Fraunhofer
ISST

Fraunhofer Institute for Software and Systems Engineering ISST

Trust-based Data Ecosystems

Seeding Data Value Creation

Executive Summary

Data is increasingly recognized as a valuable asset for organizations, with its volume growing rapidly every year. Data presents numerous opportunities to enhance business processes and develop innovative business models, thereby increasing competitiveness and improving overall business operations. However, data must be actively shared, used, and analyzed to unlock its full potential, in contrast to the traditional view of data as the new oil. Unlike oil, data does not have value when it is at rest; it needs to be put in motion and used.

Trust for risk mitigation is a key aspect of data sharing, being essential for creating value-driven data ecosystems. Trust protocols and governance structures enable secure and efficient data sharing across industries. With that, trust reduces risks and enhances the willingness to share sensitive data, accelerating innovation and collaboration.

At present, the first data ecosystems are already operational. At the same time, many more are being developed and designed by alliances of various industries and sizes. These developments are accompanied by new technological advances, standardization efforts, and even regulatory measures. They offer the opportunity to create a valuable nurturing ground but need time to grow organically.

Our vision for the future of data ecosystems involves embracing self-organization, autonomy, and trust as enablers of innovation. Instead of enforcing strict control, policymakers should create conditions that allow innovation to develop organically in constructive chaos. Constructive chaos builds on a targeted initial impulse and is then allowed to emerge freely. This development requires a balance between structured design and fostering self-organization and decentralized growth. For data spaces, we propose achieving this balance by establishing comprehensive communication and trust protocols, as well as political measures, to allow for the evolution of yet unforeseen applications and data usage patterns.

Authors

Julia Pampus

Anna Maria Schleimer

Sebastian Emons

Anna Gieß

Boris Otto

Contact

Julia Pampus
julia.pampus@isst.fraunhofer.de

Fraunhofer Institute for Software and Systems Engineering ISST
Speicherstraße 6
44147 Dortmund, Germany
www.isst.fraunhofer.de

Contents

1. Data – From Commodity to Value

Data is becoming an increasingly important commodity for organizations. At the same time, the volume of data is rising each year, with growth rates in the double-digit range. For example, between 2022 and 2023, the amount of newly created or replicated data increased from 103.66 to 132.4 zettabytes.[1] This increase in data volume presents organizations with a wide range of opportunities for value creation, both by improving existing processes and by developing new or enhanced business models. These advances allow organizations to increase their competitiveness and overall business performance but also serve to improve the common good.[2]

1.1. Towards Creating Value from Data

Data at rest poses a significant negative value for organizations, as it incurs various costs, including those associated with collection, storage, and protection. Unlike gold, which has intrinsic value, data does not have any worth when it is not being utilized. To create actual value, data must, therefore, be put into motion through sharing and usage (see Figure 1). This view contradicts earlier descriptions of data as new oil or gold. Data is not equivalent to valuable storage or productive assets and does not lose its value when shared with others. This understanding is crucial for future advancements in the field of artificial intelligence (AI). For example, many organizations have recently utilized freely accessible data published on the World Wide Web to train large language models (LLMs). Next, non-publicly available data would need to be made available for specific applications to improve the outputs of LLMs further using AI. In this context, sharing data is essential. Furthermore, access to markets is linked to data. Regulatory requirements for organizations, such as reporting obligations, are becoming increasingly strict, which requires the collection and use of data. This trend can be observed across various regions, including the European Union (EU), the United States (US), and China.

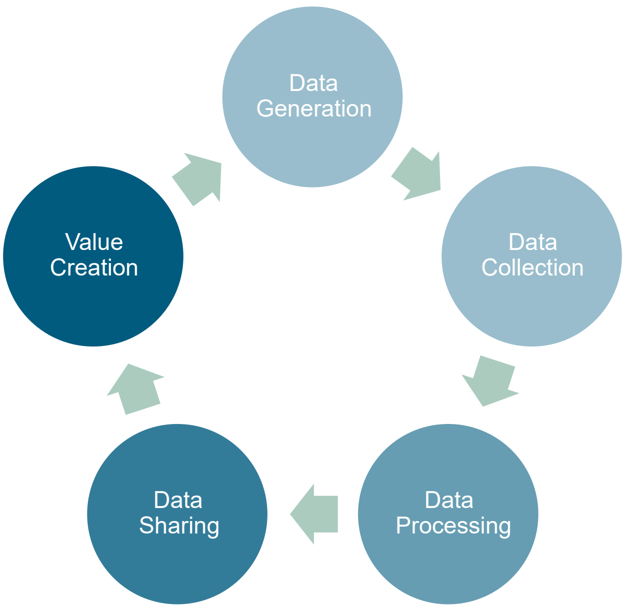

Figure 1: Continuous Process of Data Value Creation

The generation of data is a prerequisite for creating new value in the future, leading to two main pathways: First, organization-internal usage enables improvements, for example, in processes. Second, data can be made available to other organizations within data ecosystems, either in exchange for other data or for a fee. This opens numerous opportunities, including enhanced processes and the development of entirely new business models. Previous initiatives have advanced the process of value creation by increasingly collecting and processing data. For example, in the manufacturing industry, about Industry 4.0, this advancement has resulted in a growing number of sensors being used in connected production and products.

1.2. Risk versus Reward

Although data sharing creates new business opportunities and value constellations based on existing processes, it requires constantly weighing the benefits against potential risks. The more data is shared and the more stakeholders it is shared with, the less control the data rights holder has over the actual use of the data. This loss of control increases the likelihood of potential misuse. A common risk is data breaches, which can occur due to unauthorized access to sensitive information. Such incidents can lead to significant security issues and, in service-oriented offerings, result in a loss of customer loyalty and trust, potentially causing long-term reputational damage to organizations. Furthermore, compliance issues (failing to adhere to regulations) may involve legal and financial penalties.

Due to the broad potential of data sharing, organizations must find a balance between profiting from it and minimizing associated risks. First, risks must be identified. Second, measures must be taken to mitigate these risks. In the context of data sharing, it is essential to ensure that all related actions and processes take place in a controlled environment. This requires common standards at the level of data (structure or quality), data transfer, and data processing. In addition, technical measures may support application-specific data sharing by, for example, defining the context (local and temporal) of data usage and making data flows transparent.

Trust serves as a means of reducing risks to an acceptable level, thereby playing an essential role in the triangle of reward, risk, and trust (see Figure 2). The higher the trust in a data sharing party, the greater the willingness to share data of higher sensitivity and in larger quantities.[3] Ultimately, data is shared at the speed of trust. Trust ranges from an organization being confident that another organization is what it claims to be *(verifiable identities)* to compliance with restrictions on data use for specific applications *(controlled data usage)* and to ensuring that data is valid and has not been manipulated *(data integrity).* Therefore, trust is a fundamental requirement for value constellations that are driven by data sharing. Without trust, data value creation will always be limited.

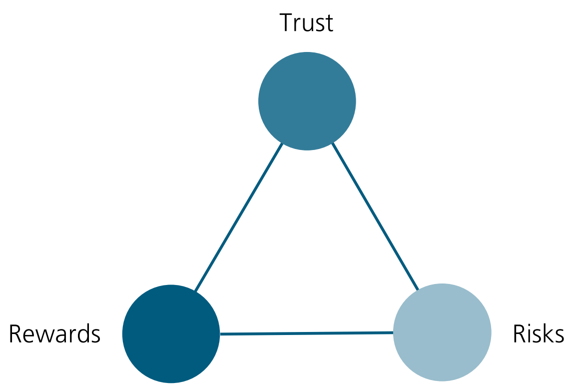

Figure 2: Balanced Triangle of Rewards, Risks, and Trust

As illustrated in Figure 3, various mechanisms (regulatory, organizational, technical) exist to strengthen trust in data sharing. First, as part of *regulatory measures,* trust is established through legal frameworks. Second, *organizational measures,* including internal policies, compliance programs, and training, play a crucial role in building trust. Third, *technical measures* include, for example, ensuring transparency and enforcing data usage decisions, i.e., policies. Organizations will exchange claims and evidence that must align with the policies of the other party until they estimate that they have accumulated sufficient evidence to make an informed decision about trust. These technical approaches can be enhanced through automation and certification processes.

