with sticky policies, which carries the data usage rights through to the target systems, is likely to take several more years."* (P14). Secondly, the lack of standardization prevents network effects (P07, P08, P17, P18, P20). Currently, there are different SSI networks based on different standards and these are not interoperable with each other, which severely limits their useability (P12, P17). SSIs only work through network effects because all relevant actors in the data ecosystem must be onboard, otherwise the operational risks outweigh the trust-enhancing potential (P17). *"The use of SSIs is a strategic decision. Today, the risk that you have a supplier who is not connected [...] outweighs the potential for managers"* (P17).

Maturity (Pilot): The experts agreed that the technology is currently at pilot level (P17, P20). One example that was specifically mentioned is the Global Legal Entity Identifier Foundation (GLEIF), which offers automated authentication and verification of legal entities based on SSI technology. GLEIF operates as the root of trust in the issuance of digital trustworthy legal identifier codes (GLEIF, 2022).

4.3 Differential Privacy (DP)

To achieve differential privacy, algorithms add random noise to a data analysis such that its results do not change significantly when a single data subject is included in or excluded from the input dataset. This provides data subjects plausible deniability, thus anonymizing the data set (Schmidt et al., 2022). DP provides mathematical guarantees of data privacy and prevents the risk of re-identification of the data by reverse-engineering the outputs (Zhao and Chen, 2022). It also allows accurate quantification of privacy loss associated with making noise-injected analysis results available, thus allowing informed trade-offs between privacy and utility (Dwork et al., 2019).

In the integration phase, DP can fulfill requirement R5 (P07, P11, P14, P19). The data is modified by DP methods before sharing it with the processing system so that the presence of a specific individual's data in the dataset cannot be inferred (P07, P19). Thus, DP allows anonymized sharing of previously personal data. *"With data modification, large datasets can be anonymized and it can be ensured that no sensitive data can be extracted"* (P19).

Limitations: DP is already in use today, but according to the experts faces three main limitations. Firstly, the involved actors themselves must have a deep understanding of the technology and must actively participate in its execution (P07, P19). *"You have to actively deal with it in practice and can't just say 'I have a huge database here, I want to sell it now [...] and not worry about it anymore. You really have to sit down with the other party and see what kind of information they want'"* (P19). Secondly, DP's mathematical ability to accurately quantify privacy loss results in a governance challenge in data sharing. According to a data analytics expert, *"Another critical point with DP is that defining your epsilon value is a political determination. You have to come to an internal group consensus about the privacy risk that you're willing to absorb"* (P11). Thirdly, the same expert believes that DP usability is limited in data sharing as it is only applicable for statistical aggregation use cases. Hence, if a use case requires the linkage of data to other information, for example in predictive maintenance where the input data is linked to a specific component, it is not possible via DP. *"[It] has limited applications for aggregate data analytics. When you need linkable data – many applications require linkable data – DP is not gonna do it for you! It's good for statistical aggregations. It's not good outside of that context"* (P11).

Maturity (Product): The experts agreed that DP is already mature and mentioned users like Microsoft and Apple, ranking it at product level (P07, P11). For example, Microsoft uses DP when collecting telemetry data from windows devices (Microsoft, 2017). In B2B data sharing, for example, the US Census Bureau uses DP to anonymize their databases to enable external access (Census, 2022).

4.4 Fully Homomorphic Encryption (FHE)

Fully Homomorphic Encryption allows operations on encrypted data without ever decrypting the data, thus providing a guarantee of privacy (Aslett et al., 2015; Morris, 2013; Will and Ko, 2015). The private key to decrypt the output is held only by its defined owner and the results match those that would be achieved were the operation performed on decrypted data (Gentry, 2009).

In the usage phase, according to experts (P07, P11, P14, P19, P21), FHE can meet the requirements for privacy-preserving and secure processing of data (R6) as well as purpose limitation (R7). In the context of data sharing, FHE can be applied in different ways. For example, it can be used for statistical analyses without disclosing any input data (Aslett et al., 2015). Thereby, the data provider would encrypt their input data with a public key and the data consumer would homomorphically execute their analysis in a cloud. At no point would the cloud provider or data consumer have access to unencrypted input data, guaranteeing computational privacy and security and limiting the use of the data to its purpose. One expert summarized the theoretical trust-enhancing potential of FHE saying, *"FHE would be the holy grail"* (P19). In practice, according to one workshop participant, FHE is highly promising for some specific use cases and thereby may already have advantages over other trust-enhancing technologies: *"If there are few data sets and the case is very specific, I would go with FHE because I don't need a lot of infrastructure. [...] If you know that the use case is supported by homomorphic and the libraries are there, you don't need much more than that, [making] FHE cheaper to implement than MPC. [...] If you're going to do it [with MPC] just for that case, you're using a shotgun to kill a fly"* (P21).

Limitations: The use of FHE today has several operational limitations. A scalable, productive use is currently impossible due to the computational effort required and complex implementation (P07, P11, P19). As the efficiency of FHE continues to increase and computational power becomes cheaper, the computational effort required represents a limit to be overcome in the future (P21). Still, the custom-built FHE implementations are highly complex and usually bound to a single use case to avoid additional complexity, hindering its scalability and connectivity to other systems and requiring deep know-how: *"[To enable multiple use cases] you need to know all the edge cases. [...] Otherwise, you have one use* *case, one dataset in one specific place, and one person who knows how to run the one operation that you support. So, either you're monolithic or [...] it's going to be a mess of governance, coordination, pipeline, et cetera"* (P21). The practical technological limitation also severely limits the possible applications of FHE: *"Imagine how many use cases do you have where you only need to multiply two numbers. I would say it's 0.01% of the enterprise cases. [...] The usability is still quite poor, compared to what you would do with secure multiparty computation or trusted execution environments"* (P21).

