BIG DATA IN SMALL STEPS

TNO innovation for life

Data is seen as the new oil: an important driver of innovation and economic growth. At the same time, many find it difficult to determine the value of big data for their organization. TNO presents a stepwise big data model that supports private and public organizations to assess the potential of big data. After identification of the key challenges to deploy big data, this paper also presents organizational and technical capabilities that address these challenges.

VALUE OF BIG DATA

The importance and ubiquity of digital data available to public and private organizations has created the firm belief that 'big data' will inevitably transform most of today's businesses. We define big data as the use of large, varied, real-time, or unstructured data for innovation or to create value for organizations. It provides the opportunity to do and know things that were not possible before. The emergence of affordable powerful data analytics and storage tools enables organizations to maximize work processes, conduct fine-grained analyses of customers and markets, and personalize and optimize products and services. But now that big data technologies and tools have become mainstream, the actual value of big data is determined by the exploitation strategy organizations deploy to create economic or societal value.

Many organizations struggle to develop a coherent strategy for deploying big data. The fundamental rethink of current business propositions, target groups or even market segments that is needed to successfully deal with the advent of big data, requires a multidisciplinary and coordinated approach. TNO therefore presents a model of big data developments, which serves as inspiration for developing a big data strategy. Based on an exploration of state-of-the-art big data developments, four phases of data-driven innovation and associated value propositions are identified: efficiency, effectiveness, new solutions and transformation. The model in this paper not only helps organizations to assess their own big data maturity and innovation potential, but also points out the most pressing and significant organizational and technological challenges that need to be addressed.

The *datafication* of society

- ✓ *Tweets generated every day*
- ✓ *Patient health information*
- ✓ *Customer relations systems*
- ✓ *Sensor data for traffic management*
- ✓ *Open government data*
- ✓ *Search engine data*
- ✓ *Satellite imagery*
- ✓ *DNA sequences*

A BIG DATA MATURITY MODEL

To capture the value of big data for organizations, we present a maturity model for big data developments. Based on an exploration of existing big data developments, this model shows the growth paths that may result from exploiting data. Consisting of four phases, the model can be used to show the subsequent steps organizations have taken (descriptive) or may take (prescriptive) when using big data. Organizations can use this to determine their own position and growth path. Assessing the current and preferred phase within the model leads to a set of necessary actions.

The key determinant to position organizations within the model is the organization's capability – or readiness – to exploit big data to create economic or societal impact – or both. Although in the more 'mature' phases the potential of big data is larger, at the same time organizational and technological complexity also likely increase, as will be discussed in more detail further on. Organizational complexity increases from mere cost cutting activities to (re)identifying core business within one organization and from the delivery of new solutions to shifting roles within a network of organizations. Such organizational challenges commonly go hand in hand with an increase of technological complexity, which increases with the combination of different data sources, for example combining large, real-time, varied and unstructured data.

EFFICIENCY

The most common application of big data analytics is an extension of traditional business intelligence, delivering more detailed and faster availability of relevant insights. Its main goal is cost reduction. It is primarily directed at gaining insight in and improvement of internal organizational practices such as more efficient production processes, resource allocation and financial monitoring. These innovations can be lucrative, but the value they generate primarily benefits the organization and its shareholders. Although this kind of data-driven innovation is often used for internal processes within an organization, it can also be used for optimization of inter-organizational processes.

Example of *Efficiency*

The IJkdijk, a 'smart levee', is the result of a shared research program in which a dike in the north of the Netherlands was equipped with sensor technology. The collected data can be analyzed and visualized to improve dike monitoring and water management. The smart application of the monitoring philosophy leads to significant cost savings and deferred investments for water management authorities.

EFFECTIVENESS

A second line of innovation with big data within organizations occurs when focus shifts from cost reduction to creating new value. Its main goal is to support and (re)enforce the core business of the organization. Examples are often found in marketing and sales, product development, and strategic decision-making. The explosion of data has led to a multitude of data sources that increase insight into customers' behaviour and attitudes (social media have become indispensable sources) as well as market developments. Many organizations automatically target advertisements to (online) audiences based on individual profiles that have been assembled by integrating various data sources that contain information about (groups of) individuals. Additionally, building on these kinds of insights, big data has made it possible to move away from one-size-fits-all solutions that have been standard procedure in product development for a long time.

