Exploring Requirements and Methods for Designing Sovereign Data Sharing

Julia Pampus

Fraunhofer Institute for Software and Systems Engineering ISST, Dortmund, Germany

Abstract

As the use of digital data becomes increasingly important in modern society, concerns about ownership and control over that data have emerged. The design and development of software systems for data sharing — covering integration and usage — focus on implementing security, privacy, and more recently, data sovereignty. However, there is a lack in scientific methods and best practices for designing software systems that establish this sovereignty. Existing literature discusses data sovereignty requirements but lacks clear underlying concepts for specifying, classifying, or eliciting these requirements. Therefore, our research explores how data sovereignty, as a collection of both functional and non-functional requirements, can be addressed in the software development and requirements engineering processes to support industrial data sharing. Based on literature reviews, case studies, and interviews with subject matter experts, we propose a method for analysing and implementing data sovereignty requirements. By defining and classifying these requirements, we streamline and unify their implementation, which is particularly beneficial for building trust in the collaborative establishment of sovereign data ecosystems. Consequently, the findings of our research have important implications for requirements engineers and decision-makers in data sharing contexts.

Keywords

data sovereignty, requirements engineering, data sharing

1. Motivation & Problem Statement

Where data is involved, the concepts of 'security' and 'privacy' are closely linked. Especially in a world where data is almost more valuable than gold, this data must be protected at all costs. Reports of data leaks and misuse are frequently reported in the media. As a result, both research and practice are increasingly focusing on mechanisms to protect data, including access restrictions, data flow control, prevention of inferences that could reveal confidential information, or encryption methods [1]. In addition to this heightened awareness of security, the idea of *sovereignty* is gaining more attention. Sovereignty refers to the autonomy and self-determination when it comes to technology choice and data management. One important aspect of sovereignty is *data sovereignty*, which is defined as "a natural person's or corporate entity's capability of being entirely self-determined with regard to its data" [2, p.71]. This concept of data sovereignty is receiving increasing attention in research, particularly in the context of trust-based data ecosystems and in applications such as the European Data Union Strategy[1]. Unlike the General Data Protection Regulation, which focuses on personal data, data sovereignty primarily views data "as a company asset" [3, p.3], encompassing business data, production data, and, to some extent, personal data.

The establishment of data sovereignty primarily focuses on two areas: data governance and usage control. Data governance involves creating guidelines and rules for managing data to maintain and increase its value. Usage control covers the technical implementation of these rules (i.e., policies), such as data usage constraints that determine how and by whom data can be transferred and processed. Unlike traditional access control, usage control offers the capability "to continuously monitor and control the usage of resources such as files or services" [4](p.289). In this way, usage control integrates

Joint Proceedings of REFSQ-2026 Workshops, Doctoral Symposium, Posters & Tools Track, and Education and Training Track. Co-located with REFSQ 2026. Poznan, Poland, March 23-26, 2026

julia.pampus@isst.fraunhofer.de (J. Pampus)

0000-0002-0877-7063 (J. Pampus)

[1] https://digital-strategy.ec.europa.eu/en/policies/data-union (last accessed on 2026-01-27)

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

published 2026-06-02

the fields of digital rights management (DRM), data leakage prevention, and access control [5].

Data spaces provide an architectural approach to achieving data sovereignty. By integrating legal, economic, and technical components, they establish a decentralised and neutral infrastructure of protocols and frameworks for data sharing among organisations. We motivate our problem statement from research (see Section 2) and practical experiences in building such data spaces. The development of the Dataspace Connector [6], the Eclipse Dataspace Components[2], and similar technologies has highlighted the various aspects involved in designing data space technologies focused on ensuring data sovereignty. The complexity of technology choice, distributed systems including multiple stakeholders, and various requirements increases the need for a significant but often insufficiently considered topic: requirements engineering. Requirements engineering is the process of eliciting, analysing, specifying, validating, and managing the requirements of a software system. Rupp et al. [7] outline seven common challenges with this process, which we also observed in the context of sovereign data sharing: unclear objectives, communication problems, insufficient quality, edge solutions, changing requirements, inaccurate planning, and comprehensive dependencies.