Due to the outlined limitations, practical applications for FHE are limited to specific, sensitive use cases (e.g., healthcare (Munjal and Bathia, 2022)): *"For a while we saw FHE being touted as the solution to all privacy challenges in the world. It has some utility but may be limited to a specific context"* (P11).

Maturity (Research): For specific, highly sensitive use cases, FHE is already in use today. For example, one expert (P07) mentioned that Microsoft employs homomorphic encryption (HE) in its Edge browser to operate its password library (Microsoft, 2021). Microsoft is also offering a cloud-based HE library, called SEAL, developed and maintained by cryptography researchers (Wood et al., 2020). First applications of SEAL are TenSEAL, applying FHE in private deep learning (Github, 2022) or Intel's HE Transformer for its future neural networks (Wood et al., 2020). However, for widespread application FHE is currently at research level to overcome the aforementioned limitations (P07, P11, P19).

4.5 Trusted Execution Environment (TEE)

Trusted execution environments (TEEs), also called confidential computing in a cloud context, protect data in use (Mulligan et al., 2021). TEEs are secure hardware areas within the processor, physically isolated from the applications and device operating system (Sabt et al., 2015). Consequently, even if the host system is compromised, the data in a TEE cannot be accessed (P06, P19). TEEs provide confidentiality and, in most cases, integrity (Sabt et al., 2015). So-called remote attestations provide cryptographic proof that the user interacts with a genuine TEE on a remote system and that the data and programs inside the TEE have not been tampered with (Hynes et al., 2018; Ménétrey, 2022; Sabt et al., 2015). Additionally, a secure communication channel between the input system and the TEE can be built by remote attestation, by establishing a secure hardware associated encryption key (Pereira Pires, 2019). This channel allows the remote unit to provide the TEE with encrypted secrets (e.g., training data, private keys) (Wagner et al., 2020; Gueron, 2016).

In the usage phase, TEEs, similar to FHE, can fulfill the requirement for a secure and privacy-preserving processing of data (R6) as well as technological enforceability of the purpose limitation (R7) (P14, P13, P19, P20). TEEs can technologically ensure that no decrypted data can leak out during processing since decryption can only be performed within the secure TEE and hardware-based mechanisms prevent an attacker from accessing the TEE (Mulligan et al., 2021). Furthermore, using remote attestation, the TEE demonstrates both its own authenticity and that the processing of programs within the environment are in line with the agreement (P14, P21). The data usage can therefore be carried out in a secure, privacy-preserving environment (Hynes et al., 2018) as the data consumer and TEE provider cannot access the unencrypted data at any point (P14). One expert explained: *"You can provide trust within a function in a program running in that trusted environment and ensure that other programs cannot read, access, or modify that function"* (P06). Another expert added, *"[They] follow the principle of creating a non-compromisable area on a machine that may be compromised"* (P19).

In contrast to FHE and MPC, the use of TEEs has the advantage that the computations are performed on only one machine, making the use of TEEs more efficient and less costly (P06, P16, P19). "*That's great, you can [solve] a lot of problems very efficiently, because you still compute efficiently on a single machine"* (P19). The ease of implementation is also an argument for the use of TEEs as it does not require the establishment of an infrastructure, as with MPC, or massive know-how (P06, P13, P19). The TEE options offered by cloud providers also serve to simplify their use compared to other technologies, as many companies already have their data stored in a cloud (P07). One expert, a privacy tech startup founder, commented on this, saying, *"Of my technology budget, I would invest 50% in TEE"* (P16).

Limitations: The trust-enhancing potential of TEEs is contrasted with three limitations: an operational, a technological, and a trust limitation. The operational limitation is because the data must not leave the

TEE. This means that all confidential calculations must be performed using code in the TEE, which in turn limits the available use cases as it complicates integration with existing systems (e.g., ERP systems) (P14). Enabling existing systems to be TEE-compatible is sometimes difficult or even impossible (P07). This limitation also applies in part to FHE, MPC and DP, and according to one expert could cause difficulties in the future: *"It's going to be a bigger problem as time passes by, because especially now, we are moving from [...] discrete machine learning to a more streaming-based learning"* (P21).

Secondly, computation in the TEE in one single machine has implications for the scalability and performance of calculations due to the limited available memory (P14, P19). When asked if big data analyses are possible in TEEs today, one expert answered: *"I would bet [the size of datasets in TEEs] usually is not larger than a handful of gigabytes. I wouldn't expect a TEE with terabytes of data – definitely not"* (P21). However, major infrastructure providers are investing a lot in TEEs, making them capable of handling larger applications in the near future (P04). This is also underlined by one expert: *"The wide application of TEEs is definitely closer to practice than homomorphic encryption"* (P19).

