

# Data Spaces: Design, Deployment, and Future Directions



Edward Curry, Simon Scerri, and Tuomo Tuikka

**Abstract** Digital transformation, data ecosystems, and Data Spaces are inevitable parts of our future. The book aims to educate the reader on data sharing and exchange techniques using Data Spaces. It will address and explore the cutting-edge theory, technologies, methodologies, and best practices for Data Spaces for both industrial and personal data. The book provides the reader with a basis for understanding the scientific foundation of Data Spaces, how they can be designed and deployed, and future directions.

**Keywords** Data Spaces · Data ecosystem · Big Data value · Data innovation

## 1 Introduction

Digital transformation creates a data ecosystem with data on every aspect of our world. The rapidly increasing volumes of diverse data from distributed sources create significant opportunities for extracting valuable knowledge. Data ecosystems can create the conditions for a marketplace competition among participants or enable collaboration among diverse, interconnected participants that depend on each other for their mutual benefit. A data space can provide a clear framework to support data sharing within a data ecosystem. For example, industrial Data Spaces can support the trusted and secure sharing and trading of commercial data assets with automated and robust controls on legal compliance and remuneration of data

---

E. Curry (✉)

Insight SFI Research Centre for Data Analytics, University of Galway, Galway, Ireland  
e-mail: [edward.curry@insight-centre.org](mailto:edward.curry@insight-centre.org)

S. Scerri  
metaphacts, Walldorf, Germany

T. Tuikka  
VTT Technical Research Centre of Finland, Oulu, Finland  
e-mail: [tuomo.tuikka@vtt.fi](mailto:tuomo.tuikka@vtt.fi)

owners. Personal Data Spaces enforce legislation and allow data subjects and data owners to control their data and its subsequent use.

Many fundamental technical, organizational, legal, and commercial challenges exist in developing and deploying Data Spaces to support data ecosystems. For example, how do we create trusted and secure Data Spaces and privacy-aware analytics methods for secure sharing of personal data and industrial data? How can small- and medium-sized enterprises get access to Data Spaces and technology? How can we support the utility trade-offs between data analysis and privacy? What are user-friendly privacy metrics for end-users? What are the standardization challenges for Data Spaces, including interoperability? How do Data Spaces ensure secure and controlled sharing of proprietary or personal data? What are the necessary technical, organizational, legal, and commercial best practices for data sharing, brokerage, and trading?

The book aims to educate the reader on data sharing and exchange techniques using Data Spaces. The book will address and explore the cutting-edge theory, technologies, methodologies, and best practices for Data Spaces for both industrial and personal data. In addition, the book provides the reader with a basis for understanding the scientific foundation of Data Spaces, how they can be designed and deployed, and future directions.

The chapter is structured as follows: Sect. 2 defines the notion of data ecosystems. Section 3 introduces the concepts of Data Spaces and their role as a platform and their role for sharing industrial and personal data. Section 4 discusses common European Data Spaces and outlines how their foundations have been established by the Big Data Value Public-Private Partnership (PPP) with the data platform projects. Section 5 details the book's structure in the three key areas of design, deployment, and future directions, together with an analysis of the contribution of the chapter's Data Spaces in terms of value, data, technology, organization people, governance, and trust. Finally, Sect. 6 provides a summary.

## 2 Data Ecosystems

A data ecosystem is a sociotechnical system enabling value to be extracted from data value chains supported by interacting organizations and individuals [1]. Data value chains are oriented to business and societal purposes within an ecosystem. The ecosystem can create the conditions for a marketplace competition among participants or enable collaboration among diverse, interconnected participants that depend on each other for their mutual benefit.

Digital transformation is creating a data ecosystem with data on every aspect of our world, spread across a range of intelligent systems, with structured and unstructured data (e.g., images, video, audio, and text) that can be exploited by data-driven intelligent systems to deliver value.

There is a need to bring together data from multiple sources within the data ecosystem. For example, smart cities show how different systems within the city

(e.g., energy and transport) can collaborate to maximize the potential to optimize overall city operations. At the level of an individual, digital services can deliver a personalized and seamless user experience by bringing together relevant user data from multiple systems [2]. This requires a system of systems (SoS) approach to connect systems that cross organizational boundaries, come from various domains (e.g., finance, manufacturing, facilities, IT, water, traffic, and waste), and operate at different levels (e.g., region, district, neighborhood, building, business function, individual).

Data ecosystems present new challenges to the design of data sharing that require a rethink in how we should deal with the needs of large-scale data-rich environments with multiple participants. There is a clear need to support knowledge sharing among participants within data ecosystems. Meeting these challenges is critical to maximizing the potential of data-intensive intelligent systems [3].

### 3 Data Spaces

The term “dataspace” or “data space” can now be seen as an umbrella term categorizing several closely related concepts. First introduced by Franklin, Halvey, and Maier in 2005 [4] within the data management community, a data space can contain all the data sources for an organization regardless of its format, location, or model. Each data source (e.g., database, CSV, web service) in the data space is known as a participant. The Franklin et al. data space can model the relations (or associations) between data in different participants. In its purest form, a data space is a set of participants and the inter-relations between them [4]. The modeling of the data space can capture different types of relations among participants, from mapping the schemas between two participants to capturing that Participant A is a replica of Participant B.