NEW SOLUTIONS

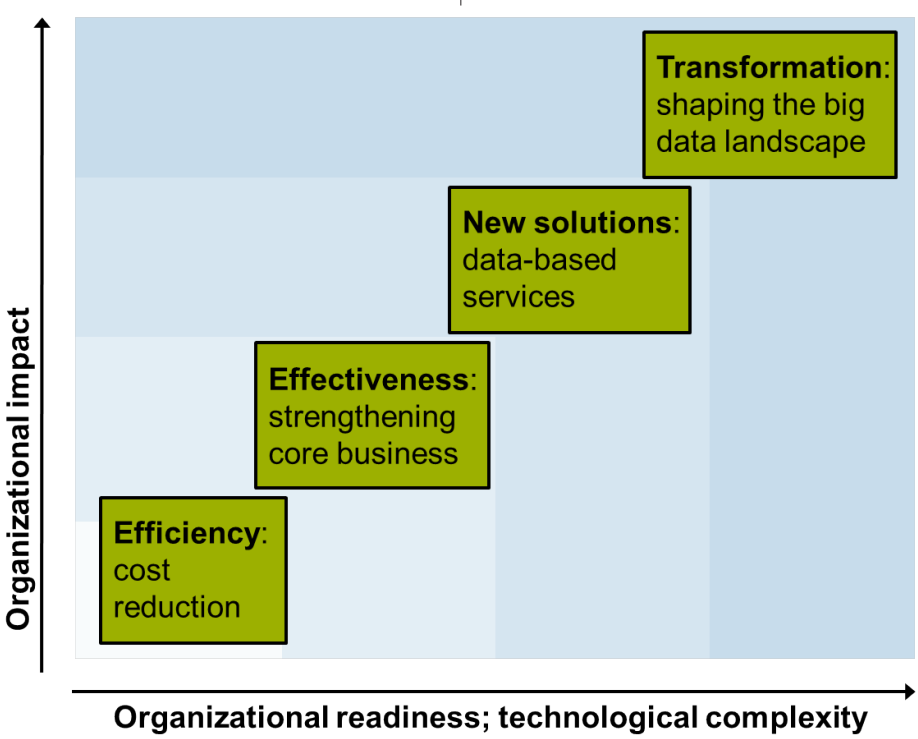
The datafication of products changes their nature and the way they are consumed, or how value networks are organized. In this phase of data-driven innovation, data constitutes a key ingredient of the product or service. The main goal in this phase is to develop new value propositions based on and containing data analyses for the customers of an organization. Data products are able to collect information and transform this information in to actionable output for the customer. Such data-driven innovations enable the creation of new products and services and offer new value to consumers. Data has thus become a vital part of the value proposition.

Example of *Effectiveness*

Walmart is a prolific user of big data to target its customers and stock and present products in its stores. It combines and crunches its proprietary data on customer behavior (e.g., purchasing history), combined with social media data and public data on the web, to create a so-called Social Genome of its customers for marketing purposes. Walmart knows exactly when it has to target what kinds of (combinations of) products to which customers. It can interpret what their customers are saying online about specific products, enabling them not only to reach the customer but also his friends, and to promote that exact product and offer a targeted offers in the process.

TRANSFORMATION

Organizations that are able to use data to redefine propositions can take pole position to reshape markets and profit from these changes. They could even



The big data maturity model

Example of *New Solutions*

Big data has a big impact on agriculture. Data analytics can, for instance, support farmers in deciding how to treat their crops in order to yield the best results. This idea inspired John Deere, a manufacturer of tractors and agricultural equipment, to enhance its most recent series of tractors with sensors that collect data from the machine, the soil and the crops it processes. The data that it collects is then analyzed and augmented with additional data about the weather and crop features, and presented on an online platform that can be used by farmers on both desktop computers and handheld devices.

set in motion systemic transformations (e.g., in agriculture and food, and lifestyle and health) that substantially distort the composition of whole ecosystems, forcing other companies to adapt as well. Organizations in this phase move across the network to incorporate, lay off or shift to other roles within the value chain. More often than not, there is a need for a completely new business model. As big data technologies become a commodity, more transformative innovations become increasingly important in terms of competitive differentiation.

ORGANIZATIONAL IMPLICATIONS

The four phases of data-driven innovation are progressive: higher levels require a better understanding of how data can be

Example of Shaping the data landscape

Nike has redesigned some of its products as 'data products'. It introduced the online Nike+ platform, the Nike+ sensor that can be clipped on running shoes, an app that tracks runs, and more recently the FuelBand, a wristband that tracks activities and burned calories. Using data collected by apps and sensors, Nike provides a multitude of information and services related to physical activities. Although its core value proposition – supporting people to be physically active and healthy – has not changed, Nike is increasingly shifting towards propositions centered around data that enable consumers to set physical goals and to track their personal progress while including social elements.

Companies need to be aware of the way data might impact the environment in which they act, which is the basis for phase four. Nike has created data-driven products and is leading the datafication of sports and active lifestyles. But with the massive adoption of smartphones and apps it also has to deal with new kinds of competitors as services that offer data-supported physical exercise, like RunKeeper or Runtastic, abound. In a new market of wearable technology Nike might have to compete with Google (Google Glass) and Apple that are also its partners for the use and distribution of its apps on the iPhone and Android phones. Nike also offers an API that allows trusted third parties to develop apps based on this data-driven platform.

used to capture value for the organization and more organizational readiness. And while value for the organization increases, organizational and technological complexity likely increase at the same time. Naturally, certain phases of data-driven innovation can be seen to take place simultaneously, for instance streamlining a production process (efficiency phase) while using data to develop new products (new solutions phase). However, the more mature phases generally pose more complex challenges as they include multiple stakeholders, various kinds of expertise and strategic decision-making.