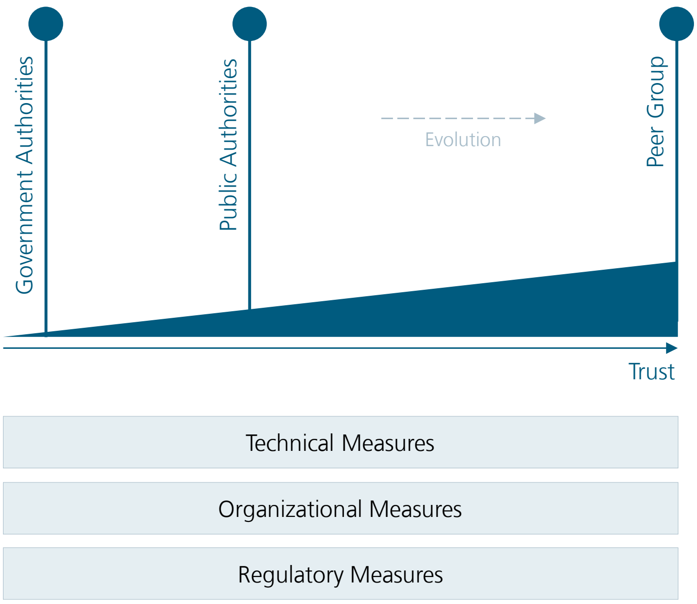

Figure 3: Gradual Building of Trust

Most importantly, trust is built gradually (see Figure 3). In the initial stage, organizations are generally cautious and do not trust anyone. Here, the state plays a central role by establishing the foundation for regulatory measures through laws and regulations, serving as officially approved roots of trust chains, i.e., trust anchors. As part of the trust-building process, the trust context of organizations can be expanded to include public authorities, such as public institutions, which serve as trust anchors. The highest level of trust is achieved when mutual trust is established within a peer group. Then, trust is built through direct interactions and shared experiences, creating a powerful form of trust that is bound to a network of trust anchors, without further reliance on central authorities. However, this is also the most challenging level of trust to achieve.

2. Trust for Data Sharing

Trust is the key element in collaboration. As outlined in Section 2, in the context of data sharing, it is essential to build trust between individual stakeholders and to support this building process through regulatory, organizational, and technical means. A pattern for creating trusted environments for data sharing is the concept of *data ecosystems*.

2.1. Data Ecosystems

Ecosystems can be found in many areas of nature. They are self-regulating systems in which everyone has their role to play (rights and duties), and a sustainable living environment is created. In ecosystems, organisms interact in symbiosis, creating an ecologically stable equilibrium.[4] This concept from nature has been applied to the economy to a limited extent so far. Currently, ecosystems mainly exist in two forms: (1) a group of a limited number of organizations that sign *a partnership agreement* to jointly offer customers a better solution than they could offer on their own; (2) an organization that provides a set of services along with *a service agreement* that creates a supporting ecosystem for the customer. Furthermore, many ecosystems evolve around centralized platforms controlled by organizations that benefit from both their competitors and customers. Typically, such partnership agreements or service contracts outline clear requirements regarding architecture, governance, and other guidelines that contradict natural ecosystems. They are designed for stable, long-term business relationships and close integration of information and communication systems.

Data ecosystems aim to achieve something different: Organizations in data ecosystems are loosely connected and can flexibly establish data sharing activities – and repeatedly adjust them according to their current business needs.[5] The autonomy and self-determination of the parties involved characterize this type of flexibility in data sharing. This is also reflected in the way agreements are made. All data sharing scenarios, using scalable and interoperable infrastructures, are governed by trust and build the foundation for value creation (see Figure 4).

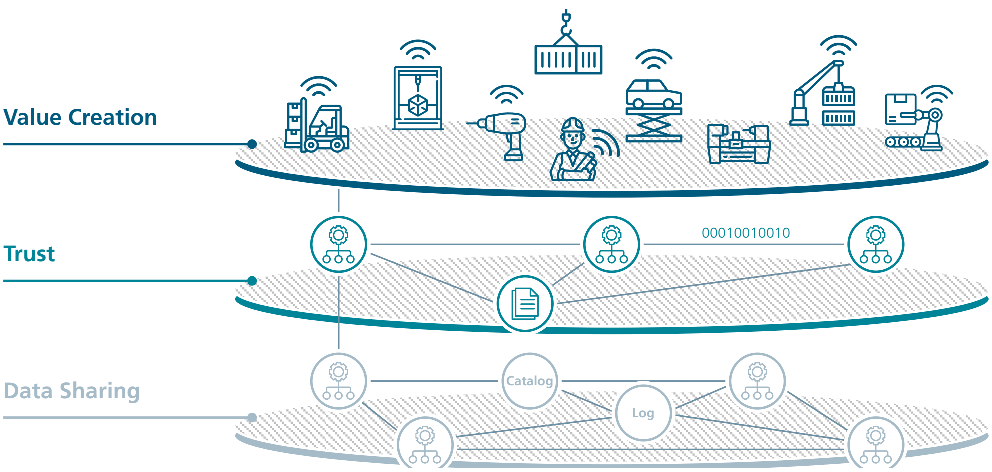

Figure 4: Layered Model of Data Sharing, Trust, and Value Creation

The vision of trust-based data ecosystems encompasses not just collaboration between a few organizations but entire industries, from raw material suppliers to recycling organizations, working together in innovative and new ways. These collaborations enable the exchange of data for other data, financial compensation, or other benefits. The distributed power of data rights owners and users creates independent balances and new environments for value creation. However, data ecosystems are not yet as advanced as natural ecosystems. To participate in data ecosystems and achieve added value that exceeds initial investments, stakeholders require different types of support. It is not sufficient to have access to advanced technologies; further measures and incentives are needed.

2.2. Data Spaces as Design Principle

One design principle for implementing data sharing infrastructures in data ecosystems is the concept of data spaces. Data spaces serve as a mechanism for establishing trust contexts for data sharing between organizations. They consist of a combination of legal, economic, and technical components, creating "a decentralized, neutral framework of protocols and frameworks"[6]. This framework is technology-agnostic and allows for flexibility in implementation by building upon established international standards. With this, data spaces aim to facilitate interoperability between various systems, enabling seamless data flows among organizations. These data flows are not restricted to data space technologies; they can occur through peer-to-peer channels. This flexibility opens up a wide range of applications, including various architectural patterns and processes that involve both humans and machines.

A crucial aspect of building trust contexts and ensuring the autonomy of organizations participating in data spaces is the ability to define policies regarding how data can be shared and used. Policy definition is a powerful and flexible tool that enables organizations to express requirements in the form of constraints, conditions, and obligations. By defining and negotiating these policies, organizations can maintain autonomy by deciding with whom to share their data. This process also enhances their accountability, particularly in meeting the expectations of their data sharing partners, i.e., enforcing the negotiated policies. With a focus on decentralization and neutrality (both technically and governmentally), data spaces avoid centralized services, such as intermediaries, by design. Each organization independently decides whom to trust and implements the necessary processes and mechanisms to verify the reliability of other parties.

2.3. Value Creation in Practice

The value of data spaces is already being proven in practical applications. The following projects demonstrate how concepts are translated into actionable steps that deliver tangible benefits for organizations.