The data space concept has gained traction with several groups exploring its usefulness for managing data from different domains and regions within a global data ecosystem. These works have provided many definitions for a data space, as captured in Table 1. For example, the Big Data Value Association (BDVA) view of Data Spaces is any ecosystem of data models, datasets, ontologies, data sharing contracts, and specialized management services (i.e., as often provided by data centers, stores, repositories, individually, or within “data lakes”), together with soft competencies around it (i.e., governance, social interactions, business processes) [16]. These competencies follow a data engineering approach to optimize data storage and exchange mechanisms, preserving, generating, and sharing new knowledge.

**Table 1** Definitions of a “dataspace” from literature (Adapted from Curry [5])

Definition	Source
“Dataspace are not a data integration approach; rather, they are more of a data co-existence approach. The goal of dataspace support is to provide base functionality over all data sources, regardless of how integrated they are.”	Halevy et al. [6]
“A data space is defined as a decentralised infrastructure for trustworthy data sharing and exchange in data ecosystems based in commonly agreed principles.”	Nagel [7]
“A dataspace system manages the large-scale heterogeneous collection of data distributed over various data sources in different formats. It addresses the structured, semi-structured, and unstructured data in coordinated manner without presuming the semantic integration among them.”	Singh [8]
“to provide various of the benefits of classical data integration, but with reduced up-front costs, combined with opportunities for incremental refinement, enabling a “pay-as-you-go” approach.”	Hedeler et al. [9]
“enable agile data integration with much lower upfront and maintenance costs.”	Hedeler et al. [10]
“A dataspace system processes data, with various formats, accessible through many systems with different interfaces, such as relational, sequential, XML, RDF, etc. Unlike data integration over DBMS, a dataspace system does not have full control on its data, and gradually integrates data as necessary.”	Wang et al. [11]
“Dataspace Support Platforms envision data integration systems where the amount of upfront effort is much smaller. The system should be able to bootstrap itself and provide some useful services with no human intervention. Over time, through user feedback or as sources are added and the data management needs become clearer, the system evolves in a pay-as-you-go fashion.”	Das Sarma et al. [12]
“Dataspace is defined as a set of participants and a set of relationships among them.”	Singh and Jain [13]
“Real-time Linked Dataspace combines the pay-as-you-go paradigm of dataspace with Linked Data, Knowledge Graphs, and real-time stream and event processing capabilities to support the large-scale distributed heterogeneous collection of streams, events, and data sources.”	Curry [14], Curry et al. [15]
“any ecosystem of data models, datasets, ontologies, data sharing contracts and specialised management services (i.e., as often provided by data centres, stores, repositories, individually or within ‘data lakes’), together with soft competencies around it (i.e., governance, social interactions, business processes).”	Scerri et al. [16]

### ***3.1 Data Spaces: A Platform for Data Sharing***

Data-driven Artificial Intelligence is revolutionizing many industries, including transportation and logistics, security, manufacturing, energy, healthcare, and agriculture, by providing intelligence to improve efficiency, quality, and flexibility. Data sharing is a critical enabler for competitive AI solutions. Data for AI is recognized as an innovation ecosystem in the European AI, data, and robotics framework [17]. In addition, data sharing and trading are enablers in the data economy, although closed and personal data present particular challenges for the free flow of data.

Platform approaches have proved successful in many areas of technology [18], from supporting transactions among buyers and sellers in marketplaces (e.g., Amazon), innovation platforms that provide a foundation on top of which to develop complementary products or services (e.g., Windows), to integrated platforms which are a combined transaction and innovation platform (e.g., Android and the Play Store).

The idea of large-scale “data” platforms has been touted as a possible next step to support data ecosystems [3]. An ecosystem data platform would have to support continuous, coordinated data flows, seamlessly moving data among intelligent systems. The design of infrastructure to support data sharing and reuse is still an active area of research [19]. The following two conceptual solutions—Industrial Data Spaces (IDS) and Personal Data Spaces (PDS)—introduce new approaches to addressing this particular need to regulate closed proprietary and personal data.

#### **3.1.1 Industrial Data Spaces (IDS)**

IDS has increasingly been touted as potential catalysts for advancing the European data economy as solutions for emerging data markets, focusing on the need to offer secure and trusted data sharing to interested parties, primarily from the private sector (industrial implementations). The IDS conceptual solution is oriented toward proprietary (or closed) data. Its realization should guarantee a trusted, secure environment where participants can safely and legally monetize and exchange their data assets within a clear legal framework. A functional realization of a continent-wide IDS promises to significantly reduce the existing barriers to a free flow of data within an advanced European data economy. Furthermore, the establishment of a trusted data sharing environment will have a substantial impact on the data economy by incentivizing the marketing and sharing of proprietary data assets (currently widely considered by the private sector as out of bounds) through guarantees for fair and safe financial compensations set out in black-and-white legal terms and obligations for both data owners and users. The “opening up” of previously guarded private data can thus vastly increase its value by several orders of magnitude, boosting the data economy and enabling cross-sectoral applications that were previously unattainable or only possible following one-off bilateral agreements between parties over specific data assets.

