

Data sharing in industrial ecosystems

Driving value across entire production lines



Preface

Industrial ecosystems for data sharing have the potential to power tremendous growth, helping companies optimize existing processes and making new products and businesses possible. Handling and sharing data is a key component of these ecosystems. Potential ecosystem members face a number of obstacles, from concerns about protecting confidential data, to the lack of a common definition of data, to the technical challenges of sharing data securely and maintaining control. While industrial ecosystems are still in their infancy, enough success cases exist to identify both the main stumbling blocks and the factors that make these approaches successful. Companies that take proactive steps now to tap into the power of ecosystems can secure a significant competitive advantage.

Access to data and its systematic collection and processing have become key differentiators for industrial companies. Thanks to advances in connectivity and edge computing, they can use previously siloed data to optimize machine performance and even integrate data further to drive value across entire production lines. However, few players have the in-house skills and resources to systematically uncover such sources of value. As a result, both solutions equipment manufacturers and system integrators have started developing solutions as additional services, an attractive new source of income that helps balance their declining margins in many sectors. In extreme cases, entirely new business models have emerged that have retired existing approaches. Rolls Royce's "power by the hour" model of selling uptime instead of turbines is a good example. Analyzing usage and service data beyond primary purpose of data for direct control enables GE to turn around their entire business model. Another example is Caterpillar's Cat Connect service, which increases fuel and materials efficiency, productivity, and safety.

But the need to share data beyond organizational boundaries has turned out to be one of the main barriers to full-scale adoption. In a connected world, sharing data can be a powerful enabler for all sides: the parties who supply data and the providers developing new services or even disrupting markets with more attractive offerings. Such innovation often requires data from different sources and, potentially, different organizations. These could include component suppliers, machine integrators, or machine users. In addition, context data (such as information on environmental conditions) might be added from other previously unconsidered sectors. The expected result: ecosystems that create winning situations for all participants. Even competing organizations may be able and willing to contribute to such ecosystems if confidentiality is guaranteed and the value they stand to gain exceeds the required investment.

At the same time, win-win data ecosystems are not feasible in every case. For example, if certain players are extremely dominant, they can dictate not only which data formats smaller players must use but also how they participate in the economic benefits. One such dominant player is Walmart, which manages the sharing of supplier data through Retail Link, its proprietary data platform.

Prof. Dr.-Ing. Boris Otto

Managing Director, Fraunhofer Institute
for Software and Systems Engineering,
Holder of the Chair of Industrial Information
Management at the TU Dortmund University

Prof. Dr. Niko Mohr

Partner, McKinsey & Company,
Dusseldorf

Dr. Matthias Roggendorf

Partner, McKinsey & Company,
Berlin

Tobias Guggenberger

Research Associate,
Chair of Industrial Information Management
at the TU Dortmund University

April 2020.

Laying the groundwork for multiparty industrial data sharing

Companies have been sharing data in bilateral, trusting relationships for decades. For example, they may trade information in closed supply chains to make vendor-managed inventories possible or in the finance space to automate invoicing and payment flows.

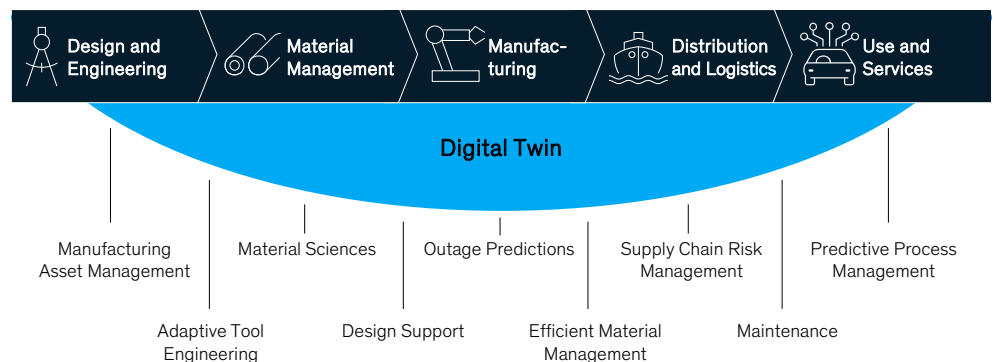
However, an ecosystem is far more complicated. For example:

- The data involved is often more complex, and the parties do not align on data ontology in advance. If no industry standard exists, parties may negotiate proprietary standards.
- Mechanisms for ensuring data is used as agreed are much harder to enforce and monitor.
- Value from data needs to be shared within the ecosystem to keep parties interested in contributing. Imbalances quickly lead to dropouts, which may endanger the stability of business models.