Problem Statement: The novelty of data sovereignty currently results in limited research on requirements analysis and software system design for sovereign data sharing, leaving practitioners without adequate models, methods, and best practices.

2. Related Work

To scientifically explore requirements and methods for designing sovereign data sharing, we conducted an integrated literature review in early 2023. This type of review is characterised by a certain "provocativeness" [8](p.364). During the search process, researchers examine and critique the existing literature on a specific topic, providing new perspectives for further research. As a result, they may identify alternative models or conceptual frameworks that directly address the research gaps identified [8]. Based on our experiences in projects focused on implementing data sovereignty (cf. Section 1), we posed a hypothesis: *Current literature lacks consistent and reusable models or methods for organising and analysing requirements across technical, organisational, and legal dimensions related to data sovereignty.*

Our review process adhered to the structure of systematic literature reviews in software engineering described by Kitchenham [9]. Throughout the review, we iteratively refined our search terms: We started with key terms such as 'usage control', 'data ecosystem', 'data sovereignty', and 'data governance', and then expanded our search to include related topics such as 'confidentiality', 'privacy', and 'DRM'. We subsequently concentrated on requirements engineering and design methods as specific areas of interest. Our research was guided by several questions: What aspects of data sovereignty are addressed in existing research? What requirements exist for sovereign data sharing? What approaches can support the requirements analysis and design processes? Which stakeholders are part of the design process?

Table 1 outlines our data collection process. We conducted an initial search of common scientific databases, including Scopus, the ACM Digital Library, and IEEE Xplore, using the specified search terms. We included both scientific and practice-oriented papers while excluding articles of other types and those with no access or thematic relevance. Afterwards, we filtered the result set based on the type, formatting, and title of the articles. We examined the frequency of terms used in keywords and abstracts, leading us to identify and exclude articles that do not focus on data sovereignty. The final result set consists of 28 articles, indicating that only a small subset of the initially discovered articles explicitly address requirements engineering and software system design for sovereign data sharing.

A substantial number of articles on data sovereignty originate from the field of data business and economics, primarily focusing on manufacturing data. The concept of data sovereignty is closely related to the topics of data sharing, confidentiality, and privacy. Relevant mechanisms for implementing requirements in this domain involve policy specifications and enforcement. Numerous studies examine different policy languages and their technical interpretations and implementations, particularly regarding access and usage control. Both concepts have their own models, frameworks, and enforcement

[2] https://projects.eclipse.org/projects/technology.edc (last accessed on 2026-02-16)

Table 1
Process of Data Collection

Step	Criteria	Result Set
1. Initial Search	search terms	584
2. Filtering	publication type, formatting, title	510
3. Coding & Filtering	keywords, abstract	510
4. Classification	requirements development, design process	28

mechanisms. On a technical level, the discussions often intersect with considerations of security and trust, which, in turn, form a distinct area of research focused on trusted computing and secure systems.

The literature and selected case studies typically address data sovereignty requirements only briefly [10, 11, 12, 13, 14, 15, 16, 5, 17, 18]. The identified requirements often consist of both functional and non-functional requirements, emphasising the technical functionalities of information systems: e.g., role-restricted or system-restricted access control [5], a need for parallel processing of many requests [5], and the scalability of deployments [5, 10]. These diverse requirements indicate that a clear definition of data sovereignty requirements is lacking, making it challenging to determine how these requirements align with typical software system requirements, particularly concerning security and scalability. Furthermore, there are no formal guidelines or processes available for defining and implementing such requirements, specifically in relation to data sovereignty.