Notable advances in IDS include the highly relevant white paper and the reference architecture<sup>1</sup> provided by the International Data Spaces Association (IDSA). In addition, the layered databus, introduced by the Industrial Internet Consortium,<sup>2</sup> and the MindSphere Open Industrial Cloud Platform<sup>3</sup> are all examples of the need for data-centric information-sharing technology that enables data market players to exchange data within a virtual and global data space.

The implementation of Data Spaces needs to be approached on a European level, and existing and planned EU-wide, national, and regional platform development activities could contribute to these efforts as recognized by the European data strategy (*Communication: A European Strategy for Data*, 2020).

### 3.1.2 Personal Data Spaces (PDS)

So far, consumers have trusted, including companies like Google, Amazon, Facebook, Apple, and Microsoft, to aggregate and use their personal data in return for free services. While EU legislation, through directives such as the Data Protection Directive (1995) and the ePrivacy Directive (1998), has ensured that personal data can only be processed lawfully and for legitimate use, the limited user control offered by such companies and their abuse of a lack of transparency have undermined the consumer's trust. In particular, consumers experience everyday leakage of their data, traded by giant aggregators in the marketing networks for value only returned to consumers in the form of often unwanted digital advertisements. This has recently led to a growth in the number of consumers adopting adblockers to protect their digital life. At the same time, they are becoming more conscious of and suspicious about their personal data trail.

In order to address this growing distrust, the concept of personal Data Spaces (PDS) has emerged as a possible solution that could allow data subjects and data owners to remain in control of their data and its subsequent use.<sup>4</sup> PDS leverages “the concept of user-controlled cloud-based technologies for storage and use of personal data.” However, consumers have only been able to store and control access to a limited set of personal data, mainly by connecting their social media

---

<sup>1</sup> Reference Architecture Model for the Industrial Data Space, April 2017, [https://www.fraunhofer.de/content/dam/zv/de/Forschungsfelder/industrial-data-space/Industrial-Data-Space\\_Reference-Architecture-Model-2017.pdf](https://www.fraunhofer.de/content/dam/zv/de/Forschungsfelder/industrial-data-space/Industrial-Data-Space_Reference-Architecture-Model-2017.pdf)

<sup>2</sup> The Industrial Internet of Things, Volume G1: Reference Architecture, January 2017, [https://www.iiconsortium.org/IIC\\_PUB\\_G1\\_V1.80\\_2017-01-31.pdf](https://www.iiconsortium.org/IIC_PUB_G1_V1.80_2017-01-31.pdf)

<sup>3</sup> *MindSphere: The cloud-based, open IoT operating system for digital transformation*, Siemens, 2017, [https://www.plm.automation.siemens.com/media/global/en/Siemens\\_MindSphere\\_Whitepaper\\_tcm27-9395.pdf](https://www.plm.automation.siemens.com/media/global/en/Siemens_MindSphere_Whitepaper_tcm27-9395.pdf)

<sup>4</sup> See a Commission paper on “Personal information management services—current state of service offers and challenges” analyzing feedback from public consultation: <https://ec.europa.eu/digital-single-market/en/news/emerging-offer-personal-information-management-services-current-state-service-offers-and>

profiles to various emerging Personal Information Management Systems (PIMS). More successful (but limited in number) uses of PDS have involved the support of large organizations in agreeing to their customers accumulating data in their own self-controlled spaces. The expectation here is the reduction of their liability in securing such data and the opportunity to access and combine them with other data that individuals will import and accumulate from other aggregators. However, a degree of friction and the lack of a successful business model are still hindering the potential of the PDS approach.

A recent driver behind such a self-managed personal data economy is the General Data Protection Regulation (GDPR), which constitutes the single pan-European law on data protection, which requires companies dealing with European consumers to (1) increase transparency, (2) provide users with granular control for data access and sharing, and (3) guarantee consumers a set of fundamental individual digital rights (including the right to rectification, erasure, and data portability and to restrict processing). This creates new opportunities for PDS to emerge. Furthermore, the rise of PDS and the creation of more decentralized personal datasets will also open up new opportunities for SMEs that might benefit from and investigate new secondary uses of such data by gaining access to them from user-controlled personal data stores – a privilege so far available only to large data aggregators. However, further debate is required to understand the best business models (for demand and supply) to develop a marketplace for personal data donors and the mechanisms required to demonstrate transparency and distribute rewards to personal data donors. Finally, questions around data portability and interoperability also have to be addressed.