2.3.1. Enhancing Automotive Supply Chain Efficiency

Catena-X[7] is an open data ecosystem that connects organizations across the entire automotive industry's value chain. It provides standardized processes and enables use cases that require data sharing across organization boundaries. Current use cases include product carbon footprint, battery and product passports, resilient supply chains, parts traceability, and master data management.

The *Predictive Unit Real-Time Information Service (PURIS)* is a use case in the areas of supply chain management and quality. It standardizes the exchange of partner-specific key supply chain information in the short term (up to four weeks) between suppliers and customers (one-up/one-down principle). In the past, data along the value chain was only exchanged in specific scenarios based on standards established by the German Association of the Automotive Industry[8], which enabled bilateral data exchange within a limited environment. With PURIS, organizations are motivated to share additional data with multiple partners, including the latest stock information, demand, and planned production volumes. By sharing data, partner organizations gain insight into current stock availability, the availability of data for short-term material requirements, expected production volumes, delivery information, and details on delivery times. With this, PURIS creates added value by improving efficiency in the supply chain, making more accurate data available across organization boundaries and eliminating and potentially preventing costly bottlenecks.

2.3.2. Advancing Collaborative Engineering in Manufacturing

Factory-X[9] is an open data ecosystem for factory outfitters and operators. It builds on the standards developed in Catena-X and the principles developed in "Plattform Industrie 4.0"[10]. It provides the groundwork for the manufacturing industry to share data across organization boundaries horizontally and vertically.

The Factory-X ecosystem supports eleven key use cases, each enhancing existing supply chain-focused solutions by extending them to the shop floor. One key use case is *Collaborative Engineering,* which demonstrates how cross-organizational, real-time collaboration in engineering can drive innovation and efficiency in the manufacturing sector. The use case enables real-time, cross-manufacturer-connected product development. In this context, information on machines and systems, such as technical specifications, operating data, or maintenance manuals, is shared automatically and securely via standardized business-to-business interfaces. This reduces manual effort, minimizes errors, and accelerates innovation cycles. Using uniform data models and standards (e.g., Asset Administration Shell[11]) enables consistent information logistics across the entire life cycle – from design to commissioning and service. Changes to data can be tracked in real-time and are shared with all involved partner organizations.[12]

2.3.3. Integrating Transport and Tourism for Future Mobility Solutions

EONA-X[13] is an open data ecosystem in transport, mobility, logistics and tourism, connecting services across these sectors. Its goal is to connect different means of transportation for holistic planning, targeting business-to-business and business-to-customer interactions. A prominent use case in this context was implemented for the organization of the Olympic Games.

The EONA-X ecosystem emphasizes the importance of data sharing in delivering a seamless travel experience for participants in large events while ensuring security and enhancing infrastructure efficiency. By utilizing digital services in conjunction with shared data, relevant information can be linked throughout the entire travel value chain, enabling effective monitoring and response to real-time needs. For example, during the Olympics at the Paris Airport, EONA-X facilitated the sharing of passenger data. With the help of this data, the airport can track the movements of delegations in real-time, providing insights into their locations and schedules. As a result, additional services related to mobility and tourism were developed. In the area of mobility, EONA-X enabled organizations to manage fleets of vehicles and coordinate their movements, improving both traffic flow and safety within the city. In addition, the tourist experience was enriched by ensuring a smooth journey from arrival to sightseeing and, ultimately, to departure. This improvement was achieved through data sharing between different service providers along the travel route, creating a reliable source of information for tourists.[14]

3. Stimulating Growth and Trust

Although first data ecosystems and data spaces that create business value are already in place, there remains a strong need to expand value creation from data. Especially in times of global economic crisis, the development of a data economy cannot be taken for granted. At the same time, it is becoming increasingly important to maintain economic growth and resilience. While leveraging data offers many economic benefits, creating data value and implementing targeted measures requires a shift in thinking.

3.1. Evolving the Data Economy

Numerous measures and initiatives have been started to leverage the enormous potential for value creation from data and data sharing. These initiatives are characterized by their scope and ambitious goals, as well as by their complexity. The landscape of data initiatives involves various stakeholders, technology providers, user groups, technologies, and standards. This raises the question of how to manage such high complexity effectively and whether it is even possible or sensible to do so. What measures and processes are necessary to act purposefully and efficiently to unlock significant benefits? Considering complex systems and the creative potential of chaos provides valuable insights into data value creation and the development of data ecosystems. This perspective highlights the need for a balance between control, thoughtful design and engineering tasks, and organic evolution that enables rapid and dynamic development.

3.1.1. Leveraging Chaos and Complexity for Innovation

A single innovation may have a selective effect or trigger changes in a particular area. Yet, wide-ranging transformations and significant leaps in innovation require an impact on an entire system. This perspective on an overall system, such as the data economy, involves considering individual system components, including selected technical building blocks, individual platforms, or even individual applications, as well as the parts of a system as a whole. The focus on the whole rather than on the parts is at the heart of systems theory[15] and is also a key requirement in research on information systems in today's complex ecosystems.[16]

Some technological evolutions may naturally unfold and become established over time, much like the concept of a balanced market or a healthy ecosystem in nature. However, with many current issues, waiting is not an option, as changes are urgently needed or will not happen without intervention. This also applies to the data economy. The common problems of not utilizing available data, inefficient systems and processes, or missed opportunities lead to the overall data economy being in a state that does not allow for further evolution or can only make progress with very low efficiency. Especially in times of global political uncertainty and economic crises, organizations cannot afford a slow evolution and continuing underutilization of data. Intentional efforts are required to transform a singular innovation into widespread systemic change, encouraging technological advancement and facilitating the emergence of a wide range of innovations.

To achieve a broad impact, developments with far-reaching effects must be initiated rather than implemented in a thorough and controlled manner. A strictly controlled approach would not be possible due to the high complexity of the data landscape, given the many areas of application and systems involved. Moreover, such centrally controlled evolution would exceed the capabilities of a single entity and its influence on the day-to-day technologies used in the economy. The complexity results from the many agents, such as system components, but also processes, people, or organizations[17]. However, this complexity is not a weakness; on the contrary, it has many advantages that can be used to foster a data economy: complex systems have many valuable properties, including the ability to self-organize and evolve. Furthermore, they can have unpredictable, far-reaching effects and explore new paths.

Explanations for evolution and the occurrence of significant changes can be found not only in complexity theory but especially in chaos theory[18]. Complexity theory focuses on complex systems and their behavior, specifically systems comprising a multitude of interdependent units. These units are often partially connected and heterogeneous, interacting through different feedback loops, which makes their prediction difficult. As the complexity of a system increases, self-organization is challenged, and the system enters a state of transition, also known as a state of emerging complexity (see Figure 5) or the *edge of chaos.*[19] The edge of chaos comes with the risk of escalating into an unpredictable catastrophe of complexity, causing increasing disintegration of the system. Concerns about the negative effects of complexity are reflected in the efforts of researchers and managers to mitigate its impact. Some even consider complexity as synonymous with a lack of influence on a system.

However, the edge of chaos is also a state of radical renewal, enabling significant steps towards positive change, even if these are unforeseen and may appear like a black box to observers. It can be compared to the concept of agility and enables generativity: the edge of chaos provides just enough stability for experimentation and creativity without falling into disintegration[20]. The chaos model also contributes to the unpredictability of technological development[21]. We refer to the edge of chaos as *a constructive chaotic state.* This state lies halfway between a system without complexity, which leads to a rigid system, and *dysfunctional chaos,* accompanied by disintegration and a complexity catastrophe with fatal negative consequences.