However, we identified several articles that focus on different aspects of potential requirements engineering methods, spanning from requirements specification to the implementation of security, privacy, and/or sovereignty. For example, Narouei et al. [19] explore how to automatically extract access control policies using natural language processing. They use semi-supervised learning and develop a framework consisting of lexical and semantic parsers. This framework helps in identifying relevant keywords in natural language and generating technically processable policies. Schaad and Mnakva [20] examine the integration of business processes in an enterprise context with usage control for design-time and runtime. They identify and classify assets involved in a business process and define and implement the necessary controls and status changes. Al-Ruithe et al. [21] propose a conceptual framework for designing cloud data governance. This framework outlines five steps: establishing the governance structure, evaluating existing mechanisms, setting up governance functions, negotiating contracts, and determining governance level agreement. Hosseinzadeh et al. [12] introduce an approach for deriving technical policies from usage control requirements. They use a graphical editor that maps form inputs to predefined policy rules.

In recent years, research on data sovereignty exploded. However, the growing number of use cases has led to more data sovereignty requirements than the development of specific methods for collecting and analysing these requirements. As part of our initial investigations, we conducted an empirical study in 2023 [22], interviewing eleven experts to identify their requirements for sovereign data sharing and the challenges they face with existing solutions. The results of this study revealed that the requirements vary widely, encompassing technical, organisational, and legal aspects. During the interviews, we observed that participants found it challenging to formulate and define data sovereignty requirements, which reinforced our problem statement (cf. Section 1). Overall, we identified four research gaps:

1. **Comprehensive analysis of data sovereignty**: There is a lack of sufficient analysis of technical aspects of data sovereignty that respect organisational and legal measures.
2. **Standardised terminology and methodological support**: A unified definition for describing data sovereignty requirements and methodological support for requirements elicitation is missing.
3. **Structured approach for requirements analysis**: There is no structured approach for analysing data sovereignty requirements for conflicts and translating these requirements into system design.
4. **Holistic reference architectures**: Comprehensive data sovereignty models addressing technical and organisational measures while accounting for cross-system dependencies and varying legal requirements are missing.

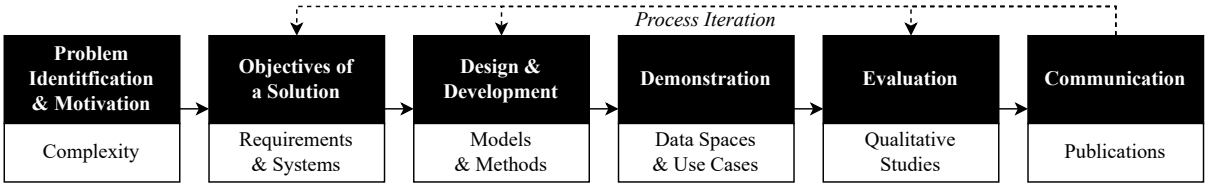

Figure 1: Research Process Following DSR by Peffers et al. [24]

3. Research Method

Our research has two primary research goals: (1) to characterise the requirements for sovereign data sharing, and (2) to develop methods for analysing these requirements and designing software systems that establish data sovereignty. Along with these goals, we define three main research questions (MRQs):

- MRQ1: What characterises data sovereignty in industrial data sharing?
- MRQ2: What are requirements for sovereign data sharing, and how can they be structured?
- MRQ3: What methods support requirements engineering for sovereign data sharing?

In our research, we follow the process of the design science research (DSR) paradigm [23], which is gaining increasing attention, particularly in the field of information systems. Our research process, illustrated in Figure 1, consists of six steps: motivation and problem identification, development of objectives, design and development, demonstration, evaluation, and communication [24]. First, we motivate our research through *problem identification*, which is derived from practical applications (cf. Section 1) and supported by the literature (cf. Section 2). As we formulate the *objectives of our solution*, we examine existing requirements for sovereign data sharing. Based on these requirements, we *design and develop* multiple artefacts, such as models and methods, each addressing one of our MRQs. We then *demonstrate* and *evaluate* the applicability of each artefact, ultimately making our findings available to both research and practice in the form of a *publication.*