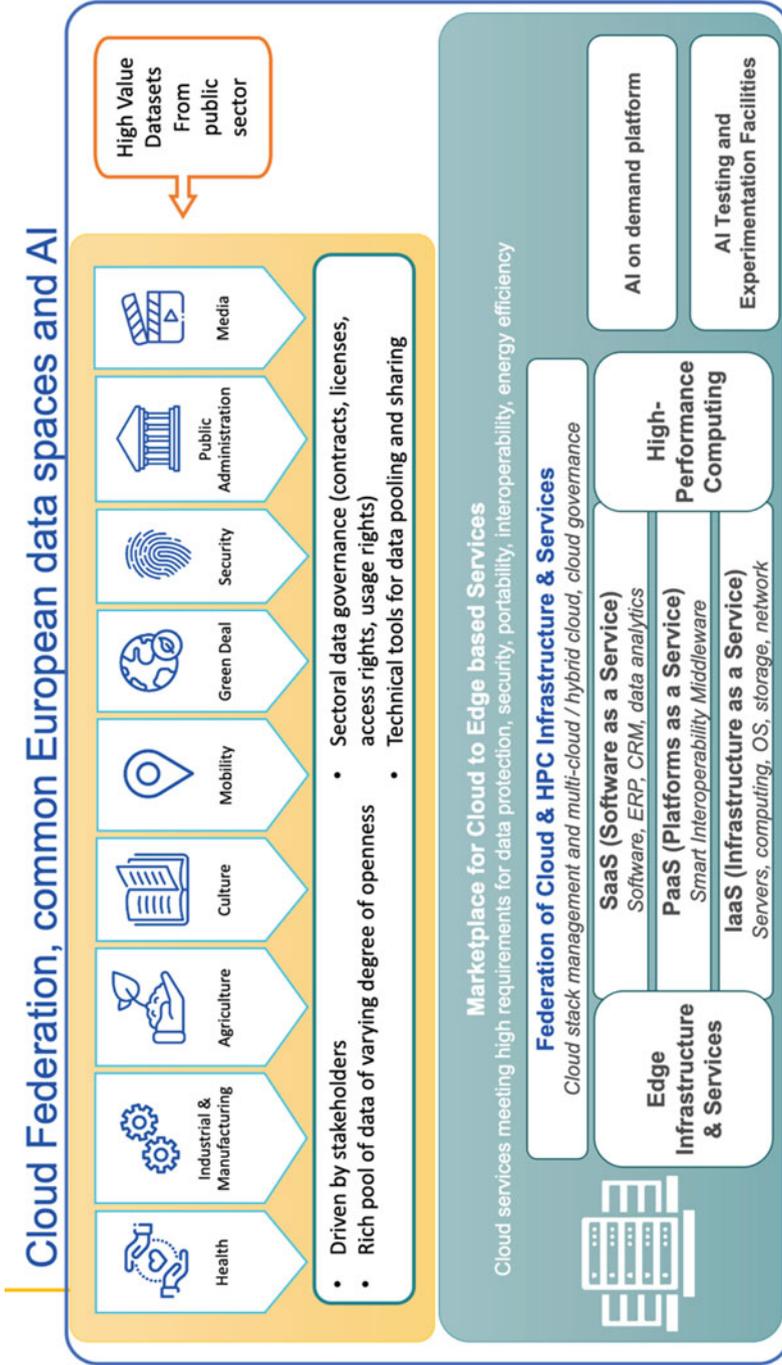
## 4 Common European Data Spaces

The European strategy for data aims at creating a single market for data that will ensure Europe’s global competitiveness and data sovereignty. The strategy aims to ensure:

- Data can flow within the EU and across sectors.
- Availability of high-quality data to create and innovate.
- European rules and values are fully respected.
- Rules for access and use of data are fair, practical, and clear and precise Data Governance mechanisms are in place.

Common European Data Spaces will ensure that more data becomes available in the economy and society while keeping companies and individuals who generate the data in control (*Communication: A European Strategy for Data*, 2020). Furthermore, as illustrated in Fig. 1, common European Data Spaces will be central to enabling AI techniques and supporting the marketplace for cloud and edge-based services.

As the first concrete steps toward implementing common European Data Spaces, a set of research and innovation actions for data platforms have been funded as part



**Fig. 1** Overview of cloud federation, common European Data Spaces, and AI (Communication: A European Strategy for Data, 2020)

of the Big Data Value PPP. Data platforms refer to architectures and repositories of interoperable hardware/software components, which follow a software engineering approach to enable the creation, transformation, evolution, curation, and exploitation of static and dynamic data in Data Spaces. In the remainder of this section, we describe the Big Data Value PPP, the Big Data Value Association, and the data platform project portfolio of the PPP.

#### ***4.1 The Big Data Value PPP (BDV PPP)***

The **European contractual Public-Private Partnership on Big Data Value (BDV PPP)** commenced in 2015. It was operationalized with the Leadership in Enabling and Industrial Technologies (LEIT) work program of Horizon 2020. The BDV PPP activities addressed the development of technology and applications, business model discovery, ecosystem validation, skills profiling, regulatory and IPR environments, and many social aspects.

With an initial indicative budget from the European Union of €534M by 2020, the BDV PPP had projects covering a spectrum of data-driven innovations in sectors including advanced manufacturing, transport and logistics, health, and bioeconomy [20]. These projects have advanced state of the art in key enabling technologies for Big Data value and non-technological aspects such as providing solutions, platforms, tools, frameworks, best practices, and invaluable general innovations, setting up firm foundations for a data-driven economy and the future European competitiveness in data and AI [21].