The third limitation of the technology is the centralization of trust to the chip manufacturer. For the technology to be of value, the data provider and consumer must trust the hardware and the chip manufacturer's remote attestation (P06, P13, P19). Yet, there have been security breaches with these chips previously (P14, P19). This residual risk is thus a trade-off for simple and efficient implementation. In reality, one expert (P20) argues companies' trust in TEEs is high enough that they enable the vast majority of use cases. Only for ultra-sensitive use cases (e.g., health care) would technologies with mathematical privacy guarantees (e.g., MPC or FHE) have to be used. *"You need to trust the manufacturer. So Intel or AMD or someone who provides this TEE"* (P06).

Maturity (Product): The experts consider TEEs as relatively advanced and therefore at early product level (P04, P07, P19, P20). According to P20, they are definitely more mature than MPC and FHE. All the large cloud providers, Google, AWS and Microsoft offer confidential computing and invest in it (AWS, 2022; Google, 2022; Microsoft, 2022). Other tools aim to enable TEE-compatibility of existing systems that are not themselves attached to TEEs (e.g., Anjuna (2022)). Based on these services, the first data spaces are currently being developed, seeking to enable secure data exchange (P20).

4.6 Secure Multiparty Computation (MPC)

Secure multiparty computation (MPC) is a cryprographic technique that allows ecosystem actors to jointly compute functions without disclosing their input data to each other (Choi and Butler, 2019). MPC systems are mostly based on secret sharing protocols, through which each data owner splits its input data into encrypted parts called secret shares (Agahari et al., 2022). Shares are then sent to the other MPC parties to perform arithmetic computations in a distributed manner (Garrido et al., 2022). In the end, all parties' sub-results are combined, and the final output is made available to the parties or only one designated data consumer (Damgård et al., 2016). Thus, the input data can be processed privacy-preservingly (Agahari et al., 2022).

In the usage phase, MPC can fulfill the requirement for privacy-preserving and secure processing of data (R6) as well as for purpose limitation (R7) (P06, P14, P17, P19). The MPC performs only one agreed-upon, specific calculation at a time. Depending on the utilized MPC protocol, the data consumer is the only one able to retrieve the final result and, hence, exercises sovereignty over its use. The data consumer can, however, not only act as consumer but also as data provider (e.g., by inputting its machine learning model into the computation). The MPC parties as well as any other party would not have access to the input data at any time, therefore, a privacy-preserving, secure processing of the data can be guaranteed (P19). *"With MPC we can execute arbitrary functions on secret data. So, one learns nothing from the secret data except what one would trivially learn from the result"* (P19). In contrast to TEEs and similarly to FHE, according to an expert, the mathematical privacy guarantee of MPC systems raises the trust level of secure and privacy-preserving processing to a level required for sensitive use cases. Whereas, in case of TEEs, this would not be possible due to their residual risk: "*TEEs only allow a limited trust level in the ecosystem in which they exist due to their limitations. The advantage of MPC is that we can mathematically assure that the calculations are correct and executed securely"* (P19).

Limitations: Currently, MPC has limitations in its efficiency, the required infrastructure, and governance. Firstly, while being computationally less complex than FHE, MPC still has considerably higher communication and processing costs than TEEs (P06, P07, P14, P19). MPC today still struggles with performance and scalability issues. Nevertheless, advances in computational power and efficiency have brought the technology close to practical application in recent years (P19). A trusted computation expert described the performance disadvantage compared to TEEs as follows: *"The disadvantage of MPCs is, of course, performance. [...] we need a cluster of participants. [...] And then they don't just calculate locally, but a protocol is executed. This means that during the calculation there is interaction between the parties executing the calculation and this is significantly slower and more complicated than if we would simply send it to a TEE, which then does everything and sends the result back"* (P19).

Secondly, similarly to FHE, MPC also has protocol-dependent limitations in terms of the computations that can be performed (P06, P07, P14, P19). On the one hand, as stated by one workshop participant, the supported operations are much wider than in current FHE: "*I've seen [MPC] systems that allow you to run pretty much everything you want."* (P21). On the other hand, MPC protocols are often designed and optimized for specific use cases rather than general purpose calculations. One interviewee compared the limitations of MPC and FHE as follows: *"The difference is that if you do MPC, you know that at the end of that [complex implementation], you may have a bit more freedom to run computation on the data. [...] You may end up supporting cases that you didn't think of at the beginning, whereas homomorphic encryption [...] restrict you on the operations."* (P21). Thirdly, an ecosystem infrastructure must be established for MPC use (P20, P21). One workshop participant stated (P21) that all data providers and data consumers involved must be brought onboard with this infrastructure, which requires a standardization effort. Furthermore, such construction of MPC infrastructure is associated with high adoption costs due to its complexity and the required development of know-how: "*It's just not feasible for standard enterprise engineers to manage those things"* (P21). Another expert summarized the current use of the MPC thus: *"Theoretically, a lot is possible, but practically, MPC is used today to chop up data and create sums from the separate calculations" (P14).*

Maturity (Proof-of-concept): According to the experts, the use cases will be potentially numerous in the future, but are relatively few today due to the described limitations (P14, P17). Nevertheless, the technology currently has a lot of traction and is making great progress (P17, P19). *"In reality, MPC is currently making the leap from theory to practice"* (P19). Today, MPC is used in specific use cases, but not for widespread data sharing applications, ranking at proof-of-concept level (P14, P17, P19). For example, Cybernetica is using MPC in its Sharemind MPC application to allow users to securely process data (Cybernetica, 2022). Another data exchange initiative that was mentioned by an interviewee is the Carbyne Stack. This is an open-source cloud stack for building scalable MPC applications, promising to allow secure collaboration on private data with any number of clients (Carbyne Stack, 2022).