4. Proposed Solution

We started our research journey by elaborating the characteristics of software systems that establish the sovereignty of data rights holders in providing and using their data (cf. MRQ1). Through a questionnaire survey of 18 experts, we discovered that data sovereignty is technically realised as functional and non-functional software system requirements. These requirements primarily focus on security, intervenability, and interoperability [25, 26]. In line with Haley et al. [27], focusing on security, we consider data sovereignty as a soft goal within the requirements engineering process and the software system design phase. We subsequently aimed to define data sovereignty theoretically and distinguish it from the concept of privacy [28]. Data sovereignty can be analysed in various contexts, including economic, organisational, regulatory, and technical aspects. We compared sovereignty and privacy against the criteria of the taxonomy of privacy-enhancing technologies by Heurix et al. [29], and conclude that privacy can be considered as a subset of data sovereignty. While the fundamental principles are the same, the scenarios, objectives, and types of data involved differ [28]. In our work, we therefore focus primarily on industrial data sharing without specifically addressing personal data.

With our newly established theoretical knowledge base, we examined the concrete requirements related to data sovereignty that arise in practice (cf. MRQ2). We conducted practical desk research [30] and a literature review [31]. Our investigation covered a wide range of requirement types, leading us to analyse and derive multiple data sharing requirement models [32, 30]. These models incorporate data, actors, and policies, with policies being the most important element. We identified three types of policies (prohibition, permission, and duty) with different focuses on actor, time, and/or environment constraints [30], and derived a grammar that can be used for the (automatic) analysis of these policies.

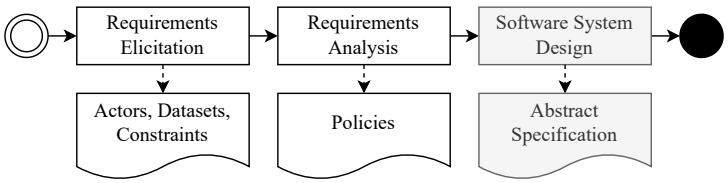

Figure 2: Basic Structure of a Method for Analysing Requirements for Sovereign Data Sharing

Consequently, we studied the surrounding environment and the influences of the individual actors and elements within our developed requirements models to achieve extensive interoperability. In line with the different contexts of data sovereignty, we delimit the business environment, legal environment, and system environment as part of a holistic conceptual model for sovereign data sharing [33, 31].

Addressing MRQ3, we propose a requirements engineering method that outlines a three-step process formatted as an *agenda* [34]. Each step specifies input and output values as well as validation conditions, which allow for structured and potentially automated checks of completeness and correctness. Our method must meet several requirements: it should be agile and allow iterative adjustments, involve ongoing collaboration among various types of stakeholders within and across organisations, and place particular emphasis on defining data flows, data types, and policies [31]. Our proposed method includes three steps: elicitation, analysis, and design, as illustrated in Figure 2. In the first step, requirements are gathered, guided by established patterns [32]. In the second step, these requirements are analysed, with a key focus on policies. A crucial sub-step in this phase is conflict analysis, where potential conflicts are evaluated [30]. In the third and final step, a system design is created based on the collected requirements, specifically addressing the design of the policy system, which is the core component of software systems used for sovereign data sharing. The outputs of this process include various components and process information that can be utilised in subsequent software development.

In summary, our proposed solution addresses the outlined problems (cf. Sections 1 and 2) as follows: (1) By characterising data sovereignty through its technical aspects and delimiting it from related concepts, we provide a foundational understanding of dependencies among data, actors, policies, and system components in data sharing. (2) Our requirement models establish a unified vocabulary for describing and categorising data sovereignty requirements. This supports stakeholder collaboration and communication. (3) Our requirements engineering method provides requirements engineers with structured guidance, specifically addressing the identification of requirement conflicts. (4) Our holistic conceptual model for sovereign data sharing bridges the gap between requirements and implementation.

5. Research Plan

In upcoming research, we will focus on two core activities: (a) finalising our proposed solution for analysing requirements for sovereign data sharing, and (b) evaluating this solution.

Method A key component of our future work involves defining the third step of the requirements engineering process, known as the abstract specification (see Figure 2). The challenge lies in finding the right balance between usability and comprehensiveness of the method for varying complexities in IT landscapes. The complexity of policies and their implementation is reflected in software specifications. Subsequently, we will integrate the methods we have developed over the last few years — including the requirements elicitation process [32], the policy analysis process, including conflict resolution [30], and the design process — to create a unified method for eliciting and analysing sovereignty requirements. We will define comprehensive validation conditions to verify the outputs of each process step.