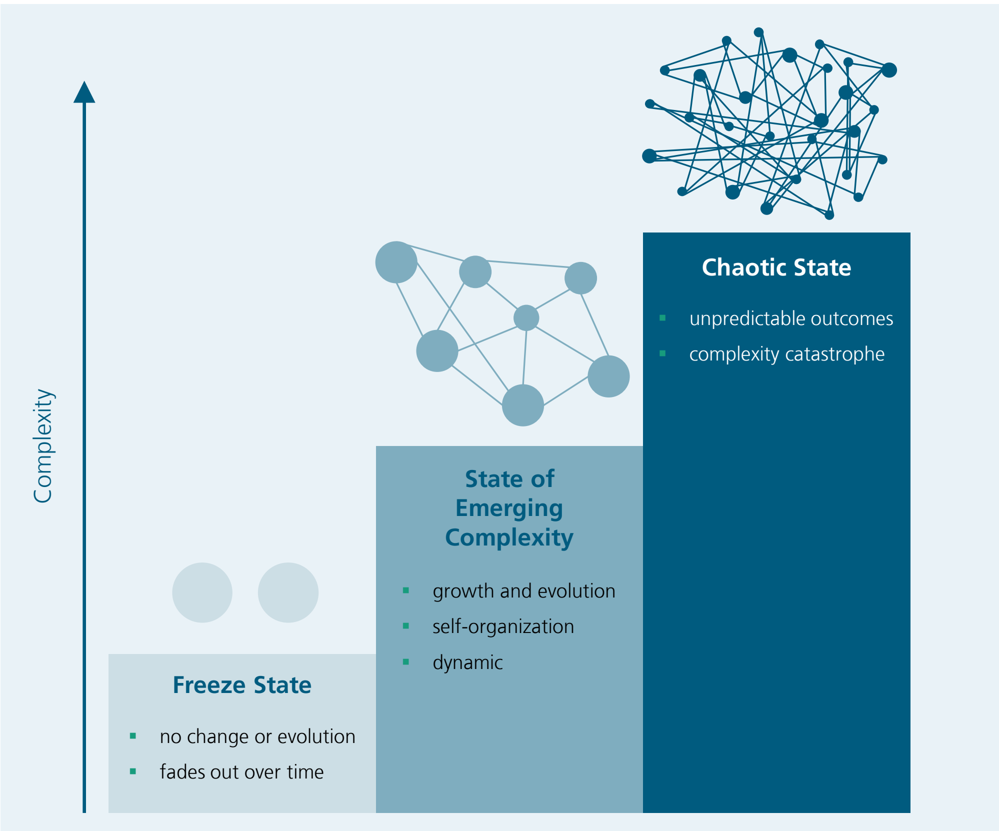

Figure 5: Complexity Stages

3.1.2. Chaos as Enabler for Data Value

Chaos theory enables us to understand and leverage the dynamics of chaotic systems and apply these insights to data value creation in trust-based data ecosystems. Most importantly, chaos is not the absence of order but rather hidden patterns. These patterns can be too complex to be understood by humans or rigidly planned. Although the theory originates in the natural sciences and was initially used to describe physical phenomena in geology, it is also applicable to information systems[22]. The butterfly effect, first researched by Lorenz in the 1960s[23], describes how small changes can have significant and unpredictable effects. It is famously illustrated by the example of a butterfly flapping its wings in Brazil, which can cause a tornado in Texas[24]. When developing a system, it is not only its initial conditions that are important but also feedback loops and attractors that either amplify or weaken the performance of a system.

Applied to the context of data value creation, this means that a well-thought-out starting condition implies further development and positive effects. Therefore, it is essential to differentiate between constructive chaos and its detrimental effects and to allow the necessary freedoms for patterns to develop. It must also be defined which initial conditions are required and what they entail, as well as identified which attractors serve as influencing forces. This approach involves a combination of controlled, well-thought-out starting conditions while also allowing for organic evolution to occur. The key is to embrace designable openness, i.e., the conscious use of uncertainty and ambiguity as productive forces. To leverage the value creation potential of data and establish trust-based data ecosystems, policymakers, standardization institutions, and governance bodies must develop a framework that fosters emergence rather than control.

History of the World Wide Web

The evolution of the World Wide Web illustrates the interplay between strictly defined initial conditions, i.e., communication protocols, and the emergence of constructive chaotic patterns.

In 1969, the Advanced Research Projects Agency Network was developed by an agency of the United States Department of Defense with the primary goal of facilitating communication and data sharing among researchers. This laid the foundation for the Transmission Control Protocol (TCP), which included concepts for internetworking, leading to the development of the Internet protocol suite, commonly referred to as TCP/IP. This protocol later became the standard protocol for Internet communication. Originating from this foundation, the network expanded, interconnected with other networks and participating organizations, and additional services such as the Domain Name System. The public access to the network, increased access to computers by businesses and private individuals, as well as commercialization options led to significant growth. The demand for a connection to the World Wide Web led to the construction of large data centers, the installation of oversea cables, and the use of telegraph lines as a connection to the World Wide Web. As a result, infrastructure providers and operators emerged, but above all, many previously unknown business models and new organizations based exclusively on information transfer via the web. Facebook, eBay, Amazon, Myspace, Etsy, and Google are just a few popular examples from the business-to-customer sector. In addition, numerous services in the business-to-business industry, as well as new information and communication service providers, now account for a substantial share of global economic output.

The example of the global Internet clearly shows how little was known about the resulting business models and their patterns when the first protocols were developed. The emergence of new business models was not controlled or planned top-down, for example, by a government or consortia. Developments may be random, highly dependent on the respective situation and not repeatable.

Originally intended as a means of exchanging information between researchers, the World Wide Web evolved over the years into applications that have transformed the economy like never before. From new Internet and platform business models to corporate data and online shops to social media, the Internet has transformed the economy and society, creating entirely new digital markets. The emergence of various previously unimaginable applications based on protocols and standards illustrates the potential that data ecosystems can have for the next step in global value creation and innovation.

3.2. Designing the Constructive Chaos

To obtain a useful mental model for designing data value creation in trust-based data ecosystems, we use chaos theory and define two complex subsystems (see Figure 6). The first covers a technical perspective (bottom-up) that considers how a set of protocols can "grow" to a wide range of *applications,* i.e., use cases. The second concept is a top-down perspective that addresses how policy strategies can determine the impact on applications.

Each pattern is divided into complexity levels: At level zero, a few strictly defined and carefully engineered items form the root cause, referred to as the initial conditions. As complexity increases, both the number of elements and the number of connections between them increase. They are not subject to detailed control but arise and develop on their own. In the mental model for trust-based data ecosystems, these items are the applications.

