

End of previous Forum article

Barbara Engels

## Detours on the Path to a European Big Data Economy

Big data is about acting on a large scale and ignoring national boundaries. As envisioned by the European Commission, national boundaries are supposed to disappear under the Digital Single Market. While the European Union could indeed become a large market for big data, it is not there yet. Many small and medium-sized companies (SMEs) are not aware of the significant potentials and promises of big data for their business activities and models.<sup>1</sup> They have good reasons for their ignorance, among them the uncertain regulatory situation in Europe. Politicians on all levels should give more attention to the wants and needs of SMEs, since they are vital to the economic performance of all EU member states. SMEs make up 99.8% of all enterprises, 57.4% of value added and two-thirds of employment. They form the backbone of the European economy. If they could tap the potential of big data, this could really be the next big thing.

### Big data – big potential

Big data has a wide spectrum of advantages for companies of any size and sector, which is oftentimes mentioned

<sup>1</sup> SMEs are defined as businesses which employ fewer than 250 employees and have an annual turnover of less than €50 million and/or a balance sheet total less than €43 million. See [http://ec.europa.eu/growth/smes/business-friendly-environment/sme-definition/index\\_en.htm](http://ec.europa.eu/growth/smes/business-friendly-environment/sme-definition/index_en.htm).

**Barbara Engels**, Cologne Institute for Economic Research, Germany.

but still not properly understood. One way to promote understanding of the importance of big data is to re-emphasise the benefits it could offer.

Particularly in the context of Industry 4.0 or the Industrial Internet, data constitutes the engine of economic growth and innovativeness. Companies can store large amounts of data digitally and analyse and process it “on demand” using real-time algorithms. Recognised patterns help optimise processes, develop a better understanding of target markets and potential future target markets, and improve forecasts of supply and demand as well as of the competitive situation. This enables decision-making that is more informed, faster and more efficient, which helps even small companies effectively position themselves in international markets. Companies can develop new business models and optimise their products and services according to the current and potential future needs and wants of their customers. They can become innovators and tremendously increase their productivity, their responsiveness and eventually their profitability.

And there is no end to this big data transformation: progress in machine learning procedures has led to significant improvements in core technologies such as speech and face recognition, translation, and image interpretation. These are all functions that could be integrated into companies’ business models. The analysis and use of big data is a crucial strategic factor for companies in the digitalisation process. Big data is a must in a digital world.

Despite the fact that there is no alternative to digital transformation, many SMEs are still reluctant to develop a digital strategy. According to an analysis of 46 digitalisation studies by the Cologne Institute for Economic Research (IW Köln),<sup>2</sup> the degree of digitalisation of German SMEs is difficult to assess. Many different measures and perceptions are relied upon, depending on the study – which in turn depends on the agenda of the institution responsible for the study and the message it is trying to convey. Digitalisation certainly does not make sense everywhere and always. What is undisputable, however, is that there is still a lot of untapped potential – and digitalisation as such is never finished, as it is a highly dynamic and self-innovating process.

According to the IW Köln analysis, the main challenges or barriers to digitalisation as seen by German SMEs are IT and cyber security, high costs and a lack of know-how – half of the analysed studies name these factors as primary obstacles.<sup>3</sup> A third of the studies attribute lagging digitalisation to missing standards. A fifth of the studies claim that companies are hesitant to fully pursue digitalisation because the regulatory framework is unclear and is constantly in flux – companies cannot be sure of which laws and rules they have to follow if they employ digital technologies.

The broader use of big data clearly faces quite similar obstacles. In the following, some light is shed on how these barriers keep small companies in particular from realising the benefits of big data.