Evaluation We will present our method supported by scientific foundations and appropriate demonstrations and evaluations. To evaluate our overall requirements engineering process, we will adhere to the guidelines for industry-based case studies in software engineering as proposed by Verner [35].

We plan to focus our case study on Catena-X, an ecosystem in the automotive industry that aims to enhance data sharing and interoperability throughout the value chain. We will gather data through expert interviews. The findings will help determine whether our proposed solution is comprehensive and effective in practice, as well as identify any limitations that may require further investigation. Key challenges include recruiting sufficient domain experts for interviews and ensuring generalisability beyond the automotive domain.

Acknowledgments

This work was partly supported by the career and development program Fraunhofer TALENTA, which is funded by the Fraunhofer-Gesellschaft.

Declaration on Generative AI

During the preparation of this work, the author used DeepL and Grammarly in order to: Translation, grammar, and spelling check. After using these tools/services, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] D. E. Denning, P. J. Denning, Data Security, ACM Computing Surveys 11 (1979) 227–249.
[2] B. Otto, S. Lohmann, S. Auer, G. Brost, J. Cirullies, A. Eitel, T. Ernst, C. Haas, M. Huber, C. Jung, et al., Reference Architecture Model for the Industrial Data Space, Fraunhofer-Gesellschaft, Munich 88 (2017).
[3] B. Otto, Organizing Data Governance: Findings from the Telecommunications Industry and Consequences for Large Service Providers, Communications of the Association for Information Systems 29 (2011).
[4] J. Schuette, G. S. Brost, LUCON: Data Flow Control for Message-Based IoT Systems, in: 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), 2018, pp. 289–299.
[5] J. Zrenner, F. O. Möller, C. Jung, A. Eitel, B. Otto, Usage control architecture options for data sovereignty in business ecosystems, Journal of Enterprise Information Management 32 (2019) 477–495.
[6] J. Pampus, B. Jahnke, R. Quensel, Evolving Data Space Technologies: Lessons Learned from an IDS Connector Reference Implementation, in: Leveraging Applications of Formal Methods, Verification and Validation. Practice, volume 13704 of *Lecture Notes in Computer Science*, 2022, pp. 366–381.
[7] C. Rupp, M. Simon, F. Hocker, Requirements Engineering und Management, Requirements Engineering und Management (2009) 10.
[8] R. J. Torraco, Writing Integrative Literature Reviews: Guidelines and Examples, Human Resource Development Review 4 (2005) 356–367.
[9] B. Kitchenham, Guidelines for performing Systematic Literature Reviews in Software Engineering, Technical Report, Keele University, 2007.
[10] I. Akaichi, S. Kirrane, Usage Control Specification, Enforcement, and Robustness: A Survey, 2022.
[11] T. Ermakova, B. Fabian, R. Zarnekow, Security and privacy system requirements for adopting cloud computing in healthcare data sharing scenarios, 2013.
[12] A. Hosseinzadeh, A. Eitel, C. Jung, A Systematic Approach toward Extracting Technically Enforceable Policies from Data Usage Control Requirements, in: Proceedings of the 6th International Conference on Information Systems Security and Privacy, 2020, pp. 397–405.

