How data sovereignty enables the next future of automotive

DEEP DIVE INTO DATA SPACES – #1

T Systems

Let's power higher performance

Table of contents

Use of data science in the automotive industry:

The last few years have been challenging and disruptive for many industries and obviously the whole of the automotive value chain was impacted. Just as cars have evolved from being mainly mechanical devices to sophisticated computerized mobility machines, the OEMs have to evolve their business model from being car manufacturers to facilitators of mobility services. Data analytics is playing a vital role in this evolution. And we have heard from many experts across the different tiers of the industry how they are trying to cope with these changes and challenges.

Data science with artificial intelligence, machine learning and analytics in the automotive industry is not just limited to creating features like self/ autonomous driving. Companies across the automotive industrial value chain can leverage the power of data analytics to improve research, design engineering and supply chain processes, and thus evolve new business models to remain ahead in the competition. Apart from developing new mobility features, AI can also help to improve efficiency and product automation, in cutting costs and to better serve customers.

The automotive ecosystem is evolving with the fast-changing relationship between auto players and software providers. Data - recorded from supply chains to sensors - is at the helm of these changing roles of the stakeholders. The magnitude and complexity of this data is as immense as the possibilities it presents. Also, it is here that we observe the challenges of data collaboration between different entities across the value chain.

Industry observers, consultants and analysts are all harping about the world of potential that AI can bring to automotive. But they are equally weary about data security and sovereignty. The industry is yet to start getting the returns on the Billion Euro AI promise. Why is it so and what is happening? How can data sovereignty help the auto industry to win at the ever-evolving market landscape?

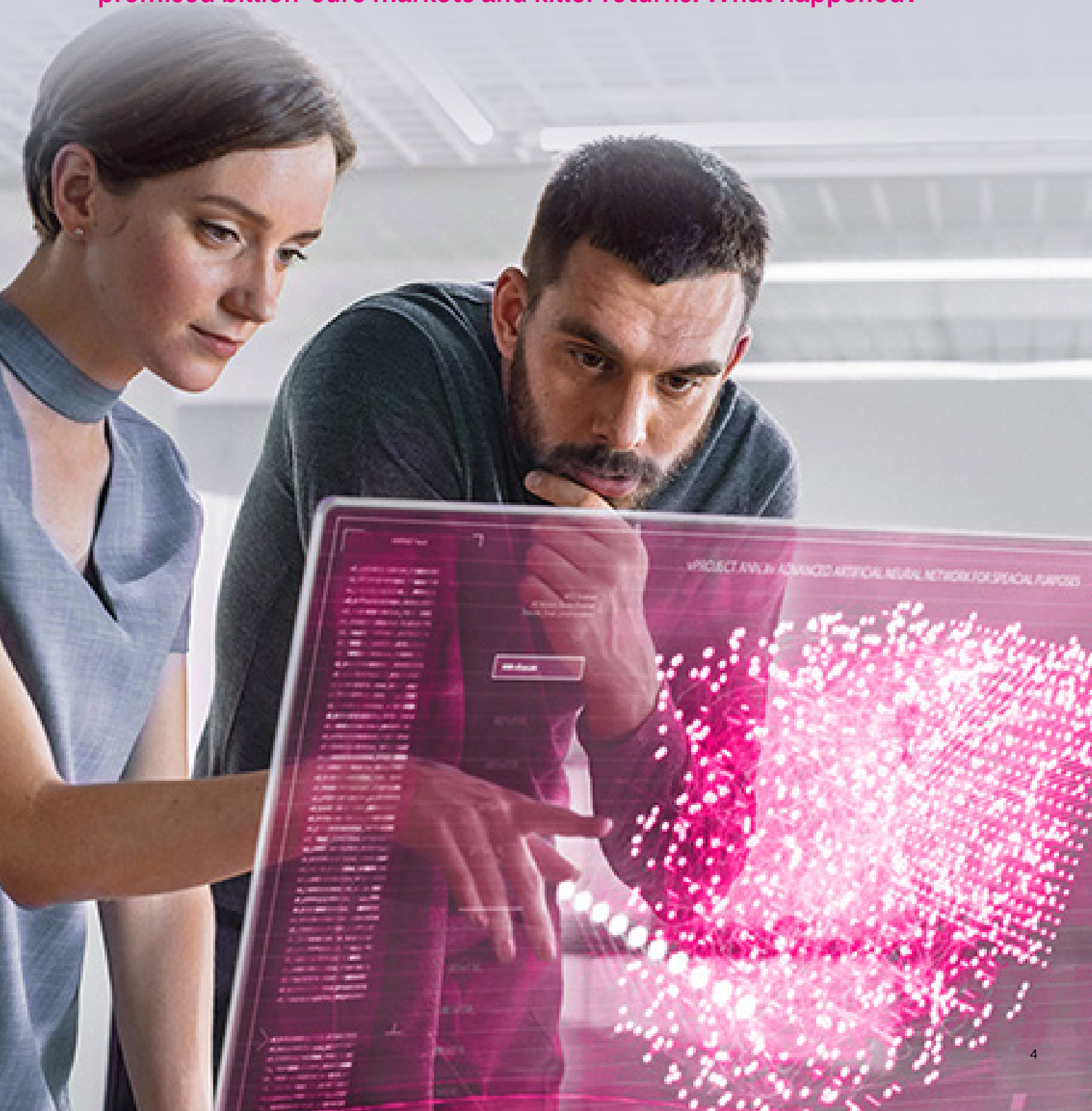
For answers to these questions, enjoy our deep dive, and explore our solutions and insights on data collaboration and sovereignty.