#### ***4.2 Big Data Value Association***

The Big Data Value Association (BDVA) is an industry-driven international not-for-profit organization that grew over the years to over 220 members all over Europe, with a well-balanced composition of large-, small-, and medium-sized industries as well as research and user organizations. BDVA has over 25 working groups organized in Task Forces and subgroups, tackling all the technical and nontechnical challenges of Big Data value.

BDVA served as the private counterpart to the European Commission to implement the Big Data Value PPP program. BDVA and the Big Data Value PPP pursued a common shared vision of positioning Europe as the world leader in creating Big Data value. BDVA is also a private member of the EuroHPC Joint Undertaking and one of the leading promoters and driving forces of the AI, Data, and Robotics Partnership planned for the next framework program MFF 2021–2027.

The mission of the BDVA was “to develop the Innovation Ecosystem that will enable the data-driven digital transformation in Europe delivering maximum economic and societal benefit, and, to achieve and to sustain Europe’s leadership on Big Data Value creation and Artificial Intelligence.” BDVA enabled existing regional multi-partner cooperation to collaborate at the European level by providing tools and knowledge to support the co-creation, development, and experimentation of pan-European data-driven applications and services and knowledge exchange. The BDVA developed a joint Strategic Research and Innovation Agenda (SRIA) on Big Data Value [22]. Initially, it was fed by a collection of technical papers and roadmaps [23] and extended with a public consultation that included hundreds of additional stakeholders representing the supply and demand sides. The BDV SRIA defined the overall goals, main technical and non-technical priorities, and a research and innovation roadmap for the BDV PPP. In addition, the SRIA set out the strategic importance of Big Data; described the data value chain and the central role of ecosystems; detailed a vision for Big Data value in Europe in 2020; analyzed the associated strengths, weaknesses, opportunities, and threats; and set out the objectives and goals to be accomplished by the BDV PPP within the European research and innovation landscape of Horizon 2020 and at national and regional levels.

### ***4.3 Data Platform Project Portfolio***

The data platform projects running under the Big Data Value PPP umbrella develop integrated technology solutions for data collection, sharing, integration, and exploitation to create such a European data market and economy [22]. The Big Data Value PPP portfolio covers the data platform projects shown in Table 2. This table gives an overview of these projects, the type of data platform they develop and the domain, respectively, and the use cases they address. These projects are briefly summarized below based on open data from <https://cordis.europa.eu/>.

**Table 2** Portfolio of the Big Data Value PPP covering data platforms

Chapter	Project name	Type	Domains/use cases
	BD4NRG: Big Data for Next Generation Energy	Industrial	Energy
	BD4OPEM: Big Data for OPen innovation Energy Marketplace	Industrial and personal	Energy
Ch 3	DataPorts: A Data Platform for the Cognitive Ports of the Future	Industrial	Transport and logistics
Ch 3	DataVaults: Persistent Personal Data Vaults Empowering a Secure and Privacy-Preserving Data Storage, Analysis, Sharing and Monetisation Platform	Personal	Sports Mobility Healthcare Smart home Tourism
Ch 3, Ch 7	i3-Market: Intelligent, Interoperable, Integrative and deployable open-source marketplace with trusted and secure software tools for incentivising the industry data economy	Industrial	Automotive Manufacturing Healthcare
Ch 3, Ch 6	KRAKEN: Brokerage and market platform for personal data	Personal	Education Health
	MOSAICROWN: Multi-Owner data Sharing for Analytics and Integration respecting Confidentiality and Owner control	Personal	Connected vehicles, finance, marketing
Ch 5	MUSKETEER: Machine learning to augment shared knowledge in federated privacy-preserving scenarios	Industrial and personal	Smart manufacturing Healthcare
Ch 3	OpertusMundi: A Single Digital Market for Industrial Geospatial Data Assets	Industrial	Geospatial
Ch 3	PIMCITY: Building the next generation personal data platforms	Personal	Generic
	PLATOON: Digital PLATform and analytic TOOLS for eNergy	Industrial and personal	Energy
	Safe-DEED: Safe Data-Enabled Economic Development	Industrial and personal	Marketing Manufacturing
Ch 3	SmashHit: Smart dispatcher for secure and controlled sharing of distributed personal and industrial data	Industrial and personal	Connected cars Smart cities
Ch 12	SYNERGY: Big Energy Data Value Creation within SYNergetic enERGY-as-a-service Applications through trusted multi-party data sharing over an AI Big Data analytics marketplace.	Industrial	Energy
Ch 3, Ch 10	TheFSM: The Food Safety Market: an SME-powered industrial data platform to boost the competitiveness of European food certification	Industrial	Food supply chain
Ch 3	TRUSTS: Trusted Secure Data Sharing Space	Industrial and personal	Finance Telecom

## 5 Book Overview

This book captures the early lessons and experience in creating Data Spaces. The book arranges these contributions into three parts (see Fig. 2) covering Part I) design, Part II) deployment, and Part III) future directions.