[13] K. K. Kim, D. K. Browe, H. C. Logan, R. Holm, L. Hack, L. Ohno-Machado, Data governance requirements for distributed clinical research networks: triangulating perspectives of diverse stakeholders, Journal of the American Medical Informatics Association 21 (2014) 714–719.
[14] F. Larrinaga, Data Sovereignty - Requirements Analysis of Manufacturing Use Cases, 2022.
[15] D. Lee, Building an Open Data Ecosystem – An Irish Experience, 2014.
[16] S. Opriel, F. Möller, U. Burkhardt, B. Otto, Requirements for Usage Control based Exchange of Sensitive Data in Automotive Supply Chains, in: Proceedings of the 54th Hawaii International Conference on System Sciences, 2021, pp. 431–440.
[17] C. Frey, P. Hertweck, L. Richter, O. Warweg, Bauhaus.MobilityLab: A Living Lab for the Development and Evaluation of AI-Assisted Services, Smart Cities 5 (2022) 133–145.
[18] M. Altendeitering, J. Pampus, F. Larrinaga, J. Legaristi, F. Howar, Data sovereignty for AI pipelines: lessons learned from an industrial project at Mondragon corporation, in: Proceedings of the 1st International Conference on AI Engineering: Software Engineering for AI, 2022, pp. 193–204.
[19] M. Narouei, H. Takabi, R. Nielsen, Automatic Extraction of Access Control Policies from Natural Language Documents, IEEE Transactions on Dependable and Secure Computing (2018) 1.
[20] A. Schaad, A. Monakva, Annotating Business Processes with Usage Controls, in: Proceedings of the WWW2012 workshop on Data Usage Management on the Web, 2012, pp. 23–28.
[21] M. Al-Ruithe, E. Benkhelifa, K. Hameed, A Conceptual Framework for Designing Data Governance for Cloud Computing, Procedia Computer Science 94 (2016) 160–167.
[22] M. Hellmeier, J. Pampus, H. Qarawlus, F. Howar, Implementing Data Sovereignty: Requirements & Challenges from Practice, in: Proceedings of the 18th International Conference on Availability, Reliability and Security, 2023, pp. 1–9.
[23] A. R. Hevner, S. T. March, J. Park, S. Ram, Design Science in Information Systems Research, Management Information Systems Quarterly 28 (2004) 75–105.
[24] K. Peffers, T. Tuunanen, C. E. Gengler, M. Rossi, W. Hui, V. Virtanen, J. Bragge, The design science research process: A model for producing and presenting information systems research, in: Proceedings of the First International Conference on Design Science Research in Information Systems and Technology (DESRIST 2006), 2006, p. 25.
[25] J. Pampus, M. Heisel, An Empirical Examination of the Technical Aspects of Data Sovereignty, in: Proceedings of the 19th International Conference on Software Technologies, 2024, pp. 112–122.
[26] J. Pampus, M. Heisel, Feature-Oriented Requirements Analysis for Sovereign Data Sharing, in press.
[27] C. B. Haley, J. D. Moffett, R. Laney, B. Nuseibeh, A framework for security requirements engineering, in: Proceedings of the 2006 international workshop on Software engineering for secure systems, 2006, pp. 35–42.
[28] J. Pampus, A Delimitation of Data Sovereignty From Privacy, Empowering Digital Sovereignty: Balancing Privacy and Global Connectivity: Balancing Privacy and Global Connectivity (2025) 1.
[29] J. Heurix, P. Zimmermann, T. Neubauer, S. Fenz, A taxonomy for privacy enhancing technologies, Computers & Security 53 (2015) 1–17.
[30] J. Pampus, M. Heisel, A Pattern for Analysing Data Usage Requirements in Data Sharing, in press.
[31] J. Pampus, M. Heisel, Designing Software Systems for Sovereign Data Sharing: A Model-based Approach, in press.
[32] J. Pampus, M. Heisel, Pattern-based Requirements Elicitation for Sovereign Data Sharing, Procedia Computer Science 254 (2025) 147–156.
[33] J. Pampus, M. Heisel, Towards Consistent Policy Enforcement in Dataspaces, in: Proceedings of the 14th International Conference on Data Science, Technology and Applications, 2025, pp. 560–566.
[34] M. Heisel, Agendas – A Concept to Guide Software Development Activites, in: Proc. Systems Implementation 2000, Chapman & Hall London, 1998, pp. 19–32.
[35] J. M. Verner, J. Sampson, V. Tosic, N. A. Bakar, B. A. Kitchenham, Guidelines for industrially-based multiple case studies in software engineering, in: 2009 Third International Conference on Research Challenges in Information Science, IEEE, 2009.