5 Discussion and Conclusion

Previous studies have indicated that technology has the potential to enhance digital trust and shed light on the trust-enhancing technologies landscape. However, there is currently little knowledge about the trust-enhancing potential of digital technologies in the data sharing process of ecosystems and no systematic design for their integration into these. This paper aims at closing this gap by taking a DSR approach. Driven by clearly defined criteria, our technology selection allows us to show how a combination of technologies (SSI, DP, FHE, TEE and MPC) can enhance digital trust across the data sharing process and that these technologies have wide-ranging capabilities which are partly interchangeable (e.g., FHC, TEE, MPC) in specific phases (e.g., usage phase). However, we also identify three main limitations that continue to limit the full potential of the technologies today: standardization (SSI, MPC), computational performance (FHE, TEE, MPC) and ease of implementation (DP, FHE, MPC). Considering the highly dynamic nature of IT artifacts in part due to innovation and changing regulations (Orlikowski and Iacono, 2001), we assume that the five technologies identified will mature quickly and that the limitations will start to decrease. At the same time however, technologies beyond the scope of our study are also expected to gain relevance. Therefore, our collection of identified technologies needs to be adjustable in the future. Hence, based on our approach, we have identified and suggest two principles for technology integration that guide the recombination and reintegration of trust-enhancing technologies: (1) *complementarity* and (2) *customization*.

Complementarity: Within the five most promising trust-enhancing technologies that we identified, none had the potential to fulfill all relevant trust requirements (R1–R7) comprehensively (cf. Figure 3). If one of the requirements is neglected or insufficiently fulfilled, it becomes the weakest link in the data sharing process and decreases the maximum level of trust. Accordingly, a complementary collection of technologies is required to mutually fulfill the relevant trust requirements (R1–R7).

Customization: As our analysis (cf. Figure 3) shows, three technologies were able to satisfy the requirements R6 and R7. These technologies are in a ratio of substitution. The final choice of technology ultimately depends on the use case. Whether FHE, TEE or MPC are to be selected for a secure calculation depends on the required efficiency for the calculation, on the sensitivity of the input data or on the multifariousness of the use case.

With our research results, we make a contribution on both a theoretical and a practical level. On the theoretical level, the novel insights from our explorative research approach on the application of trust-enhancing technologies for data sharing add to the existing literature on the role of digital technologies for trust building (Agahari et al., 2022; Lumineau et al., 2023; Mubarak and Petraite, 2020). Up until now it has been unclear what trust-enhancing potential digital technologies are capable of yielding in the data sharing process and how they can be integrated into this process within data ecosystems. Moreover, our study addresses the dimensions of data privacy, cybersecurity and data sovereignty, thus contributing to the emerging literature stream on digital trust (Sahut et al., 2022). With our study, we shed light on the topic from a technological point of view and identify which technologies require further research to reduce their limitations.

With our findings, we also aim to contribute to the application of trust-enhancing technologies in practice; investment in these technologies has doubled within the last three years (Chui et al., 2022). Moreover, studies have shown that trust is a strategic factor for the success of companies, especially in the context of data ecosystems (Jiang et al., 2021; Boehm et al., 2022; Kluiters et al., 2023). This contribution should support executives in their strategic investment decisions. Recent studies indicate that companies must be able to evaluate the suitability of such technologies in advance (Mubarak and Petraite, 2020; Sahut et al., 2022). As we propose a collection of five technologies based on a profound evaluation, executives can steer their investment portfolio accordingly.

Due to its explorative nature, our study does not come without limitations, which themselves open new research avenues. Firstly, as our study is a conceptual study and no larger scale demonstrations or evaluations have been conducted as yet, an experimental study could check whether the trust enhancement is sufficient for data providers to share their data in data ecosystems. If such a study were to disprove this hypothesis, challenges regarding further ecosystem actors (e.g., consumers) may lead to additional requirements to be satisfied. Secondly, due to our sampling approach, our results might be industry limited and not generalizable to other world regions. Third, the experts provided indications for potential operational limitations of trust-enhancing technologies. To address these limitations, our systemic perspective (see Figure 1) needs to be broadened by incorporating data providers and consumers as sub-systems within data ecosystems. In this vein, follow-up studies may not only incorporate trust requirements for integration design, but also technical requirements that may stem, for instance, from challenges related to the integration of data from legacy systems (e.g., ERP systems) placed within a sub-system (i.e., data provider and consumers). We anticipate future design, demonstration and evaluation of our artifact in existing legacy systems to obtain operational context specification. Beyond additional technical considerations, future research should also focus on governance mechanisms (Oliveira et al., 2019).

The artifact that we have designed in this article is not meant to offer a unique solution to enhance trust for data sharing in data ecosystems. Nevertheless, we believe that the insights from our article regarding the integration of trust-enhancing technologies into the data sharing process in data ecosystems are helpful for both academics and practitioners to unlock the potential of data ecosystems.

References

Abraham, C., Sims, R. R., Daultrey, S., Buff, A. and Fealey, A. (2019). “How digital trust drives culture change,” *MIT Sloan Management Review* 60 (3), 1-8.