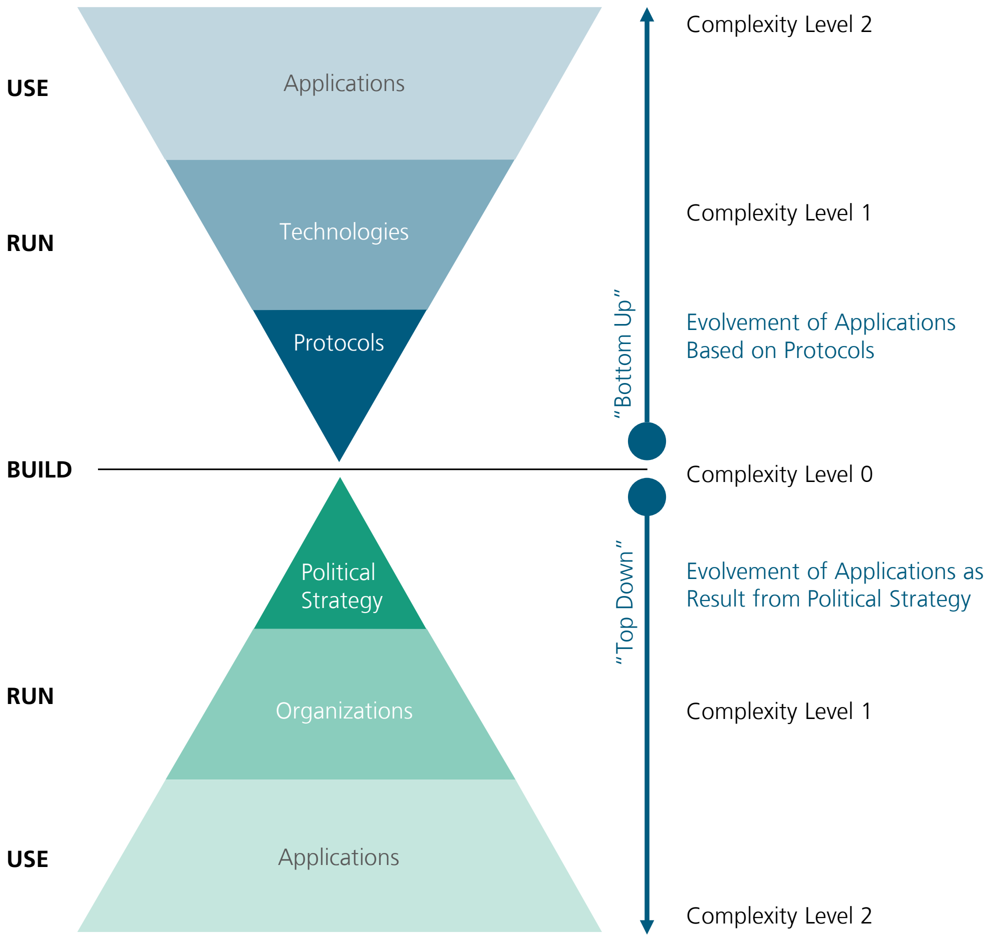

Figure 6: Overview on Complexity Levels and Perspectives

3.2.1. Protocols and Technologies

Building the technical foundation involves establishing common (trust) protocols that enable interoperability across different technologies and organizations within the same or different application domains (see Figure 7).

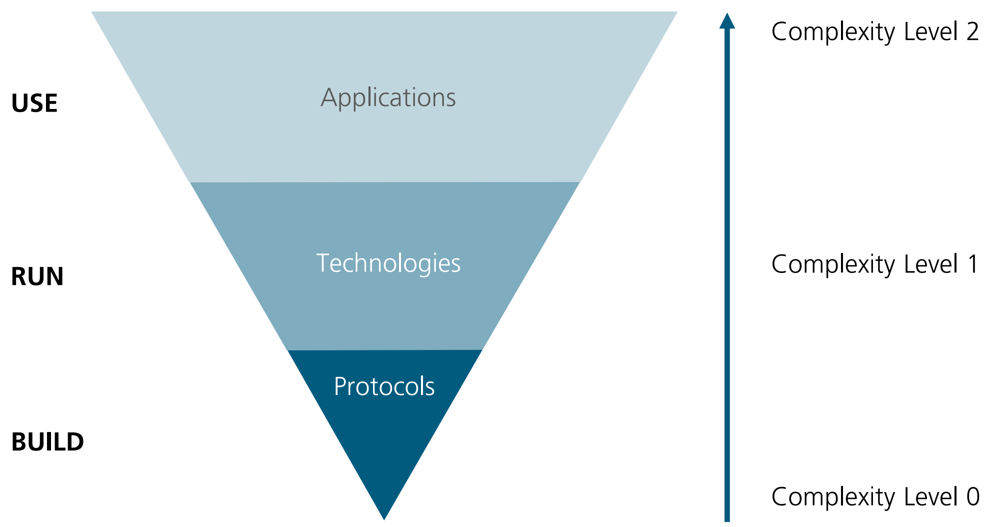

Figure 7: From Protocols to Applications

Unified protocols are crucial for maintaining stability within a system. They provide structured guidelines that create order and consistency, ensuring that all organizations adopting them follow the same rules and formats. This leads to technologies that are predictable and *interoperable*. In the context of data sharing within trust-based data ecosystems, protocols are essential for building this trust. Their open development encourages broader adoption and, over time, enhances their maturity through diverse contributions.

- **Data Offering and Discovery**
  Organizations can autonomously decide what data they want to share and consume, as well as the conditions under which it can be used, i.e., policies, by connecting with potential data sharing partners.

- **Policy Negotiation**
  Organizations can negotiate policies in a self-determined manner. They can define the attributes to be checked during policy validation when accessing services and throughout the entire data lifecycle.

- **Data Transfer Orchestration**
  Organizations can manage peer-to-peer data flows across various systems, enabling them to implement a range of applications.

- **Issuing Claims**
  Organizations can issue claims to assert the validity of their data, identity, or other statements. This process is best carried out in a decentralized manner using verifiable credentials.

The usage of *standards* is essential for creating reliable and scalable solutions. Open standards ensure that information, such as policies and identities, is consistently understood across different systems, enabling accurate and meaningful interpretation. Consequently, standards serve as the foundation for technical adoptions of trust protocols. These adoptions, i.e., technologies, can meet a wide range of requirements and vary in nature.

- They must integrate diverse data sources across platforms and domains.
- They should manage identities across multiple jurisdictions to comply with legal and regulatory requirements.
- They need to effectively process different types of data transfer, including streaming and big data, to support a wide range of applications.
- They should implement policy management across multiple infrastructure layers to ensure compliance with requirements.

As a guiding principle, decentralization forms the basis for system resilience and trust by minimizing reliance on central components and major platforms. This approach enables systems to adapt and maintain stability during disruptions, providing the flexibility required to navigate complex environments and applications.

The Role of Open Source

Open source plays an important role in the development of interoperable software systems within data ecosystems.

- Trust and transparency: Open source projects are inherently transparent, fostering trust and collaboration, which are essential in trust-based data ecosystems.
- Neutral collaboration format: By now, "open source" no longer means that private individuals develop code in their spare time. The contemporary open source landscape provides environments with processes, legal frameworks, and support structures that enable organizations to collaborate on commercializable codebases and specifications. Such collaborations require a neutral format with strong governance models. In established open source foundations, global collaboration is possible, with everyone having the same rights based on the principle of meritocracy.
- Standardization and interoperability: The outputs of open source projects can be viewed as the result of the work of a consortium, similar to the initial steps of a standardization process. By combining open source and standardization paths, processes can be accelerated and standards brought to market more quickly. Speed is essential in the digital world: more and more traditional, lengthy standardization processes do not meet the needs of the fast-paced digital and data economy. In addition, standards in the context of open source are usually interwoven with a working (compatible) implementation, which increases actual interoperability. In the case of data ecosystems, for example, the Eclipse Foundation's liaison with the International Standardization Organization is a notable example.[25]

3.2.2. Political Strategies and Organizations

The second perspective covers the relationship from political strategies "down" to applications. Here, political strategies form a stable foundation for the organic growth of organizations and their business models, thereby enabling value-creating applications (see Figure 8).

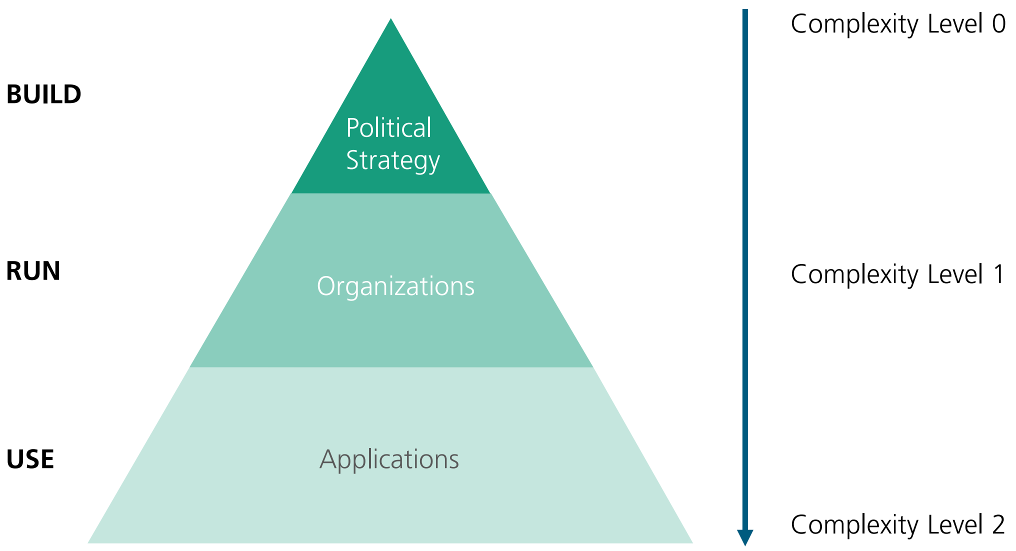

Figure 8: From Regulations to Applications

Different political approaches to data and data sharing strategies exist worldwide, and these strategies affect organizations and individuals in various ways.