On the positive side, opportunities from industrial ecosystems are substantially higher than for data-sharing setups that function as one-trick ponies. As data sharing expands beyond single use cases and data usage, participants may explore different types of innovation. These may involve blending data with other sources as mentioned above, applying machine learning or artificial intelligence, or extracting value from streams of real-time data that might diminish as the data becomes stale. With the proliferation of IoT devices and Industry 4.0 adoption, ever-growing amounts of data – with largely unexplored value in many cases – are becoming available.

Industrial data ecosystems also play an important role in the successful use of digital twins. A digital twin is a virtual representation of a machine, a production line, or even an entire production environment such as a factory or (as in semiconductor or battery production) a network of factories. While bilateral data sharing might be sufficient for implementing a digital twin of a single machine, extending the scope requires a more advanced approach: a shared digital twin involves multiple partners and suppliers, each with specific interests, available data, and the internal capabilities for leveraging and sharing information.

Exhibit 1:
A digital twin comprises data about all lifecycle phases of a real-world object



Technical solutions are needed for receiving, storing, and sharing – and, depending on ecosystem policies and rules, processing and analyzing – data from the ecosystem’s different members. Many industrial companies have extensively considered specialized data and IoT platforms to help them master their data-sharing challenges. And they have a wealth of options: in recent years, more than 1,000 such offerings have hit the market¹. However, many industry decision

¹ VDMA Plattformökonomie im Maschinenbau (https://www.vdma.org/documents/15012668/26471342/RB_PUB_18_009_VDMA_Plattform%C3%B6konomie-06_1530513808561.pdf/f4412be3-e5ba-e549-7251-43ee17ec29d3)

makers feel that these platforms only partially meet expectations. While they generally provide a solid basis for hosting massive data volumes or showing real-time data streaming patterns, they fall short in terms of collaboration and sharing capabilities. Often, they are only designed for single-party use. Big cloud hyperscaler often offer more flexible platform services and allow companies to share data more openly while staying in full control of these data assets.

Key challenges in exchanging data in industrial ecosystems

Data ecosystems obviously have clear benefits, so what is holding companies back from taking full advantage of them? The reasons are manifold. Some involve organizational, technical, or legal barriers or restrictions, while others are fueled by typical ways of thinking and experience-based concerns. They can be divided into four major categories:

- **Challenges related to culture and mindset.** New value opportunities depend on sharing data in larger industrial ecosystems that may involve parties with limited trust in one another. Participants might be able to analyze shared data to derive confidential business information. Considering every eventuality is difficult. For example, it seems reasonable to share welding position data so that another party can optimize production pathways, but a competitor could potentially use that data to analyze the resulting shapes – and determine in advance what new products the first company is planning. Typically, however, data in isolation rarely creates a competitive advantage.

Another major reason that companies do not engage in data-sharing models is the perceived lack of control when data leaves their premises. For example, software licensing has reached a high level of maturity (in terms of both contracting and technical measures) but data sovereignty is still hard to guarantee and implement.

The need to identify enough value to justify investment can also be an obstacle. Some organizations have difficulty breaking out of existing thinking patterns, while others lack the capabilities to run the kind of structured assessment needed to drive engagement in data ecosystems. Often, this requires a collaborative give-and-take governance model at odds with traditional buyer-supplier-relationships.