Agahari, W., Ofe, H. and De Reuver, M. (2022). “It is not (only) about privacy: How multi-party computation redefines control, trust, and risk in data sharing,” *Electronic Markets* 32, 1577-1602.

Aguiar, M., Pidun, U., Lacanna, S., Knust, N., Williams, M. and Candelon, F. (2021). *Discovering the Tools and Tactics of Trust in Business Ecosystems*. BCG Henderson Institute. URL: https://web-assets.bcg.com/69/a7/91fc1e9544068e923e398560dfee/bcg-discovering-the-tools-and-tactics-of-trust-in-business-ecosystems-jun-2021.pdf.

Ali, M., Dhamotharan, R., Khan, E., Khan, S. U., Vasilakos, A. V., Li, K. and Zomaya, A. Y. (2015). “SeDaSC: Secure data sharing in clouds,” *IEEE Systems Journal* 11 (2), 395-404.

Anjuna (2022). *Protect Data With Anjuna Confidential Computing Software*. URL: https://www.anjuna.io/protect-data-with-anjuna-confidential-cloud-software (visited on November 16, 2022).

Aslett, L. J., Esperança, P. M. and Holmes, C. C. (2015). *A review of homomorphic encryption and software tools for encrypted statistical machine learning*. Department of Statistics, University of Oxford. URL: https://arxiv.org/abs/1508.06574.

Aws (2022). *AWS Nitro Enclaves*. URL: https://aws.amazon.com/de/ec2/nitro/nitro-enclaves/ (visited on November 16, 2022).

Azkan, C., Möller, F., Meisel, L. and Otto, B. (2020). “Service dominant Logic Perspective on Data Ecosystems-a Case Study based Morphology,” in: *Proceedings of the 28th European Conference on Information Systems (ECIS)*, Marrakech, Morocco.

Bachmann, R. (2003). “Trust and power as means of coordinating the internal relations of the organization: A conceptual framework,” in: Nooteboom, B. & Six, F. (eds.) *The trust process in organizations: Empirical studies of the determinants and the process of trust development.* Cheltenham, UK: Edward Elgar Publishing.

Badewitz, W., Kloker, S. and Weinhardt, C. (2020). „The Data Provision Game: Researching Revenue Sharing in Collaborative Data Networks,“ *2020 IEEE 22nd Conference on Business Informatics (CBI)*, Antwerp, Belgium.

Ballatore, M., Toumi, D. and Arena, L. (2022). „Blockchain-Based Data Sharing System: An Experimental Analysis of Behavioural Features Affecting Inter-Organisational Cooperation,“ *European Conference on Information Systems (ECIS)*, Timisoara, Romania.

Benbasat, I. and Zmud, R. W. (2003). “The identity crisis within the IS discipline: Defining and communicating the discipline's core properties,” *MIS quarterly* 27 (2), 183-194.

Beverungen, D., Hess, T., Köster, A. and Lehrer, C. (2022). “From private digital platforms to public data spaces: implications for the digital transformation,” *Electronic Markets* 32, 493-501.

Bitkom (2022). *Datenschutz in der deutschen Wirtschaft: DS-GVO & internationale Datentransfers*. Bitkom Research. URL: https://www.bitkom.org/sites/main/files/2022-09/Bitkom-Charts%20Datenschutz%2027%2009%202022_final.pdf.

Boehm, J., Grennan, L., Singla, A. and Smaje, K. (2022). *Why digital trust truly matters*. McKinsey Digital. URL: https://www.mckinsey.com/capabilities/quantumblack/our-insights/why-digital-trust-truly-matters.

Braud, A., Fromentoux, G., Radier, B. and Le Grand, O. (2021). “The road to European digital sovereignty with Gaia-X and IDSA,” *IEEE Network* 35 (2), 4-5.

Carbyne Stack (2022). *Award-Winning Cloud Native Secure Multiparty Computation*. URL: https://carbynestack.io (visited on November 16, 2022).

Casadesus-Masanell, R. and Hervas-Drane, A. (2020). “Strategies for managing the privacy landscape,” *Long Range Planning* 53 (4), 1-11.

Census (2022). *Understanding Differential Privacy*. URL: https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/differential-privacy.html (visited on November 16, 2022).

Chen, Y., Kreulen, J., Campbell, M. and Abrams, C. (2011). “Analytics ecosystem transformation: A force for business model innovation,” in: *Proceedings 2011 Annual SRII Global Conference*, San Jose, California USA.

Choi, J. I. and Butler, K. R. B. (2019). “Secure Multiparty Computation and Trusted Hardware: Examining Adoption Challenges and Opportunities,” *Security and Communication Networks* 1, 1-28.

Chui, M., Roberts, R. and Yee, L. (2022). *McKinsey Technology Trends Outlook 2022*. McKinsey & Company. URL: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-top-trends-in-tech.

Culnan, M. J. (2019). “Policy to avoid a privacy disaster,” *Journal of the Association for Information Systems* 20 (6), 848-856.

Curry, E. and Sheth, A. (2018). “Next-generation smart environments: From system of systems to data ecosystems,” *IEEE Intelligent Systems* 33 (3), 69-76.

Cybernetica (2022). *Sharemind MPC (Multi-Party Computation)*. URL: https://cyber.ee/products/sharemind-mpc (visited on November 16, 2022).