Dr. Christian Hort
Senior Vice President Automotive
T-Systems International GmbH

Prof. Dr. Christoph Schlueter Langdon
Data Analytics Executive and Scientist
T-Systems International GmbH

Winning with data sovereignty: From data icebergs to data products, data factories, and data space supply chains

Where are the returns from data science and artificial intelligence? Industry observers, consultants, analysts, and investment banks have promised billion-euro markets and killer returns. What happened?



Data iceberg – only 10 percent is visible

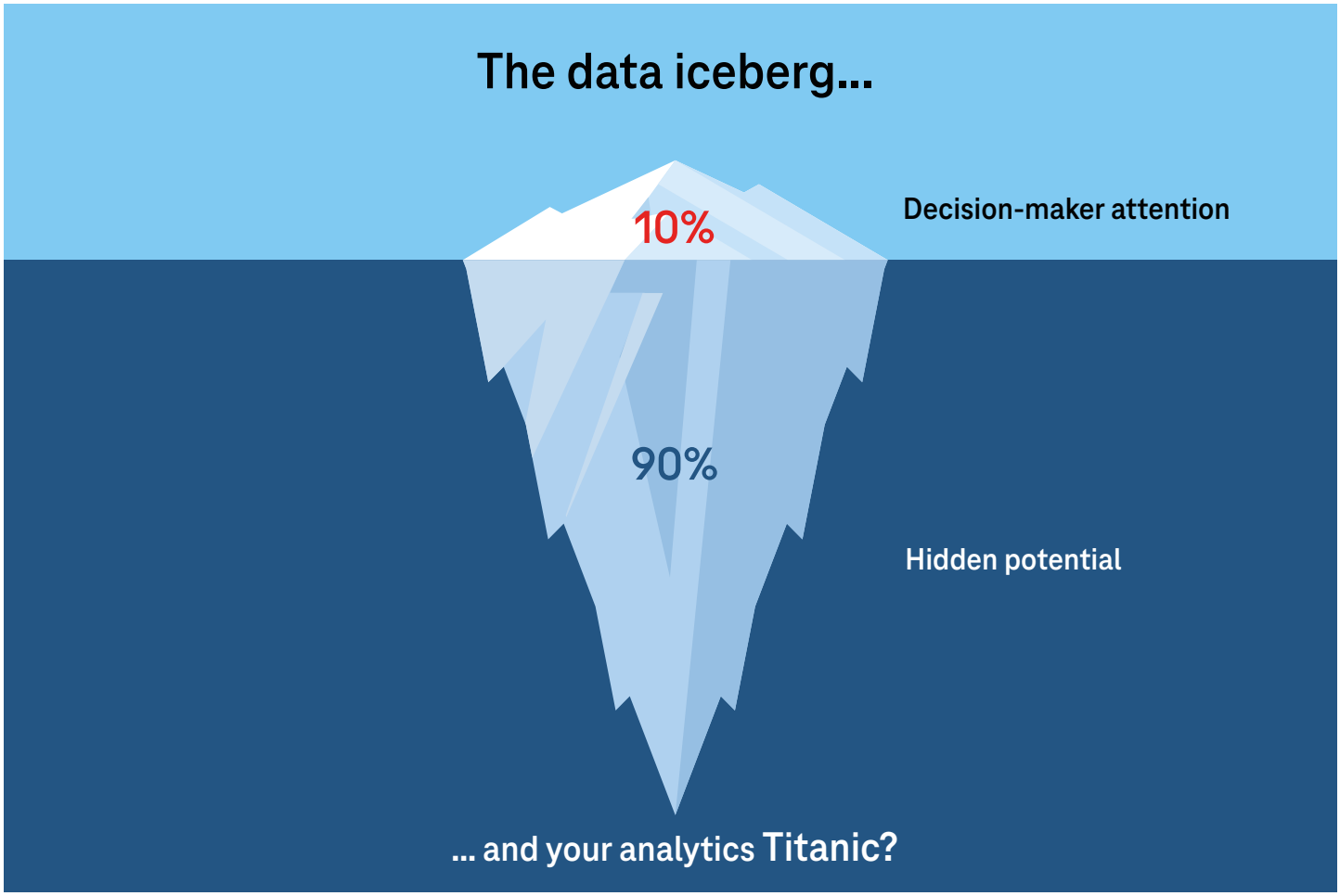


Figure 1: Data Iceberg - hidden potential

To understand what is going wrong with data, think of data as an iceberg (see Figure 1). We humans are highly visual creatures and can see danger from afar (Kaas 2013, 1992), but data is like an iceberg, meaning we only see perhaps 10 percent of its total effort and potential. The rest remains hidden, invisible to the human eye. Our brain cannot automatically translate reams of numbers into something meaningful – instead we need theory and tools to visualize it, such as descriptive statistics and charting applications. This biological shortcoming is compounded by a lack of proper training. When we ask an audience who among them has attended a statistics class, all hands go up. For data, no hands. Hardly anybody has had proper training on data. Curiously, if we ask the audience something about statistics, such as regarding a t-test there is silence. If we ask about data, all hands go up. Somehow, it feels easy to engage with, easy to jump to conclusions. This could explain early stumbles with business returns from data science investments.

Source - *Figure 1: Chris s. Langdon, Drucker Customer Lab, CGU, https://research.cgu.edu/drucker-customer-lab/

Fixing the data bottleneck: three key themes