- **Challenges related to the need for ontologies.** Data can only be shared if it can be easily interpreted and quickly integrated with other data sources. It must be unambiguously understood by all ecosystem members. Thus, a common language or ontology that all parties agree on is required. In some industries, such as banking or healthcare, such data languages were established decades ago to enable interoperability and the leveraging of standard software. In the industrial sector, however, individual ontologies and standards exist – such as eCl@ss and the Asset Administration Shell, to name just two examples. What is needed, though, is a comprehensive architecture of standards that includes different types of data such as master data, reference data, manufacturing and supply chain event data, etc. This architecture stack cannot be developed by an individual ecosystem member, but rather requires consensus within a community of practice.

An industrial data ecosystem promising a significant competitive advantage could be a strong impetus for companies to pragmatically define standards for sharing data beyond single-purpose cases and allow other parties to interpret it correctly. Generic ontologies for exchanging sensor and robotics data that can be easily extended are on the rise due to advances in connected devices and the need for interoperability in heterogeneous environments. However, industrial environments are often a mix of older systems and new IoT-based sensors, making it even harder to reach agreement on common standards.

- **Technological challenges to designing and mastering the required platform and services.** Even when companies understand the value of sharing data in industrial ecosystems, they often lack the skills to make their data available in an effective and efficient way. The challenge starts with internal technical interfaces for supplying data and continues with developing an environment to process and integrate data from different sources and run models using advanced analytics methods. If companies decide to build new products based on such data, they also need the capabilities to productionize the resulting models and ensure effective commercial and technical operations, including guaranteeing service levels and scaling up to serving potentially millions of customers. These capabilities typically do

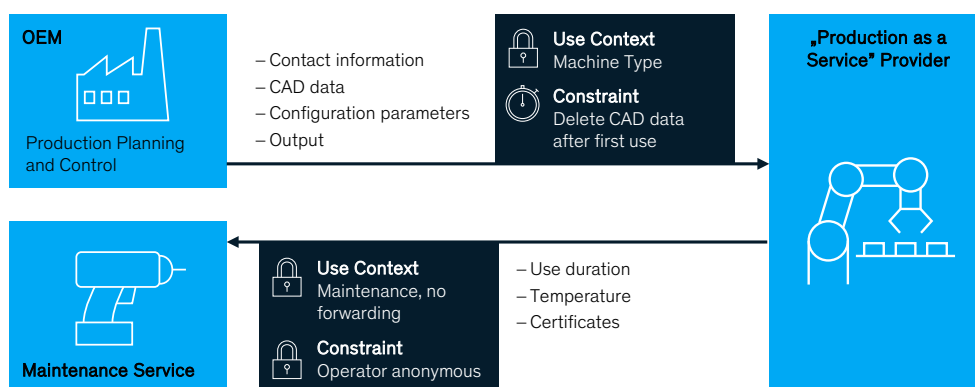
not exist in engineering-focused industrial clients, so they need to be systematically built either as new entities or with external help.

— **Challenges in effectively managing data to make timely, automated access to it possible.**

The last challenge relates to the internal capabilities required to provide data from internal applications that fits its intended purpose, is of sufficient quality, and is available in a timely fashion. A lot of data sits in silos and is not even accessible. Other data may be used for its primary purpose (for example, on the shop floor) but not for any others. Also, many organizations, especially in the industrial space, have limited capabilities for managing their data assets strategically. As a result, they have trouble distinguishing between data that can be easily shared and strategic data that poses significant opportunities but entails the risk of exposing intellectual property to untrusted entities. Before these companies participate in a data ecosystem, they need to establish basic data governance principles and systematically develop their employees' data literacy.

So far, only niche solutions in highly selective segments have been developed (see boxes for examples). In many cases these data-sharing ecosystems were created by a dominant player with the power to mandate that other market participants join the network.

Exhibit 2:
Data sovereignty is achieved by attaching usage conditions to shared data – as the additive manufacturing scenario shows



Example 1: Lufthansa AVIATAR²

Lufthansa's AVIATAR is a platform for the aviation industry. It integrates data from sources on airline operations, aircraft, maintenance systems, and more to create a comprehensive tool for fleet management. As it works independently of organizational or technical boundaries, it can be used to optimize all operations of any imaginable configuration of fleets.