Dalmolen, S., Bastiaansen, H., Kollenstart, M. and Punter, M. (2019). “Infrastructural sovereignty over agreement and transaction data ('metadata') in an open network-model for multilateral sharing of sensitive data,” in: *Proceedings of the 40th International Conference on Information Systems (ICIS)*, Munich.

Damgård, I., Damgård, K., Nielsen, K., Nordholt, P. S. and Toft, T. (2016). „Confidential benchmarking based on multiparty computation,“ *International Conference on Financial Cryptography and Data Security*, Christ Church, Barbados.

Dwork, C., Kohli, N. and Mulligan, D. (2019). “Differential privacy in practice: Expose your epsilons!,” *Journal of Privacy and Confidentiality* 9 (2), 1-22.

European Commission (2020). *A European strategy for data*. Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN.

Fadler, M. and Legner, C. (2022). “Data ownership revisited: Clarifying data accountabilities in times of big data and analytics,” *Journal of Business Analytics* 5 (1), 123-139.

Garrido, G. M., Sedlmeir, J., Uludağ, Ö., Alaoui, I. S., Luckow, A. and Matthes, F. (2022). “Revealing the landscape of privacy-enhancing technologies in the context of data markets for the IoT: A systematic literature review,” *Journal of Network and Computer Applications* 207, 1-29.

Gassmann, O. and Ferrandina, F. (2021). “Connected Business: Creating Value in the Networked Economy,” in: Gassmann, O. & Ferrandina, F. (eds.) *Connected Business.* Cham, Switzerland: Springer.

Gelhaar, J. and Otto, B. (2020). „Challenges in the Emergence of Data Ecosystems,“ *Twenty-Third Pacific Asia Conference on Information Systems*, Dubai, UAE.

Gentry, C. (2009). *A fully homomorphic encryption scheme.* PhD thesis, Stanford University.

Giddens, A. (1990). “The consequences of modernity,” *Cambridge: Polity*.

Github (2022). *TenSEAL: A library for doing homomorphic encryption operations on tensors*. URL: https://github.com/OpenMined/TenSEAL (visited on November 16, 2022).

GLEIF (2022). *GLEIF – Global Legal Entity Identifier Foundation*. URL: https://www.gleif.org/de/ (visited on November 16, 2022).

Goldberg, I., Wagner, D. and Brewer, E. (1997). “Privacy-enhancing technologies for the internet,” in: *Proceedings IEEE COMPCON 97. Digest of Papers*, San Jose, California, USA.

Google (2022). *Confidential Computing concepts*. URL: https://cloud.google.com/compute/confidential-vm/docs/about-cvm (visited on November 16, 2022).

Gueron, S. (2016). “A memory encryption engine suitable for general purpose processors,” *Cryptology ePrint Archive*, 1-14.

Guo, B., Deng, X., Guan, Q., Tian, J. and Zheng, X. (2018). “An incentive mechanism for cross-organization data sharing based on data competitiveness,” *IEEE Access* 6, 72836-72844.

Hansen, M., Hoepman, J.-H., Jensen, M. and Schiffner, S. (2015). *Readiness Analysis for the Adoption and Evolution of Privacy Enhancing Technologies: Methodology, Pilot Assessment, and Continuity Plan*. European Union Agency for Network and Information Security (ENISA). URL: https://www.enisa.europa.eu/publications/pets.

Hartmann, P. M., Zaki, M., Feldmann, N. and Neely, A. (2016). "Capturing value from big data–a taxonomy of data-driven business models used by start-up firms," *International Journal of Operations & Production Management* 36 (10), 1382-1406.

Hevner, A. R., March, S. T., Park, J. and Ram, S. (2004). "Design science in information systems research," *MIS quarterly* 28 (1), 75-105.

Huang, X., Xu, C., Wang, P. and Liu, H. (2018). "LNSC: A security model for electric vehicle and charging pile management based on blockchain ecosystem," *IEEE access* 6, 13565-13574.

Hynes, N., Dao, D., Yan, D., Cheng, R. and Song, D. (2018). "A demonstration of sterling: a privacy-preserving data marketplace," *Proceedings of the VLDB Endowment* 11 (12), 2086-2089.

ISO 27032 (2012). *Information technology — Security techniques — Guidelines for cybersecurity. (ISO/IEC Standard No. ISO/IEC 27032:2012)*. URL: https://www.iso.org/standard/44375.html.

Jansen, W. A. (2011). "Cloud hooks: Security and privacy issues in cloud computing," in: *Proceedings of the 2011 44th Hawaii International Conference on System Sciences*, Hawaii, USA.

Jarke, M., Otto, B. and Ram, S. (2019). "Data sovereignty and data space ecosystems," *Bus Inf Syst Eng* 61 (5), 549-550.

Jiang, Z., Thieullent, A.-L., Steve, J., Perhirin, V., Baerd, M.-C., Shagrithaya, P., Cecconi, G., Isaac-Dognin, L., et al. (2021). *Data sharing masters: How smart organizations use data ecosystems to gain an unbeatable competitive edge*. Capgemini Research Institute. URL: https://www.capgemini.com/wp-content/uploads/2021/09/Final-Web-Version-of-Report-Data-Ecosystems-1.pdf.