### Obstacle 1: Lacking data security

Many companies are reluctant to fully explore the potential of big data because they feel they will not be able to protect the data they generate, process and interpret against misuse and other breaches. In fact, many companies see cybercrime as the biggest threat to companies in Germany.<sup>4</sup> They have a valid argument, as more than one in three companies in Germany have been affected by cybercrimes such as data theft over the last two years.<sup>5</sup> According to another study, more than two out of three industrial companies (69%) in Germany were victims of data theft,

economic espionage or sabotage in 2014 or 2015.<sup>6</sup> What is more, there is a high number of unknown cases, since many companies do not report cybercrimes to official institutions, either because they do not realise that they have been attacked (or the extent to which they have been attacked) or because they fear reputational damage.<sup>7</sup> Small companies in particular do not have the know-how or the resources to identify potential data leaks and security problems. In contrast to large corporations, they usually do not have a holistic approach to security; however, they nonetheless have confidential information – or access to it – which makes them an interesting target for hackers.

The more data that companies handle, the more attractive they become to exploitation by competitors and foreign governmental services. This attractiveness can destroy companies, as cybercrime entails enormous costs. In Germany alone, the costs run to tens of billions of euros per year. In addition to the direct costs related to the loss of intellectual property and the abatement of the immediate damage, there are various indirect costs such as sales losses and reputational harm. The loss of customers' personal data is especially costly in terms of reputation. It is essential to understand cybersecurity as a prerequisite for all entrepreneurial and commercial activities in order to protect the competitiveness and stability of the digital European economy. In other words, if we cannot protect confidential data – to a certain degree, since there can be no guaranteed security – then we should reconsider whether to use big data at all. The security architecture of companies must take the security of these new data streams into account.

### Obstacle 2: High costs

Another major obstacle to digitalisation and hence big data usage are high costs. Innovations require financial investments, which entail the risk of investing in the wrong changes. As digitalisation is highly dynamic, more and more wrong decisions can be made in an ever-shorter period. What is more, if companies take digital transformation seriously, they cannot invest in a selective fashion. Instead, the need for investment in digital technologies – and hence in the generation and analysis of big data – often extends to the entire value chain. It entails not only the innovations themselves, but also accompanying measures such as

2 V. Demary, B. Engels, K. Röhl, C. Rusche: Digitalisierung und Mittelstand – Eine Metastudie, IW-Analyse No. 109, 2016, available at <https://www.iwkoeln.de/studien/iw-analysen/beitrag/vera-demary-barbara-engels-klaus-heiner-roehl-christian-rusche-digitalisierung-und-mittelstand-eine-metastudie-312107>.

3 Ibid.

4 Allianz: Allianz Risk Barometer – Die zehn wichtigsten Geschäftsrisiken 2017, Allianz Risk Pulse, 2017, available at [https://www.agcs.allianz.com/assets/PDFs/Reports/Allianz\\_Risk\\_Barometer\\_2017\\_DE.pdf](https://www.agcs.allianz.com/assets/PDFs/Reports/Allianz_Risk_Barometer_2017_DE.pdf).

5 KPMG: e-Crime in der deutschen Wirtschaft 2017 – Computerkriminalität im Visier, 2017, available at <https://home.kpmg.com/de/de/home/themen/2017/04/e-crime-studie.html>.

6 Bitkom: Datendiebstahl, Spionage und Sabotage in der Industrie, 2016, available at <https://www.bitkom.org/Presse/Presseinformation/Industrie-im-Visier-von-Cyberkriminellen-und-Nachrichtendiensten.html>.

7 B. Engels: Wirtschaftliche Kosten der Cyberspionage für deutsche Unternehmen. Cybersicherheit als Grundvoraussetzung der digitalen Transformation, IW policy paper No. 6, 2017, available at <https://www.iwkoeln.de/studien/iw-policy-papers/beitrag/barbara-engels-wirtschaftliche-kosten-der-cyberspionage-fuer-deutsche-unternehmen-341100>.

cybersecurity and staff training. At the same time, the actual investment requirements are difficult to estimate, particularly for SMEs. Profitability is often unclear because of substantial network effects, since one company's digitalisation-related revenues are dependent on the degree of digitalisation of other companies. Company A's return on big data investment is likely to be dependent on company B's digital adoption.