- One way of implementing these strategies is to communicate proposals and strategy papers to the public, both verbally and in writing (e.g., opinion documents).[26]
- Another way involves offering incentives in the form of subsidies. In this case, politicians provide funding for specific purposes to organizations (companies, research institutions, and other entities), allowing them to allocate resources according to strategic guidelines and effectively implement their plans (e.g., Horizon Europe).[27]
- Lastly, regulations can be used to oblige organizations or society as a whole to adhere to political strategies through direct requirements (e.g., EU General Data Protection Regulation (GDPR)).[28]

Globally, various approaches guide the choice of means and actions. One approach favors a self-regulating economy that minimizes intervention and makes only slight adjustments only to address unintended consequences. The second approach prioritizes individual interests and aims to distribute value creation in new expansions, thereby countering the concentration of value and power at the architectural level. The third approach emphasizes the common good, prioritizing the will of the state or the entire population over individual interests. In this case, central control is exercised through guidelines designed to steer activities in a particular direction.

Regulations play a crucial role in the context of data sharing, acting as both a driver and a restriction. For example, on the one hand, regulations such as the EU's Corporate Sustainability Due Diligence Directive[29] require data regarding working conditions to be shared throughout the supply chain. On the other hand, the GDPR defines restrictions on the use of personal data. These limitations and obligations surrounding data sharing are not exclusive to the EU; similar frameworks exist at the international level like the US and China.

In today's interconnected world, supply chains cover large parts of the world and ultimately lead to the delivery of products. As a result, cross-border data sharing is essential for organizations to conduct their business activities effectively. Cooperation between organizations does not stop at borders for products; the same principle applies to the data associated with those products. However, combining inter-organizational and cross-border data sharing introduces additional challenges that must be addressed. One significant challenge is the legal aspect, which presents new territory for jurisdictions and raises questions about the allocation of legal responsibilities. Different legal systems can complicate matters, especially when local regulations grant access rights to third parties or states. To overcome these challenges, politicians must establish clear rules and agreements with various countries and economic regions, thereby providing clarity and reducing legal obstacles that hinder the creation of value from data. In addition, cross-border data sharing raises organizational concerns. Incompatibility between identification systems across countries and economic areas makes it difficult to identify partner organizations uniquely. This lack of compatibility reduces trust, which is essential for fostering data sharing.

A Global Comparison

In the US, the emphasis on the freedom and responsibility of the individual provides significant flexibility. In the past, this has led to numerous leading digital organizations that are positioned as global providers of digital infrastructure[30]. These organizations are increasingly acting as both gatekeepers and enablers for the adoption of new technologies across various sectors. This environment encourages innovative collaboration, such as data sharing among organizations. However, it also poses the risk of dependency on individual contributors. However, due to the focus on freedom and responsibility of the individual in the US, there are currently minimal restrictions limiting the actions and influence of these organizations. In addition, there has not yet been any significant funding dedicated specifically to projects that promote data sharing initiatives and create data ecosystems.

The EU focuses on personal and ownership rights protection, and this emphasis is becoming increasingly clear in the digital sector. As the EU has fallen behind in digital technologies, it aims to create and build up a niche by regulating these digital technologies. Data sharing has been identified as a strategic area where the EU aims to lead, particularly to maintain and enhance competitive advantages in traditional industries like manufacturing. The EU Data Strategy of 2020[31] served as a public starting point for this shift. It was supported by funding initiatives at both the European and national levels. Recently, there has been a strong focus on enhancing competitiveness in the digital space, as highlighted in the Draghi report[32] from September 2024, which calls for substantial investments. In addition, data spaces are being promoted as essential infrastructure to facilitate various applications based on data sharing.

In China, data is viewed as a strategic resource and is recognized as the "fifth market production factor".[33] The government has implemented measures to establish data sharing and has established the National Data Administration to support this initiative. Furthermore, initiatives such as the "Data Factor X", the "Three-Year Action Plan"[34], "Interim Provisions on Accounting for Enterprise Data Resources", and the "Trusted Data Space Development" Five-Year Action Plan have been introduced.[35] The objective is to create over 100 trusted data spaces by the end of 2028, which will include the establishment of 47 Data Trading Centers. Currently, ten of these centers are active and operational.[36]

Data ecosystems can only run sustainably when organizations recognize and embrace their role as active shapers and contributors of data-driven value creation, alongside political strategies and regulatory frameworks. In this context, regulations do not serve as coercive instruments, but rather as structural, enabling conditions that guide entrepreneurial actions toward responsibility, innovation, and collaboration.

Organizations as Operational Implementation Units for Political Strategies

While political-strategic measures, such as the EU Data Strategy and China's "Trusted Data Space" plan, establish the framework, it is organizations that create concrete applications, platforms, and services. They translate legal requirements into governance structures, technical implementations, and operational routines. Three roles are essential in this process:

- **Data Providers**
  Organizations that own and provide relevant data sets (e.g., machine manufacturers, transport organizations, hospitals).

- **Data Consumers**
  Entities that derive value from the data provided (e.g., analysis platforms, research institutions, business service providers).

- **Services Providers**
  Institutions that extend sharing, control, and trust mechanisms (e.g., data marketplaces, certification bodies).

The translation of regulations into operational practice is done by internal business processes, such as consent management, data classification, access control, and demonstrating compliance with data usage. It is crucial to incorporate compliance-by-design principles that integrate regulatory requirements with efficiency and automation in these areas.

Incentive Systems for Organizations: From Compulsion to Motivation

A key aspect of the "top-down" approach is the question: *Why should organizations share data?* The answer lies in an increasingly hybrid incentive system that combines economic, regulatory, and ethical motivational structures.

- **Regulatory Incentives**
  Compliance with legislation such as the Data Act, GDPR, or sector-specific standards (e.g., healthcare, finance) creates implicit pressure for organizations to provide and control the use of data. In some cases, sharing data becomes a legal obligation.

- **Reputational Incentives**
  Organizations that share data transparently and responsibly position themselves as innovative, sustainable, and future-oriented – a key competitive factor in data-driven markets.

- **Economic Incentives**
  Participation in data ecosystems results in new business models, data-driven services, and efficiency improvements (e.g., predictive maintenance, CO_2 monitoring, process automation).

- **Cognitive Incentives (Autonomous Thinking)**
  Particularly in European contexts, there is a growing awareness of digital sovereignty and self-determination. Organizations are increasingly interested in retaining control over their data while actively participating in collaborative networks that focus on value creation – in the sense of responsible autonomy.

Ownership as a Central Governance Resource

Data ecosystems require a shift in the governance paradigm, transitioning from a centralized control model to one of decentralization and self-regulation. Organizations must take responsibility for their handling of data, both technically and ethically. This self-responsibility is demonstrated through various initiatives, including:

- defining internal "data steward" roles,
- establishing data-specific competence centers,
- actively participating in standardization and open source communities,
- engaging in a discourse on data ethics and corporate digital responsibility.