The boom in artificial intelligence has put data in the spotlight, but the most data is yet to come (see Figure 2). Experts know that any algorithm is only as good as the data used to create it. While this is true for traditional artificial intelligence methods, it is especially the case for more modern, bio-computational ones (see Schlueter Langdon 2020 for example). The former includes statistical methods such as k-means clustering – known as an unsupervised machine-learning method – for market segmentation in the field of marketing, for example. The latter includes methods that loosely mimic biological phenomena such as neural networks, which were

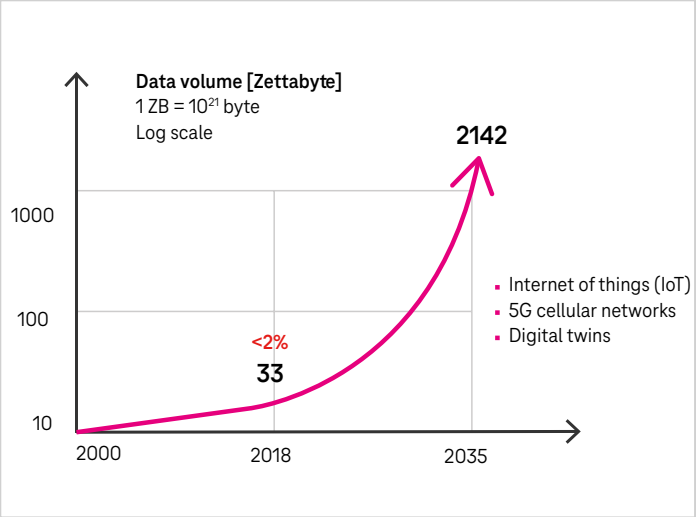


Figure 2: If you're already struggling now... it's only 2 percent of the next 15 years of data

originally derived from our understanding of how the human brain works. Today it is used in the form of deep learning for pattern analysis used for facial recognition, for example, or to power voice bots or virtual assistants like Amazon's Alexa, for example (Crosby & Schlueter Langdon 2017). In automotive, deep learning expertise is the key pillar for autonomous driving. Therefore the OEMS or Tier 1s maintain large fleets to amass data. Regardless of AI type, data holds the key to success. No surprises then that management has finally shifted its attention to data. This is good news, since it is broken and is inhibiting the widespread success of AI. There are success stories, yet they seem to be the exceptions that prove the rule. Why is that? Look no further than the time it takes to create a success story. At least 80 percent of the budget of a data analytics project is spent on data, less than 20 percent on results (see Figure 3). With analytics, we are running the Pareto principle of business efficiency – the 80/20 rule – in reverse. For AI to deliver success stories in business, this must change. And urgently so, because Figure 2 tells us that we are already struggling with less than 2% of the next 15 years of data. Without a fix all will go from bad to worse, and we can't sit it out or shift careers as data is emerging everywhere, quickly and exponentially.