Joshi, A. and Wade, M. (2020). "The Building Blocks of an AI Strategy. The AI & Machine Learning Imperative," *MIT Sloan Management Review* 10, 7-9.

Jussen, I., Schweihoff, J., Dahms, V., Möller, F. and Otto, B. (2023). „Data Sharing Fundamentals: Definition and Characteristics,“ *56th Hawaii International Conference on System Sciences*, Maui, Hawaii, USA.

Kang, J., Yu, R., Huang, X., Wu, M., Maharjan, S., Xie, S. and Zhang, Y. (2018). "Blockchain for secure and efficient data sharing in vehicular edge computing and networks," *IEEE Internet of Things Journal* 6 (3), 4660-4670.

Kluiters, L., Srivastava, M. and Tyll, L. (2023). "The impact of digital trust on firm value and governance: an empirical investigation of US firms," *Society and Business Review* 18 (1), 71-103.

Koo, D., Hur, J. and Yoon, H. (2013). "Secure and efficient data retrieval over encrypted data using attribute-based encryption in cloud storage," *Computers & Electrical Engineering* 39 (1), 34-46.

Krasikov, P., Eurich, M. and Legner, C. (2022). „Unleashing the Potential of External Data: A DSR-based Approach to Data Sourcing,“ *European Conference on Information Systems (ECIS)*, Timisoara, Romania.

Liu, Q., Wang, G. and Wu, J. (2014). "Time-based proxy re-encryption scheme for secure data sharing in a cloud environment," *Information sciences* 258, 355-370.

Luhmann, N. (1979). *Trust and Power,* Chichester: John A. Wiley and Sons.

Lumineau, F., Schilke, O. and Wang, W. (2023). "Organizational Trust in the Age of the Fourth Industrial Revolution: Shifts in the Form, Production, and Targets of Trust," *Journal of Management Inquiry* 32 (1), 21-34.

Lumineau, F., Wang, W. and Schilke, O. (2021). "Blockchain governance—A new way of organizing collaborations?," *Organization Science* 32 (2), 500-521.

Mayer, R. C., Davis, J. H. and Schoorman, F. D. (1995). "An integrative model of organizational trust," *Academy of Management Review* 20 (3), 709-734.

Ménétrey, J., Göttel, C., Khurshid, A., Pasin, M., Felber, P., Schiavoni, V. and Raza, S. (2022). "Attestation Mechanisms for Trusted Execution Environments Demystified," in: *Proceedings of the 22nd IFIP International Conference on Distributed Applications and Interoperable Systems*, Lucca, Italy.

Microsoft (2017). *Collecting telemetry data privately*. URL: https://www.microsoft.com/en-us/research/blog/collecting-telemetry-data-privately/ (visited on November 16, 2022).

Microsoft (2021). *Password Monitor: Safeguarding passwords in Microsoft Edge*. URL: https://www.microsoft.com/en-us/research/blog/password-monitor-safeguarding-passwords-in-microsoft-edge/ (visited on November 16, 2022).

Microsoft (2022). *What is confidential computing?* URL: https://learn.microsoft.com/en-us/azure/confidential-computing/overview (visited on November 16, 2022).

Mo, F., Haddadi, H., Katevas, K., Marin, E., Perino, D. and Kourtellis, N. (2021). "PPFL: privacy-preserving federated learning with trusted execution environments," in: *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, New York, USA.

Morris, L. (2013). *Analysis of partially and fully homomorphic encryption*. Department of Computer Science, Rochester Institute of Technology. URL: http://gauss.ececs.uc.edu/Courses/c5156/pdf/homo-outline.pdf.

Mubarak, M. F. and Petraite, M. (2020). "Industry 4.0 technologies, digital trust and technological orientation: What matters in open innovation?," *Technological Forecasting and Social Change* 161, 1-11.

Mühle, A., Grüner, A., Gayvoronskaya, T. and Meinel, C. (2018). "A survey on essential components of a self-sovereign identity," *Computer Science Review* 30, 80-86.

Mulligan, D. P., Petri, G., Spinale, N., Stockwell, G. and Vincent, H. J. (2021). „Confidential Computing—a brave new world," *2021 International Symposium on Secure and Private Execution Environment Design (SEED)*, Washington, DC, USA.

Munjal, K. and Bhatia, R. (2022). "A systematic review of homomorphic encryption and its contributions in healthcare industry," *Complex & Intelligent Systems*, 1-28.

Oliveira, M. I. S., Lima, G. D. F. B. and Lóscio, B. F. (2019). "Investigations into Data Ecosystems: a systematic mapping study," *Knowledge and Information Systems* 61 (2), 589-630.

Oliveira, M. I. S. and Lóscio, B. F. (2018). "What is a data ecosystem?," in: *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, Delft, Netherlands.

Orlikowski, W. J. and Iacono, C. S. (2001). "Research commentary: Desperately seeking the "IT" in IT research—A call to theorizing the IT artifact," *Information Systems Research* 12 (2), 121-134.

Otto, B. (2019). "Interview with Reinhold Achatz on "data sovereignty and data ecosystems"," *Business & Information Systems Engineering* 61 (5), 635-636.

Otto, B. (2022). "A federated infrastructure for European data spaces," *Communications of the ACM* 65 (4), 44-45.

Otto, B. and Jarke, M. (2019). "Designing a multi-sided data platform: findings from the International Data Spaces case," *Electronic Markets* 29 (4), 561-580.