Big data has several cost factors, namely its generation, storage, processing, analysis and interpretation. Data generation often happens "on the go" – for example, customers buy products and leave traces of their behaviour. The data is more or less the by-product of commercial activity; its "production costs" are hence negligible. Less negligible is the cost of transmitting and exchanging data. High-performance broadband networks need to be established all across Europe, especially in rural areas, in order to allow for real-time communication and scalable, efficient data transmission.

Also less negligible are the costs for the storage of data. Most SMEs do not have the capacity to store large amounts of data. However, cloud computing services continue to grow, offering ever-safer storage opportunities at ever-lower costs. The primary challenge lies not in the costs but rather in determining which provider to trust – as well as in potential data security problems, a recurring theme in this paper. Assuming a similar set of data protection standards, many European companies would prefer to store their data on a European server. In fact, studies show that one of the most important criteria for German companies when selecting cloud services is that the company data is stored exclusively at data centres in Germany.<sup>8</sup> The processing, analysis and interpretation of data is the largest cost factor in the big data industry. It is closely correlated with the third most significant digitalisation obstacle according to the IW Köln study – missing know-how.

### Obstacle 3: Missing know-how

Big data is not smart data. SMEs need to be able to generate sensible and usable information from the massive volume of data their business activities generate. Processing and distilling big data into smart data is a major task that could overwhelm small companies in particular. After all, data can easily be incorrectly interpreted, particularly when prediction, correlation and causation are mixed up. Statistically, it is easy to detect a non-existent correlation as long as the number of data points is large enough. Attempting

<sup>8</sup> See, for example Bitkom: Cloud Monitor 2017, <https://www.bitkom.org/Presse/Presseinformation/Nutzung-von-Cloud-Computing-in-Unternehmen-boomt.html>.

to make sense of a huge amount of data is like trying to take a sip from a firehose. Data scientists, who are said to have the "sexiest job of the 21st century",<sup>9</sup> are unaffordable for SMEs and, due to their scarcity, represent a massive investment for larger companies.

To actually make use of big data, SMEs therefore have to buy data analysis solutions in a plug-and-play fashion. Especially in the short term, it is sensible and efficient to outsource big data analysis to experts. Many start-ups have taken advantage of this and have specialised in the data business. However, apart from the ability to gather, distil and interpret data, companies must be able to recognise the specific opportunities available to them, which is hard to acquire and cannot be outsourced. Generally, digital technologies can only be fully exploited if the employees are empowered to properly use them. On-the-job training is crucial.

### Obstacle 4: The lack of common standards

Big data means that information from many sources and players is gathered and exchanged. A prerequisite for this data exchange are standards, i.e. a common language that all players can communicate in. If data cannot be transferred or processed for technical reasons, this has negative effects on the economic potential of the data. To facilitate data usage and improve its technical and semantic interoperability, common metadata schemas and standardised vocabularies must be developed. Standardised interfaces can facilitate a seamless exchange of information. Currently, however, missing standards are still a significant obstacle to digital transformation.

A recent study by the IW Köln involving 1200 industrial companies in Germany finds that most German SMEs consider common standards important for the digital transformation of their value chain.<sup>10</sup> Nevertheless, they have generally introduced such standards only if their strategy is already geared toward digital transformation. Companies that do not have a digital strategy are unlikely to introduce standards. By contrast, digitalised companies have discovered the potential of standardisation, but they are still struggling with obstacles that are mainly due to the lack of standardisation among their customers and suppliers – i.e. companies that do not have a digital strategy yet. For them,

<sup>9</sup> See T.H. Davenport, D.J. Patil: Data Scientist: The Sexiest Job of the 21st Century, in: Harvard Business Review, October 2012, available at <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>.