How can we fix it? Just as any good doctor would avoid jumping to conclusions, we present three themes that carefully follow the cycle from symptoms to diagnosis in order to arrive at the correct root-cause solution or "medicine." This is a rapidly evolving field and by no means the be-all and end-all or last word on the subject. Far from it; instead we would like to share three themes based on our insights fresh from the field:

- **Treat data as a product**
- **Create data factories**
- **Leverage data space supply chains**

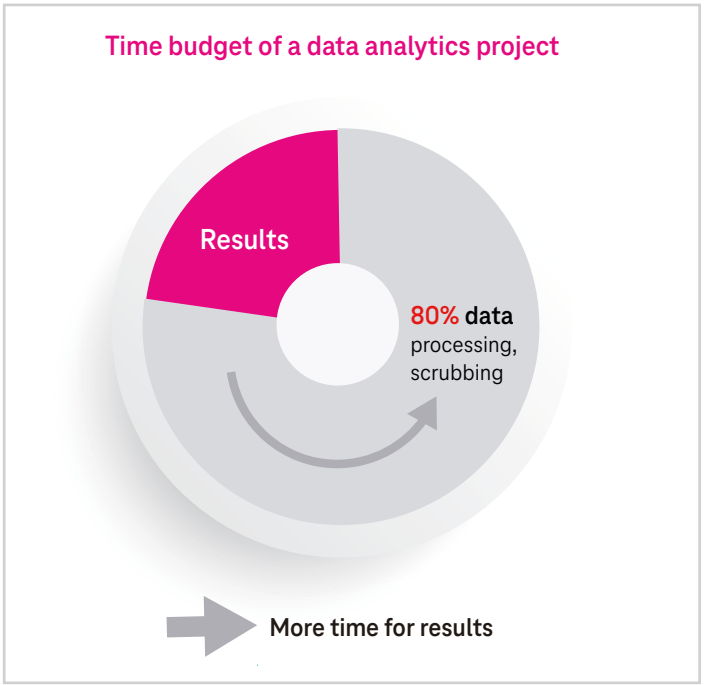


Figure 3: Pareto efficiency in reverse – 80 percent effort, only 20 percent results

Source - *Figure 2: Chris S. Langdon, Drucker Customer Lab, CGU, data from Handelsblatt (2019) and IDC (2020)
*Figure 3: Meta research by Prof. Dr. Chris S. Langdon, Peter Drucker School of Management, Claremont Graduate University, 2019, https://research.cgu.edu/drucker-customer-lab/

1. Treat data as a product: how do you even measure it, let alone monetize it?

Our first theme builds on the quantitative empirical description of the data productivity dilemma illustrated in Figure 3 and traces it to a bigger management problem. As Peter Drucker, the founder of modern management, wrote decades ago in McKinsey Quarterly: "We cannot put on a computer what we cannot quantify. And we cannot quantify what we cannot define" (Drucker 1967). Drucker was highlighting key characteristics of a product as "an article refined for sale" (Oxford English Dictionary). So, conceptualize and treat data as a product and apply the modern view of product management as a multi-functional, continuous-loop, customer-centric discipline. We know to buy eggs by the dozen, butter by the pound, and milk by the gallon. But how do we buy or sell data? By byte or by the length of a time series or share of population? Here it is, we dream of data monetization, yet are already struggling with supposedly straightforward issues such as sizing data and deciding how to measure the quantity of data. One start would be to treat data as products such as eggs, with attribute categories and grades, and assign product managers, for example (Crosby u. Schlueter Langdon 2019). We see tendencies in OEMs to adopt the data as a product mentality. Data from the cars are made accessible to other lines of business in the organizations to shape new business opportunities.

Find out more here in "Data: How to measure it?" (link), and "Data quantity or quality?" (link).