Pearson, S. and Casassa-Mont, M. (2011). "Sticky policies: An approach for managing privacy across multiple parties," *Computer* 44 (9), 60-68.

Peffers, K., Rothenberger, M., Tuunanen, T. and Vaezi, R. (2012). „Design science research evaluation," *International Conference on Design Science Research in Information Systems*, Las Vegas, USA.

Peffers, K., Tuunanen, T., Rothenberger, M. A. and Chatterjee, S. (2007). "A design science research methodology for information systems research," *Journal of Management Information Systems* 24 (3), 45-77.

Pigni, F., Piccoli, G. and Watson, R. (2016). "Digital data streams: Creating value from the real-time flow of big data," *California Management Review* 58 (3), 5-25.

Pereira Pires, R. (2019). Distributed systems and trusted execution environments: trade-offs and challenges. PhD thesis, Université de Neuchâtel.

Porter, M. E. and Heppelmann, J. E. (2014). "How smart, connected products are transforming competition," *Harvard Business Review* 92 (11), 64-88.

Russo, M. and Albert, M. (2018). *How IoT data ecosystems will transform B2B competition*. BCG Henderson Institute. URL: https://image-src.bcg.com/Images/BCG-How-IoT-Data-Ecosystems-Will-Transform-B2B-Competition-July-2018_tcm50-197926.pdf.

Sabt, M., Achemlal, M. and Bouabdallah, A. (2015). „Trusted execution environment: what it is, and what it is not,“ *14th IEEE International Conference on Trust, Security and Privacy in Computing and Communications*, Helsinki, Finland.

Sahut, J.-M., Schweizer, D. and Peris-Ortiz, M. (2022). “Technological forecasting and social change introduction to the VSI technological innovations to ensure confidence in the digital world,” *Technological Forecasting & Social Change* 179, 121680.

Saunders, M. and Townsend, K. (2018). “Choosing Participants,” in: Cassell, C., Cunliffe, A. L. & Grandy, G. (eds.) *The Sage Handbook of Qualitative Business and Management Research Methods.* SAGE Publications.

Schmidt, K., Munilla Garrido, G., Mühle, A. and Meinel, C. (2022). „Mitigating Sovereign Data Exchange Challenges: A Mapping to Apply Privacy-and Authenticity-Enhancing Technologies,“ *International Conference on Trust and Privacy in Digital Business*, Vienna, Austria.

Simon, H. A. (1957). *Models of man; social and rational,* New York: Wiley.

Song, Q., Cao, J., Sun, K., Li, Q. and Xu, K. (2021). “Try before You Buy: Privacy-preserving Data Evaluation on Cloud-based Machine Learning Data Marketplace,” *Annual Computer Security Applications Conference*, Virtual Event, USA.

Sumpf, P. (2019). *System Trust: Researching the Architecture of Trust in Systems,* Wiesbaden: Springer.

Van Blarkom, G., Borking, J. J. and Olk, J. E. (2003). *Handbook of privacy and privacy-enhancing technologies: The case of Intelligent Software Agents,* The Hague: Privacy Incorporated Software Agent (PISA) Consortium.

Von Solms, B. and Von Solms, R. (2018). “Cybersecurity and information security–what goes where?,” *Information & Computer Security* 26 (1), 2-9.

Wagner, P. G., Birnstill, P. and Beyerer, J. (2020). “Establishing secure communication channels using remote attestation with TPM 2.0,” *International Workshop on Security and Trust Management*, Guildford, UK.

Wang, F. and De Filippi, P. (2020). “Self-sovereign identity in a globalized world: Credentials-based identity systems as a driver for economic inclusion,” *Frontiers in Blockchain* 2 (28), 1-22.

Wang, K.-Y., Lin, G., Kuo, K., Lee, H.-C., Tsai, B. and Peng, W. (2020). „An empirical study of an open ecosystem model for inclusive financial services,“ *2020 IEEE International Conference on Services Computing (SCC)*, Beijing, China.

Weiss, E. (2018). *How to convince customers to share data after GDPR*. Harvard Business Review. URL: https://hbr.org/2018/05/how-to-convince-customers-to-share-data-after-gdpr.

Westin, A. F. (1967). *Privacy and Freedom,* New York: Atheneum.

Will, M. A. and Ko, R. K. L. (2015). “Chapter 5 - A guide to homomorphic encryption,” in: Ko, R. & Choo, K.-K. R. (eds.) *The Cloud Security Ecosystem.* Boston: Syngress.

Williamson, O. E. (1985). *Yhe Economic Institutions of Capitalism: Firms, markets, relational Contracting,* New York, USA: Free Press.

Wood, A., Najarian, K. and Kahrobaei, D. (2020). “Homomorphic encryption for machine learning in medicine and bioinformatics,” *ACM Computing Surveys (CSUR)* 53 (4), 1-35.

Zhao, Y. and Chen, J. (2022). “A survey on differential privacy for unstructured data content,” *ACM Computing Surveys (CSUR)* 54 (10s), 1-28.

Zöll, A., Olt, C. M. and Buxmann, P. (2021). “Privacy-sensitive business models: Barriers of organizational adoption of privacy-enhancing technologies,” in: *Proceedings of 29th European Conference on Information Systems (ECIS)*, Marrakesh, Morocco.