For organizations to fully realize the regulatory potential of data ecosystems, they must go beyond mere compliance and view themselves as proactive architects of a data-driven future.

3.2.3. The role of Applications for Value Creation

Applications represent the interaction environment for humans that collapses the chaos into a deterministic, human-understandable form. In a data ecosystem, these applications rely on shared data and trust mechanisms. The application layer serves as the point where technical principles and governmental strategies and measures are transformed into concrete value creation. Applications are developed based on specific individual and sector-specific requirements and are highly sensitive to context. Their design is not only context-dependent but also constantly evolving: a data ecosystem generates a wide range of applications tailored to the unique business needs of its participating organizations and continuously change their patterns. Complete control over this dynamic landscape is unachievable due to insufficient situational awareness and scaling challenges.

Data value is created by patterns – not by individual components. Thus, data value creation is an emergent phenomenon: data becomes valuable not in isolation but through its integration into recurring structures and value creation patterns. These patterns can include:

- Value flows along supply chains, such as through component traceability or shared sustainability metrics.
- Cooperative data economy, where data that may not be meaningful on its own gains value through sharing and aggregation.
- Cross-domain business models that combine sector-specific data spaces (e.g., mobility, tourism, and environment).

Outlook for Value-Adding Application Patterns

LLMs serve as a helpful example to illustrate how data value can be grasped through applications. With the rapid advancement of generative AI systems, particularly LLMs, a new paradigm of data-driven applications is emerging: semantically enriched, context-adaptive data interaction. While static dashboards or deterministic queries primarily characterize today's applications, LLMs provide dialogical, explainable, and flexible interfaces that are suitable for both human users and machines.

Component Change Impact Analysis in Manufacturing
When a manufacturer plans to change the size of a component, this modification can have indirect effects on other components of the product. Typically, data silos hinder transparency regarding the impact on interconnected components and partner organizations. This lack of transparency can lead to production delays, increased costs, and disruptions in the supply chain. By leveraging data ecosystems, organizations can access relevant partner data to understand the impact of their operations throughout the supply chain. This leads to quicker decision-making, cost reductions, and enhanced collaboration within the ecosystem. LLMs can help automate this process. In this scenario, the LLM would analyze technical documents to determine the implications of a component change. It would also identify which components are affected and to what extent, and it could assist in finding new suppliers for the required components.

4. Recommendations

Given the growing importance of data in today's data economy, fostering trust in data sharing is essential for supporting innovation and securing competitive advantages. The following recommendations for policymakers and decision-makers aim to establish frameworks that enable data sharing in trust-based data ecosystems while protecting the rights and interests of all parties involved.

- **Embrace Chaos for Innovation.**
  No stakeholder can or should have the mandate to control the entire system. Policymakers should accept chaos as a driver for innovation, allowing systems to self-organize and evolve. This involves setting clear goals and ways to measure and incentivize them while leaving the paths to achieving them open, thereby creating an environment in which value can emerge organically.

- **Promote Trust Protocols.**
  Establish trust protocols for organizations and data sharing to ensure a reliable trust context. This includes creating governance frameworks and rulebooks that strengthen trust among stakeholders.

- **Support Open Standards and Interoperability.**
  Encourage the development and adoption of open standards and protocols to facilitate interoperable data sharing across different platforms and sectors.

- **Facilitate Cross-Border Data Sharing.**
  Develop clear rules and agreements to overcome legal and organizational barriers in cross-border data sharing. This includes harmonizing identification systems and metadata standards to foster trust and promote collaboration across organizations, domains, and countries.

- **Encourage Decentralized Architectures.**
  Support decentralized architectures that reduce reliance on central components such as data intermediaries, thereby enhancing system resilience and adaptability. Each central system poses a security risk and reduces trust.

- **Create Incentive Systems.**
  Implement hybrid incentive systems that combine economic, regulatory, and ethical motivations to encourage organizations to share data effectively. This includes regulatory compliance, reputational benefits, and economic opportunities from data-based services.

- **Encourage Autonomy, Agency, and Responsibility.**
  Promote the concept of autonomy, self-determination, and self-responsibility among organizations, encouraging them to actively participate in collaborative value creation networks while keeping control over their data. This involves defining internal roles, participating in standardization efforts, and engaging in discussions related to data ethics.

Returning to the model of constructive chaos patterns and self-organization, the following picture emerges: various technical and political measures lead to the development of diverse applications and value creation patterns. These can build upon each other and trigger or influence other patterns. Ultimately, this can lead to a network of different focal points that organize themselves and generate an ever-increasing number of applications. The color gradient in Figure 9 indicates how the measures taken create and expand data value. Value is generated where data meets applications, where it is shared and used (highlighted in orange).

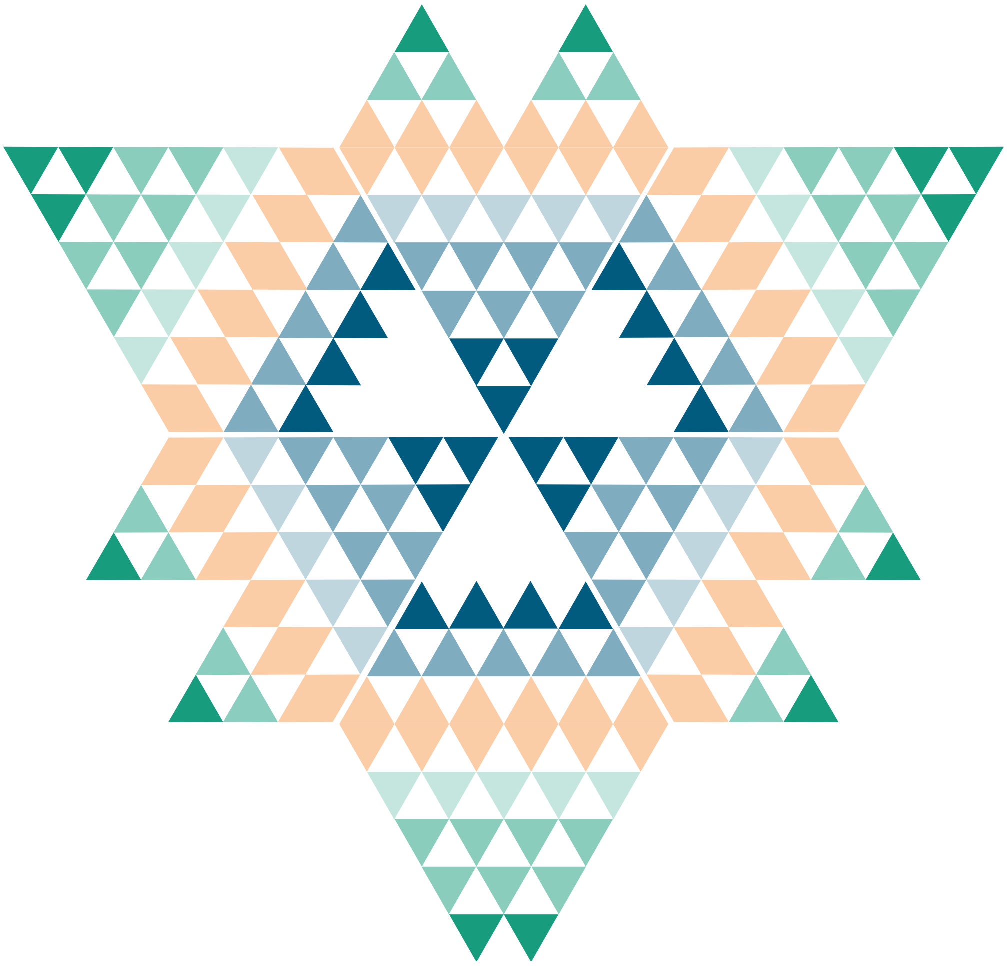

Figure 9: Abstract Representation of an Organically Grown Trust-based Data Ecosystem

References

1 Statista (2024). *Volumen der jährlich generierten/replizierten digitalen Datenmenge weltweit von 2010 bis 2023 und Prognose für 2028.* https://de.statista.com/statistik/daten/studie/267974/umfrage/prognose-zum-weltweit-generierten-datenvolumen/ [Accessed: June 3, 2025].