2. Industrialize data: create data factories for data products

Our second theme highlights the importance of industrializing data. Creating data products and then economizing on them with data factories is seen as one crucial step toward the wider adoption of advanced data analytics in business (Schlueter Langdon & Sikora 2020). The Swabian engineer, Gottlieb Daimler, invented the motorcar with the patent awarded in 1886, but it was the American businessman Henry Ford who invented the automotive business with the introduction of industrialization a few decades later (Womak et al. 1990). The need for extensive, industrialized data refinement or productization is also a theme in Humby's famous 2006 quote of "data is the new oil" (Arthur 2013). Indeed, oil refining entails massive refinery installations operating at a vast scale to make refined products such as gasoline affordable, and it is a multibillion-dollar business that is highly concentrated – a parallel that can already be observed with the cloud computing and storage business. Figure 4 provides an illustration of emerging data factory solutions, that help to manage data rights and facilitate data ingestion, data classification, data quality scoring and data governance. What we see in the Automotive industry are still separated "data factories" mainly from the car fleets adhering to the logic of figure 4, but which are not yet fully connected. The potential of interoperability remains so far untapped.

Find out more here in "Creating data factories for data products" (link); and "Success with 'small data' through specialized analytics" (link).

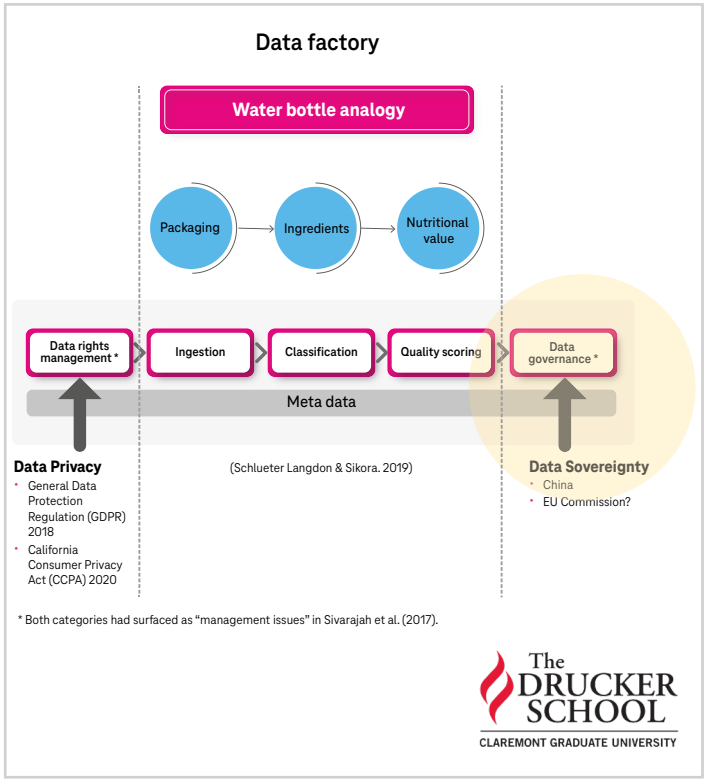


Figure 4: Data factories for data products

3. Build data space supply chains: the sovereignty solution to unleash free data flows

What is a data space? Why is it useful? When German Chancellor Angela Merkel proclaimed data spaces as a top priority for the German auto industry at the 2020 "Autogipfel" automotive summit even industry observers were caught off guard (Delhaes 2020).

Today, first data spaces have emerged or are under construction, such as the Mobility Data Space (MDS) and Catena-X. The latter is an industry-driven project with public funding, which was announced by the German Federal Minister of Economic Affairs and Energy together with the President of the German Association of the Automotive (VDA) and select automotive CEOs (more about these two data spaces and other examples in Part II).

From a management point of view think of data spaces as a supply chain for data. It can be orders of magnitude more efficient than traditional data relationships because it can replace one-to-one contracts with one-to-many or club arrangements. It's a bit like a supermarket: producers can sell to all consumers; consumers can buy all products – all without lengthy negotiations. Technologically, a data space is a U-turn in how data is made available to data consumers like software applications. It presents a shift away from centralized data storage to federated data. Instead of collecting and pouring all the data for an app into a central reservoir or lake, it may be better, less complex and more flexible if we could access it on demand.

While new technology, such as early Web services and today's Docker & Kubernetes, has increased flexibility at the application layer, little has happened with the data layer beneath (Schlueter Langdon 2006, 2003). A data space setup could now do the trick at the data layer. It would even cut the Gordian knot of data sharing between parties that do not trust each other, such as competitors, because it protects data sovereignty.

What is data sovereignty exactly? In short, data sovereignty, is the power to retain control over one's own data, which is often lost in data sharing, today. As soon as a file is sent, anything can happen to it. Data space technology, such as the International Data Spaces (IDS) standard, can help here (see Otto et al. 2019). Even if two parties do not trust each other, such as competitors, for example, can engage in a data transaction because data sovereignty is protected with (a) access controls, (b) usage controls, and (c) policy enforcement.