- The first part of the book explores the design space of Data Spaces. Then, the chapters detail organizational design for Data Spaces, data platforms, Data Governance federated learning, personal data sharing, data marketplaces, and hybrid AI for Data Spaces.
- The second part of the book explores the use of Data Spaces within real-world deployments. The chapters include case studies of Data Spaces in sectors including Industry 4.0, food safety, FinTech, health care, and energy.
- The third and final part of the book details future directions for Data Spaces, including challenges and opportunities for common European Data Spaces and privacy-preserving techniques for trustworthy data sharing.

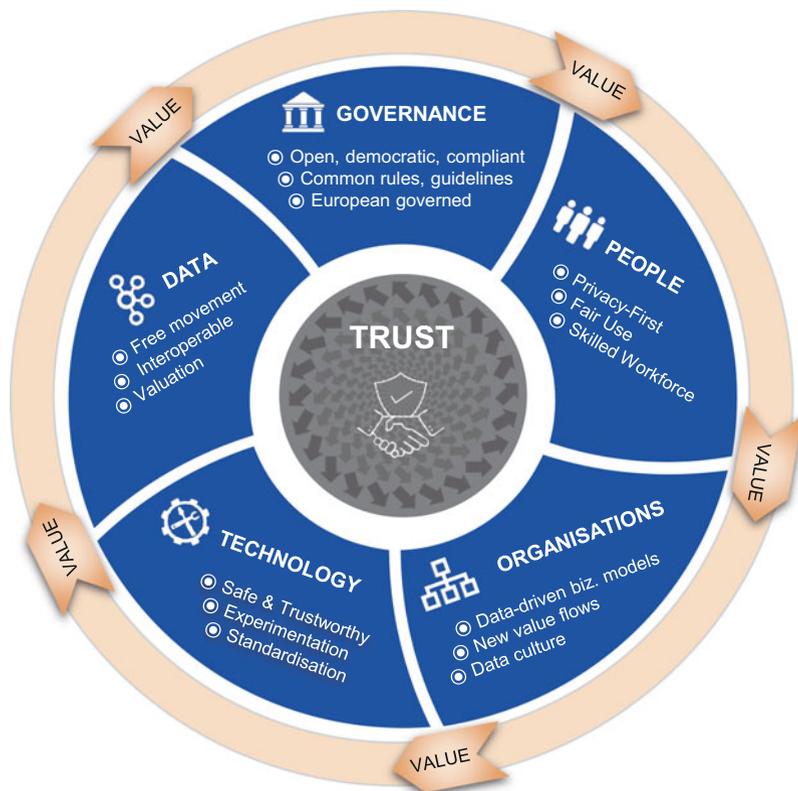
### 5.1 Chapter Analysis

As depicted in Fig. 3, the success of widespread data sharing activities revolves around the central key concept of trust: in the validity of the data itself and the algorithms operating on it, in the entities governing the data space, in its enabling technologies, as well as in and among its wide variety of users (organizations and private individuals as data producers, consumers, or intermediaries). To achieve the required levels of trust, each of the following five pillars must meet some of the necessary conditions:

- Organizations—More organizations (including business, research, and governmental) need to rethink their strategy to fully embrace a data culture that places data at the center of their value proposition, exploring new data-driven business models and exploiting new data value flows.
- Data—As a touted fifth European fundamental freedom, free movement of data relies on organizational data strategies that embed methodologies for data sharing by-design (e.g., interoperability) and clear standard guidelines that help determine the market value of data assets.
- Technology—Safer experimentation environments are needed to catalyze the maturation of relevant technology behind trustworthy data, data access, and algorithms (privacy, interoperability, security, and quality). In addition, standardization activities need to adjust for faster reaction times to emerging standards and the identification of new ones.
- People—Data sharing needs to guarantee individual privacy and offer fair value or compensation of shared personal data. For Europe to drive data sharing

<b>Part I: Design</b>						
Chapter 2 Organizational Maturity Model for Data Spaces	Chapter 3 Data Platforms for Data Spaces	Chapter 4 Data governance, implementation in Big Data environments	Chapter 5 Increasing trust within an ecosystem with Federated learning	Chapter 6 Secure, Trusted, Regulatory-Compliant, Privacy-Preserving Data Platform	Chapter 7 Connecting Data Spaces and Data Marketplaces	Chapter 8 Hybrid AI for Data Platforms
<b>Part II: Deployments</b>						
Chapter 9 A Digital Twin Platform for Industrie 4.0	Chapter 10 A framework for digital sovereignty: The European Industrial Data Space	Chapter 11 Deploying Scalable Big Data Platform to Enable a Food Safety Data Space	Chapter 12 Data Space Best Practices for Data Interoperability in FinTechs	Chapter 13 Trusted Integrated Knowledge Dataspaces for Sensitive Data Sharing & Collaboration	Chapter 14 Towards an Energy Data Platform Reference Architecture	
<b>Part III: Future Directions</b>						
Chapter 15 Common European Data Spaces: Challenges and Opportunities			Chapter 16 Privacy preserving techniques for trustworthy data sharing: Opportunities and Challenges for future research			

Fig. 2 Structure of the book



**Fig. 3** The data sharing value “wheel”—core pillars and principles of the envisioned European-governed data sharing space that generate value for all sectors of society

activities, the European workforce needs appropriate reskilling and upskilling to meet the evolving needs of the labor market.