2 De Franceschi, A., & Lehmann, M. (2015). *Data as Tradeable Commodity and New Measures for their Protection.* https://sfera.unife.it/bitstream/11392/2339106/2/592-data-as-tradeable-commodity-and-new-measures-for-their-protection.pdf [Accessed: June 3, 2025].

3 Theraux, O., & Keller, J. R. (2021). *The economic impact of trust in data ecosystems – Frontier Economics for the ODI.* https://theodi.org/insights/reports/the-economic-impact-of-trust-in-data-ecosystems-frontier-economics-for-the-odi-report/ [Accessed: June 3, 2025].

4 Cox, C. B., Moore, P. D., & Ladle, R. J. (2016). *Biogeography: An ecological and evolutionary approach* (Ninth edition). Wiley Blackwell.

5 Möller, F., Jussen, I., Springer, V., Gieß, A., Schweihoff, J. C., Gelhaar, J., Guggenberger, T., & Otto, B. (2024). Industrial data ecosystems and data spaces. *Electronic Markets, 34*(1). https://doi.org/10.1007/s12525-024-00724-0

6 International Data Spaces Association (2025). *Manifesto of International Dataspaces.* https://docs.internationaldataspaces.org/manifesto-for-international-dataspaces [Accessed: June 3, 2025].

7 Catena-X Automotive Network e.V. (2025). *Catena-X.* https://catena-x.net/ [Accessed: June 3, 2025].

8 Verband der Automobilindustrie e.V. (2025). *Regulations and standardization.* https://www.vda.de/en/topics/automotive-industry/standardization-and-technical-standards [Accessed: June 3, 2025].

9 Open Industry 4.0 Alliance Implementation GmbH (2025). *Projects: 11 dedicated use cases.* https://factory-x.org/projects/ [Accessed: June 3, 2025].

10 Federal Ministry for Economic Affairs and Energy (2025). *What is Plattform Industrie 4.0.* https://www.plattform-i40.de/IP/Navigation/EN/Home/home.html [Accessed: June 3, 2025].

11 OPC Foundation (2025). *OPC 30270: Industry 4.0 Asset Administration Shell.* https://reference.opcfoundation.org/I4AAS/v100/docs/#4.1 [Accessed: June 3, 2025].

12 Open Industry 4.0 Alliance Implementation GmbH (2025). *TP 2.03 Collaborative Information Logistics.* https://factory-x.org/use-case/tp-2-3-collaborative-information-logistics/ [Accessed: June 3, 2025].

13 EONA-X (2025). *What are we improving?* https://eona-x.eu/#page-3 [Accessed: June 3, 2025].

14 Hebant, P. (2024). *Data Spaces Symposium: EONA-X and Interoperability.* https://www.data-spaces-symposium.eu/wp-content/uploads/2024/03/14.25-Patrick-Hebant-Data-Spaces-Symposium-2024-EONA-X-and-Interoperability.pdf [Accessed: June 3, 2025].

15 Von Bertalanffy, L. (1968). *General system theory: Foundations, development, applications.* G. Braziller.

16 Wang, P. (2021). Connecting the Parts with the Whole: Toward an Information Ecology Theory of Digital Innovation Ecosystems. *MIS Quarterly, 45*(1), 397–422. https://doi.org/10.25300/MISQ/2021/15864

17 Benbya, H., Nan, N., Tanriverdi, H., & Yoo, Y. (2020). Complexity and information systems research in the emerging digital world. *MIS Quarterly, 44*(1), 1–17.

18 Merali, Y. (2006). Complexity and Information Systems: The Emergent Domain. *Journal of Information Technology, 21*(4), 216–228. https://doi.org/10.1057/palgrave.jit.2000081

19 Kauffman, S. A. (1993). *The origins of order: Self-organization and selection in evolution.* Oxford University Press, USA.

20 Wang, X., & Vidgen, R. (2007). Order and Chaos in Software Development: A comparison of two software development teams in a major IT company. In *ECIS 2007 Proceedings* (Vol. 66, 807-618). https://aisel.aisnet.org/ecis2007/66

21 Dhillon, G. S., & Ward, J. (2003). Chaos theory as a framework for studying information systems. In *Advanced Topics in Information Resources Management, Volume 2* (pp. 320–337). IGI Global.

22 McBride, N. (2005). Chaos theory as a model for interpreting information systems in organizations. *Information Systems Journal, 15*(3), 233–254. https://doi.org/10.1111/j.1365-2575.2005.00192.x

23 Lorenz, E. N. (1963). 1963: Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences 20,* 130-141.

24 Lorenz, E. N. (1972). Predictability: Does the flap of a butterfly's wing in Brazil set off a tornado in Texas? Presented before the American Association for the Advancement of Science, December 29, 1972.

25 Eclipse Foundation. (2023). *Eclipse Foundation Establishes Category A Liaison With ISO/IEC JTC 1/SC 38 for Cloud Computing and Distributed Platforms.* https://newsroom.eclipse.org/news/announcements/eclipse-foundation-establishes-category-liaison-isoiec-jtc-1sc-38-cloud [Accessed: June 3, 2025].

26 European Data Protection Board (2025). *Opinions.* https://www.edpb.europa.eu/our-work-tools/consistency-findings/opinions_en [Accessed: June 3, 2025].

27 European Commission (2021). *Horizon Europe.* https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe_en [Accessed: June 3, 2025].

28 European Parliament & Council of the European Union (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).* http://data.europa.eu/eli/reg/2016/679/oj

29 European Parliament & Council of the European Union (2024). *Directive (EU) 2024/1760 of the European Parliament and of the Council of 13 June 2024 corporate sustainability due diligence and amending Directive (EU) 2019/1937 and Regulation (EU) 2023/2859.* http://data.europa.eu/eli/dir/2024/1760/oj

30 Statista (2024). *Volumen der jährlich generierten/replizierten digitalen Datenmenge weltweit von 2010 bis 2023 und Prognose für 2028*.

31 European Commission (2019). *European data strategy: Making the EU a role model for a society empowered by data.* https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en [Accessed: June 3, 2025].

32 European Commission (2024). *The Draghi report on EU competitiveness.* https://commission.europa.eu/topics/eu-competitiveness/draghi-report_en [Accessed: June 3, 2025].

33 Mark, L., & Chang, M. (2024). *China's Data as a Fifth Market Production Factor – an Asset on Your Balance Sheet.* https://www.haynesboone.com/news/alerts/chinas-data-as-a-fifth-market-production-factor--an-asset-on-your-balance-sheet [Accessed: June 3, 2025].

34 Mark, L., & Chang, M. (2024). *China's Data as a Fifth Market Production Factor – an Asset on Your Balance Sheet.* https://www.haynesboone.com/news/alerts/chinas-data-as-a-fifth-market-production-factor--an-asset-on-your-balance-sheet [Accessed: June 3, 2025].

35 Kwok, W., & China Brief Notes. (2024). *'Trusted Data Spaces': Beijing's Latest Step Toward Digital Dominance.* https://jamestown.org/program/trusted-data-spaces-beijings-latest-step-toward-digital-dominance/ [Accessed: June 3, 2025].

36 Zheng, W. (2024). *China aims for more than 100 'trusted data spaces' by 2028 under national action plan.* https://www.scmp.com/news/china/politics/article/3287937/china-aims-more-100-trusted-data-spaces-2028-under-national-action-plan [Accessed: June 3, 2025].

Image Credits

©Pascal – AdobeStock, title