Case study 2: 365FarmNet³

365FarmNet is an innovative software platform for manufacturer-independent farm management. The service integrates multiple partners and data sources to create a comprehensive management tool. Its modular architecture combines free-to-use functions (such as data management, weather analytics, documentation, and basic planning tools) with complementary value-added services (for telematics, route optimization, soil sampling, and more). The open interface makes it possible to integrate intelligent apps from a variety of agricultural suppliers, such as machinery manufacturers or pesticide and fertilizer producers, creating a data ecosystem around the platform. The approach demonstrates that bringing an integrated value proposition to customers requires data sharing across sectors.

² <https://www.aviatar.com>

³ <https://www.365farmnet.com>

Success ingredients of successful data-sharing models

Many industrial data ecosystems are still in their nascent stages and many players are only taking their first steps in this area, often in the form of bilateral data sharing. However, the early successes that do exist provide a basis for identifying steps to increase the chances that an industrial ecosystem will prosper. As more companies participate in data ecosystems, this list will surely evolve:

- **Systematically assess business opportunities from sharing your own data and gaining access to new data.** In a first step, it is important for organizations to understand an ecosystem's potential in terms of the business it will generate rather than the data the company will share. Such an approach requires looking broadly at both internal applications and opportunities that can be cocreated with customers, or even new business they can build outside the core. In some cases, additional value may come from selling data alone, while in others it may involve some give and take. Companies often fail to articulate their expectations clearly and underestimate the time needed to achieve benefits. Also, it is crucial to understand the main actors in the potential business ecosystem – for example, who will need to play the orchestrator and who are the participants – and what mechanisms would incentivize them to share their data (for example, if value creation is unevenly distributed).
- **Define your own role and options.** After understanding the business opportunity offered by the ecosystem and its operating protocol, companies need to thoroughly assess their strategic options. These might involve becoming a data owner, a participant in an emerging ecosystem, or a provider of services. The choice of this role largely determines the required investments, the organizational setup, and the resources needed to drive implementation from both the business and technology sides.
- **Create a value- and risk-driven approach to data.** A systematic scan of corporate and transactional data can reveal the value that can be created and the intellectual property that might be at stake. Options should be considered for sanitizing or aggregating the data so it can still be used with partners. Methods for managing data value and risk should be put in place for data assets assessed as strategic.
- **Establish sustainable data governance.** Once relevant data has been identified, scaling up the approach requires suitable data governance. This involves establishing business ownership of internal data so that it can be shared externally and integrated quickly with data from new sources. Companies may also want to set up a set of data management tools, such as a data catalog (or internal data marketplace), as well as a basic tools for topics such as tracking data lineage, classifying data risk, or documenting data access policies. Data partnership managers may be needed to drive data identification and commercialization.
- **Embark on a data platform strategy.** Often, a company's legacy IT systems are highly complex, and its data is siloed in operational applications. Inflexible systems and data models determined by software vendors make the data production side of the equation difficult to change. The data consumption side, however, is easier to evolve. The key to gaining traction is a flexible data consumption architecture that enables easy blending of data, real-time data processing, and the use of analytics and AI. In many cases, the main work entails making existing data accessible to such an analytics environment rather than creating the environment itself, as cloud-based analytics solutions are fast to set up and use.
- **Apply a test-and-learn approach before scaling:** Companies should not be afraid to learn by doing. Set an aggressive goal of taking a first minimal viable product live with internal or external customers within a few weeks. Start with use cases of limited complexity that contribute to overall strategy. Find a friendly partner or customers who are excited to learn with you. Communicate your successes and grow your capabilities along the way.

The market for data-sharing platforms that guarantee data sovereignty is still nascent. Furthermore, commercial approaches are difficult to establish (except in niches) due to the need for a neutral supplying entity (a prerequisite that may interfere with commercial interests).

Players that engage early can shape the market and build highly innovative business models and products or services that competitors cannot easily copy. In some cases, such solutions become add-ons to the existing business; in others, the superiority of the new solutions could fully disrupt the industry.

A key requirement for the proliferation of industrial ecosystems is the existence of standards for data interoperability, portability, and sovereignty. Promising standardization initiatives such as International Data Spaces (IDS) and Trusted Cloud – as well as the recently started GAIA-X project – are steps into the right direction.