<sup>10</sup> B. Engels: Bedeutung von Standards für die digitale Transformation, IW Trends No. 2, 2017, available at <https://www.iwkoeln.de/studien/iw-trends/beitrag/barbara-engels-bedeutung-von-standards-fuer-die-digitale-transformation-339509>.

the advantages of standardisation are often not clear: if they introduce standards at all, they do it mainly because their customers demand it and not out of strategic considerations. Standardisation hence has to become more of a strategic factor, as this will also help to maximise the benefits of big data for SMEs.

### Obstacle 5: Regulatory framework

An uncertain regulatory framework is another entry on the list of obstacles to digitalisation. SMEs cannot tap the full potential of digitalisation and hence big data if they cannot trust the regulatory environment in which they operate. In particular, they need to be able to move, exchange and merge data freely across national boundaries. Ideally, laws concerning data generation, processing and analysis, such as data protection standards, should be identical in every country that the companies want to cover. In particular, it would be appropriate to make the European Union a single big data market that adheres to a common and standardised set of regulations.

In July 2014, the European Commission stated that public consultations had identified a lack of cross-border coordination and a fragmented and overly complex legal environment as two of the main problems with big data.<sup>11</sup> With the General Data Protection Regulation (GDPR), which came into effect in May 2016, the legal environment was supposed to become clearer. And it has, in a way: the GDPR significantly contributes to the free flow of data, since it harmonises data protection across all member states. The harmonised legal framework is an essential prerequisite for databased business models. However, there have been attempts to disturb this harmonisation. National laws that deviate from the GDPR, as well as existing data localisation rules, which require data to be stored within a certain country, counteract the European big data economy and should therefore be examined by the European Commission. A patchwork of differing regulations has to be avoided by all means – otherwise the Digital Single Market will remain a daydream.

Regardless of the advantages and disadvantages the GDPR might imply for big data usage, and no matter the degree of harmonisation that is eventually achieved, there is still a gap between this legislation and business reality. Many companies have not yet adapted to the new legal standards. A fifth of the IT companies surveyed in a recent

study have ignored the GDPR thus far, although beginning 25 May 2018 there will be high fines for breaches.<sup>12</sup>

This finding seems to be symptomatic of the mutual ignorance of the GDPR and European companies. The GDPR is concerned with personal data and hence ignores a large collection of data that is crucial to business models: machine-generated, non-personal data, such as production data. Only recently, with the “Building a European Data Economy” initiative, has the Commission started to eye machine-generated data,<sup>13</sup> which it wants companies to share among each other. The rights to the access and use of non-personal data are barely regulated at present, which results in legal uncertainty and could hinder the development of a European data economy.

The European data economy can only develop its economic potential if as much of the available data as possible can be used by as many interested parties as possible. A large number of players have an interest in using data from and about machines and products, regardless of their position in the value chain. The prerequisite for data-sharing is fair competition between all market players – including SMEs and start-ups. It has to be closely monitored whether discriminatory access to this data has led to market failure. “Codes of conduct”, established among companies of a digital ecosystem, may prevent potential market failure. In addition, the EU Commission should make use of European antitrust and competition law to take decisive action against developing data access monopolies at an early stage.

### Outlook

Big data is a big deal for the European economy because it can yield benefits for companies of all sizes. It is not an elite project whose advantages are constrained to successful large companies. On the contrary, large ships cannot reach small bays. Indeed, large companies often lack agility due to their internal hierarchy and sedate strategies. SMEs are oftentimes more agile and have more flexible infrastructures and decision-making processes. They can potentially turn data into new business ideas and therefore into profits more quickly than their larger competitors. However, they cannot do it on their own. They need help from others in order to be able to extract relevant information from the data and keep it well protected. Just as networks are the main component of digitalisation, SMEs should make networking a priority and team up in order to be big in big data.

<sup>11</sup> European Commission: Commission urges governments to embrace potential of Big Data, Press release, 2 July 2014.

<sup>12</sup> See <https://www.bitkom.org/Presse/Presseinformation/Jedes-fuenfte-IT-Unternehmen-ignoriert-bislang-Datenschutzgrundverordnung.html>.

<sup>13</sup> See <https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy>.