Due to the multidisciplinary nature of the challenges that big data and big data-driven innovations pose, we stress the need for a set of organizational and technical capabilities that organizations need to possess in order successfully set up and deploy big data.

BIG DATA CHALLENGES

Despite the potential of big data, few organizations today are able to systematically integrate the collection, combination, storage and processing – thereby capturing the value – of data into their daily business. Most notably, as a result of concerns related to privacy, data quality, and liability, much data is intentionally left unused. Based on a TNO survey among professionals in the field of big data, we highlight the five challenges that were mentioned most often. Addressing these challenges often requires a combination of organizational and technical capabilities. Furthermore, it should be noted that while all challenges occur in every phase of the model, their manifestation is likely to differ and their complexity increases.

ORGANIZATIONAL CAPABILITIES

When an organization has determined its current and desired position, it needs to set up a strategy for achieving this

Challenges for deploying big data

- ✓ *Sense-making: decision-making, profiling customers, personalizing products and services, enriching and linking dataset and visualization.*
- ✓ *Organizational readiness: ensuring that big data is embedded in the work processes*
- ✓ *Ensuring data quality: making sure that data is of sufficient quality for the specific goal*
- ✓ *Enriching and combining of datasets: ensuring that it is possible to link data*
- ✓ *Ensuring privacy and security: particularly for the use of personal data*

position. This requires organizational capabilities that are attuned to big data deployment. Dynamic capabilities that aim for innovations, such as with big data, generally aim at three activities: sensing opportunities and challenges of big data, seizing these opportunities and develop a value proposition, and transforming these opportunities into results. When these capabilities are not in place, it is hard to develop a big data strategy.

TECHNICAL CAPABILITIES

In order to realize the value proposition offered by big data, organizations also need to have a thorough understanding of their current information technology (IT)

Organizational capabilities

- ✓ *Strategy: developing a value proposition for big data*
- ✓ *Alignment: embedding big data into the work processes*
- ✓ *Liability: making agreements on liability regarding data quality*
- ✓ *Partnering and community building: forging partnerships*
- ✓ *Limiting use of and access to data: using personal data for its specific purpose only, asking data subjects for permission, limiting access to data*

capabilities and be able to set clear goals for the future. In practice, this means that organizations need to develop their IT capabilities to be able to set up a big data strategy. Still, this does not mean that it is necessary for every phase of development

of a big data strategy to invest in, for example, Hadoop clusters, since a few small-scale pilots can also spur big data developments.

Technical capabilities

- ✓ *Data analytics: business intelligence, data mining, visualization*
- ✓ *Data management: to establish an effective control framework*
- ✓ *Data quality management: address meta and master data management*
- ✓ *Linking data: adding metadata, semantics, and setting up data platforms*
- ✓ *Safety: data security, anonymization and aggregation of data*

POSITION YOUR ORGANIZATION

This paper supports organizations that wish to determine the potential of big data for their organization and that need to devise a strategy to capture this value. With the increase of readily available technologies for big data analytics the value that big data will have for organizations is mainly determined by the way in which they are able to use data to support and strengthen their core business or develop new propositions. Key to developing a successful big data strategy is to deploy a combination of organizational and technological capabilities to address key challenges.

The TNO big data maturity model can be used as inspiration to develop a big data strategy: organizations can use it to position themselves and determine which actions they need to undertake to achieve their desired position. It is not necessary to move through these phases sequentially, nor does it imply that every organization should strive for achieving the most mature phase. For some organizations cost cutting can already be lucrative. Nevertheless, it may be useful to look at the later phases and understand how this will impact your organization and change the make-up of the market.

TNO's research team:
Anne Fleur van Veenstra, Jop Esmeijer, Tom Bakker, Thomas Bachet, Frank Berkers, Freek Bomhof, Silja Eckartz, Erwin Folmer, Wessel Kraaij en Joost de Wit.

CHALLENGES	ORGANISATIONAL CAPABILITIES	TECHNICAL CAPABILITIES
'Sense-making': data to information	Strategy development	Business intelligence, data analytics, data mining, visualisation
Organizational 'readiness'	Alignment with work processes	Data management
Ensuring data quality	Agreements on liability	Data quality management
Enriching and combining datasets	Partnering, community building	Linking data; setting up data platforms
Privacy and security	Limiting use and access	Data security, anonymization and aggregation

Organizational and technical capabilities addressing key challenges to big data deployment

TNO.NL

TNO is an independent innovation organization. TNO connects people and knowledge to create innovations that sustainably boost the competitive strength of industry and the welfare of society. TNO's more than 4000 professionals work on practicable knowledge and solutions for the problems of global scarcity.

Information Society: TNO researches the impact of the information society and stimulates the introduction of new services, applications and policy, based on the latest ICT, media and aerospace technology.

Anne Fleur van Veenstra
Brassersplein 2
P.O. Box 5050
2600 GB Delft
The Netherlands

T +31 88 866 27 79
E annefleur.vanveenstra@tno.nl