- **Governance**—A European-governed data sharing space can inspire trust by adhering to the more advanced European rules, guidelines, and regulations and promoting European values. Participation should be equally open to all and subject to transparent and fair rules of conduct.

Table 3 gives an overview to which extent the contributions described in the different chapters of this book contribute to the different dimensions of the data sharing wheel.

As this table indicates, the chapters in this book provide broad coverage of the pillars of the data sharing wheel, reinforcing the relevance of these concerns.

The majority of the chapters cover the value, data, and technology pillars of the wheel which illustrate the data-driven focus of the works in Data Spaces. Governance and trust are also well covered, highlighting the importance of these pillars to the deployment and operation of Data Spaces. While organization aspects

**Table 3** Coverage of the pillars of the data sharing wheel by the book chapters

Chapter	Value	Data	Technology	Organization	People	Governance	Trust
Part I: Design							
Ch 2	X	X	X	X	X	X	X
Ch 3		X	X	X	X	X	X
Ch 4	X	X	X	X		X	
Ch 5	X	X	X			X	X
Ch 6		X	X	X		X	X
Ch 7	X	X	X				X
Ch 8	X	X	X	X		X	
Part II: Deployment							
Ch 9	X	X	X	X			
Ch 10	X	X		X		X	X
Ch 11	X	X	X				X
Ch 12	X	X	X				
Ch 13		X	X			X	X
Ch 14	X	X	X	X			X
Part III: Future directions							
Ch 15		X	X	X		X	X
Ch 16	X	X	X	X	X	X	

are well covered with an understanding of how organizations can leverage the benefits of Data Spaces to transform their business models and operations, there is a paucity of work in the area of the people pillar. Skills, reskilling, and upskilling to meet the emerging needs of Data Spaces and society.

## 6 Summary

We are now seeing digital transformation toward global digital markets and Data Spaces. However, this will not be a fast transition and may take a decade before we understand the methods and the means of mature Data Spaces. In comparison, the World Wide Web as a mature platform for trade took from the mid-1990s to well beyond 2000 before it became an everyday tool to search for information and order weekly groceries.

As this development is systemic, it requires scientific, technical, and social foundations. This book addresses and crystallizes the developments of many efforts to establish Data Spaces and learnings from the efforts. Data Spaces are feature-rich technical constructs within a social and regulation framework that support data ecosystems with fair and trusted approaches to share data. It is an ambitious goal, and therefore data ecosystems present new challenges to the design of data sharing. We need to rethink how we should deal with the needs of these large-scale data-rich environments with multiple participants. This chapter gave the

foundations of the concepts, but obviously, many challenges exist. A data space can provide a clear framework to support data sharing within a data ecosystem. This book is a step toward such a framework by delineating real experiences from pilots and experiments. The book addresses and explores the cutting-edge theory, technologies, methodologies, and best practices for Data Spaces for both industrial and personal data. It provides the reader with a basis for understanding the scientific foundation of Data Spaces, how they can be designed and deployed, and future directions.

The development of data space technology is societal. For example, the development of electricity networks required agreements on a common approach for the electricity grid. In the same manner, common agreements are needed from large and small industry, policymakers, educators, researchers, and society at large to create the basis of the data economy and common European Data Spaces. The BDV PPP has advanced the value of Big Data and AI, laying the basis for new combinations of technologies that will go beyond the digital market toward a new and productive Digital Society.