Read about our latest data space projects:

- "Welcome to 'data spaces': Storage is secondary – IT Director interview with Prof. Chris S. Langdon," link
- "Catena-X with GAIA-X: Will Data Space Be the Word of 2021?" link

Data spaces are the last key piece of the data economy puzzle. In summary, we propose three themes to unleash new revenue and growth with data: (1) Treat data as a product and apply proven product management practice, (2) create data products and industrialize the process using factories, and (3) leverage data spaces as a supply chain for data.

Stay posted for Part II: It will dig deeper, lift the hood and look into the engine compartment so to speak, it will (a) dissect first data space examples, such as the Mobility Data Space and Catena-X, explain key building blocks and essential services, and (b) extend sovereignty beyond data to include cloud sovereignty.

"As a technology partner we stand for a safe data ecosystem. Open for all companies of the European automotive industry, as well as their partners, users and suppliers. A network, that stands for data-sovereignty, whilst maintaining provider- and technology openness.

(Timotheus Höttges – Chairman of the Board of Management, Deutsche Telekom AG)

Conclusion

In conclusion, we see how data analytics is making inroads in the automotive sector and the painful productivity challenge surrounding it. This paper demonstrates how to slice and dice it, to navigate and organize it successfully. Instead of reinventing the wheel or experimenting with green field ideas, this paper is leaning on and mapping the data productivity challenge on to tried and true concepts, specifically product management (treat data as a product), factory automation (create data factories for data products), and supply chain orchestration (use dataspaces with data sovereignty protection). As dataspaces are a more novel concept, in part two, we will dive deeper into it, using lessons learned from the world's first dataspace installations in the industry. We are looking forward to sharing it with you. Please get in touch with us for any questions and feedback.

References

Arthur, C. 2013. Tech giants may be huge, but nothing matches big data. The Guardian (August 23), link

Crosby, L., and C. Schlueter Langdon. 2019. Data is a Product. American Marketing Association Marketing News (April), link

Crosby, L, and C. Schlueter Langdon. 2017. Smart Technology. American Marketing Association Marketing News (Nov-Dec), link

Delhaes, D. 2020. Merkel urges automotive groups: BMW, Daimler and VW to share data. Handelsblatt (2020–10–28), link

Drucker, P. 1967. The manager and the moron. McKinsey Quarterly (December), link

Otto, B., S. Steinbuss, A. Teuscher, S. Lohmann, et al. 2019. Reference Architecture Model Version 3.0. International Data Spaces Association, Berlin, link

Kaas, J. H. 2013. The evolution of brains from early mammals to humans. Wiley Interdisciplinary Reviews: Cognitive Science Wiley 4(1): 33–45, link

Kaas, J. H. 1992. Do humans see what monkeys see? Trends in Neurosciences 15(1): 1–3, link

Schlueter Langdon, C. 2020. Simulators in Business: Testing the impossible, discovering the probable. Working Paper (WP_DCL-Drucker-CGU_2020–05), Drucker Customer Lab, Drucker School of Management, Claremont Graduate University, Claremont, CA, link

Schlueter Langdon, C. 2006. Designing Information Systems Capabilities to Create Business Value: A Theoretical Conceptualization of the Role of Flexibility and Integration. Journal of Database Management 17(3) (July-September): 1–18

Schlueter Langdon, C. 2003. The State of Web Services. IEEE Computer 36(7): 93–95, link

Schlueter Langdon, C., and R. Sikora. 2020. Creating a Data Factory for Data Products. In: Lang, K.R., J.J. Xu et al. (eds). Smart Business: Technology and Data Enabled Innovative Business Models and Practices. Springer Nature, Switzerland, link

Womack, J., D. Jones, and D. Ross. 1990. The Machine that Changed the World: The story of Lean Production. Free Press, Simon & Schuster: New York, NY

Contact

Dr. Christian Hort
Senior Vice President Automotive
T-Systems International GmbH
christian.hort@t-systems.com

Prof. Dr. Christoph Schlueter Langdon
Data Analytics Executive and Scientist
T-Systems International GmbH
christoph.schlueter-langdon@t-systems.com

Publisher

T-Systems International GmbH
Hahnstraße 43d
60528 Frankfurt am Main
Germany

T Systems

Let's power
higher performance

Last updated: 09/2022 | Designed by: Creative Factory, TSI