## References

1. Curry, E. (2016). The big data value chain: Definitions, concepts, and theoretical approaches. In J. M. Cavanillas, E. Curry, & W. Wahlster (Eds.), *New horizons for a data-driven economy* (pp. 29–37). Springer International Publishing. [https://doi.org/10.1007/978-3-319-21569-3\\_3](https://doi.org/10.1007/978-3-319-21569-3_3)
2. Curry, E., Hasan, S., Kouroupetroglou, C., Fabritius, W., ul Hassan, U., & Derguech, W. (2018). Internet of Things enhanced user experience for smart water and energy management. *IEEE Internet Computing*, 22(1), 18–28. <https://doi.org/10.1109/MIC.2018.011581514>
3. Curry, E., & Sheth, A. (2018). Next-generation smart environments: from system of systems to data ecosystems. *IEEE Intelligent Systems*, 33(3), 69–76. <https://doi.org/10.1109/MIS.2018.033001418>
4. Franklin, M., Halevy, A., & Maier, D. (2005). From databases to dataspace: a new abstraction for information management. *ACM SIGMOD Record*, 34(4), 27–33. <https://doi.org/10.1145/1107499.1107502>
5. Curry, E. (2020). Fundamentals of real-time linked dataspace. In *Real-time linked dataspace* (pp. 63–80). Springer International Publishing. [https://doi.org/10.1007/978-3-030-29665-0\\_4](https://doi.org/10.1007/978-3-030-29665-0_4)
6. Halevy, A., Franklin, M., & Maier, D. (2006). Principles of dataspace systems. In *25th ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems - PODS '06* (pp. 1–9). ACM Press. <https://doi.org/10.1145/1142351.1142352>
7. Nagel, L. (2021). *Design principles for data spaces*.
8. Singh, M. (2013). A framework for data modeling and querying dataspace systems. In *Seventh International Conference on Data Mining and Warehousing (ICDMW)*.
9. Hedeler, C., Belhajjame, K., Paton, N. W., Campi, A., Fernandes, A. A. A., & Embury, S. M. (2010). Dataspace. In S. Ceri & M. Brambilla (Eds.), *Search computing* (pp. 114–134). Springer. [https://doi.org/10.1007/978-3-642-12310-8\\_7](https://doi.org/10.1007/978-3-642-12310-8_7)
10. Hedeler, C., Belhajjame, K., Paton, N. W., Fernandes, A. A. A., Embury, S. M., Mao, L., & Guo, C. (2011). Pay-as-you-go mapping selection in dataspace. In *Proceedings of the 2011 international conference on Management of data - SIGMOD '11* (pp. 1279–1282). ACM Press. doi:<https://doi.org/10.1145/1989323.1989476>.
11. Wang, Y., Song, S., & Chen, L. (2016). A survey on accessing dataspace. *ACM SIGMOD Record*, 45(2), 33–44. <https://doi.org/10.1145/3003665.3003672>

12. Das Sarma, A., Dong, X., & Halevy, A. Y. (2009). Data modelling in dataspace support platforms. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 5600 LNCS, pp. 122–138). doi:[https://doi.org/10.1007/978-3-642-02463-4\\_8](https://doi.org/10.1007/978-3-642-02463-4_8).
13. Singh, M., & Jain, S. K. (2011). A survey on dataspace. In *Communications in Computer and Information Science* (pp. 608–621). doi:[https://doi.org/10.1007/978-3-642-22540-6\\_59](https://doi.org/10.1007/978-3-642-22540-6_59).
14. Curry, E. (2020). *Real-time linked dataspace*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-29665-0>
15. Curry, E., Derguech, W., Hasan, S., Kouroupetroglou, C., & ul Hassan, U. (2019). A real-time linked dataspace for the Internet of Things: Enabling “pay-as-you-go” data management in smart environments. *Future Generation Computer Systems*, 90, 405–422. <https://doi.org/10.1016/j.future.2018.07.019>
16. Scerri, S., Tuikka, T., & Lopez de Vallejoan, I. (Eds.). (2020). Towards a European data sharing space.
17. Zillner, S., Bisset, D., Milano, M., Curry, E., Hahn, T., Lafrenz, R., et al. (2020). *Strategic research, innovation and deployment agenda - AI, data and robotics partnership. Third Release (Third)*. Brussels: BDVA, euRobotics, ELLIS, EurAI and CLAIRE.
18. Gawer, A., & Cusumano, M. A. (2014). Industry platforms and ecosystem innovation. *Journal of Product Innovation Management*, 31(3), 417–433. <https://doi.org/10.1111/jpim.12105>
19. Curry, E., & Ojo, A. (2020). Enabling knowledge flows in an intelligent systems data ecosystem. In *Real-time linked dataspace* (pp. 15–43). : Springer International Publishing. doi:[https://doi.org/10.1007/978-3-030-29665-0\\_2](https://doi.org/10.1007/978-3-030-29665-0_2).
20. Curry, E., Metzger, A., Zillner, S., Pazzaglia, J.-C., & García Robles, A. (2021). *The elements of big data value*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-68176-0>
21. Zillner, S., Gomez, J. A., García Robles, A., Hahn, T., Le Bars, L., Petkovic, M., & Curry, E. (2021). Data economy 2.0: From big data value to AI value and a european data space. In E. Curry, A. Metzger, S. Zillner, J.-C. Pazzaglia, & A. García Robles (Eds.), *The elements of big data value* (pp. 379–399). Springer International Publishing. [https://doi.org/10.1007/978-3-030-68176-0\\_16](https://doi.org/10.1007/978-3-030-68176-0_16)
22. Zillner, S., Curry, E., Metzger, A., Auer, S., & Seidl, R. (Eds.). (2017). *European big data value strategic research & innovation agenda*. Big Data Value Association. Retrieved from [http://www.edwardcurry.org/publications/BDVA\\_SRIA\\_v4\\_Ed1.1.pdf](http://www.edwardcurry.org/publications/BDVA_SRIA_v4_Ed1.1.pdf)
23. Cavanillas, J. M., Curry, E., & Wahlster, W. (2016). *New horizons for a data-driven economy: A roadmap for usage and exploitation of big data in Europe*. Springer Nature PP. <https://doi.org/10.1007/978-3-319-21569-3>

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

