

# BIG VALUE THANKS TO BIG DATA?

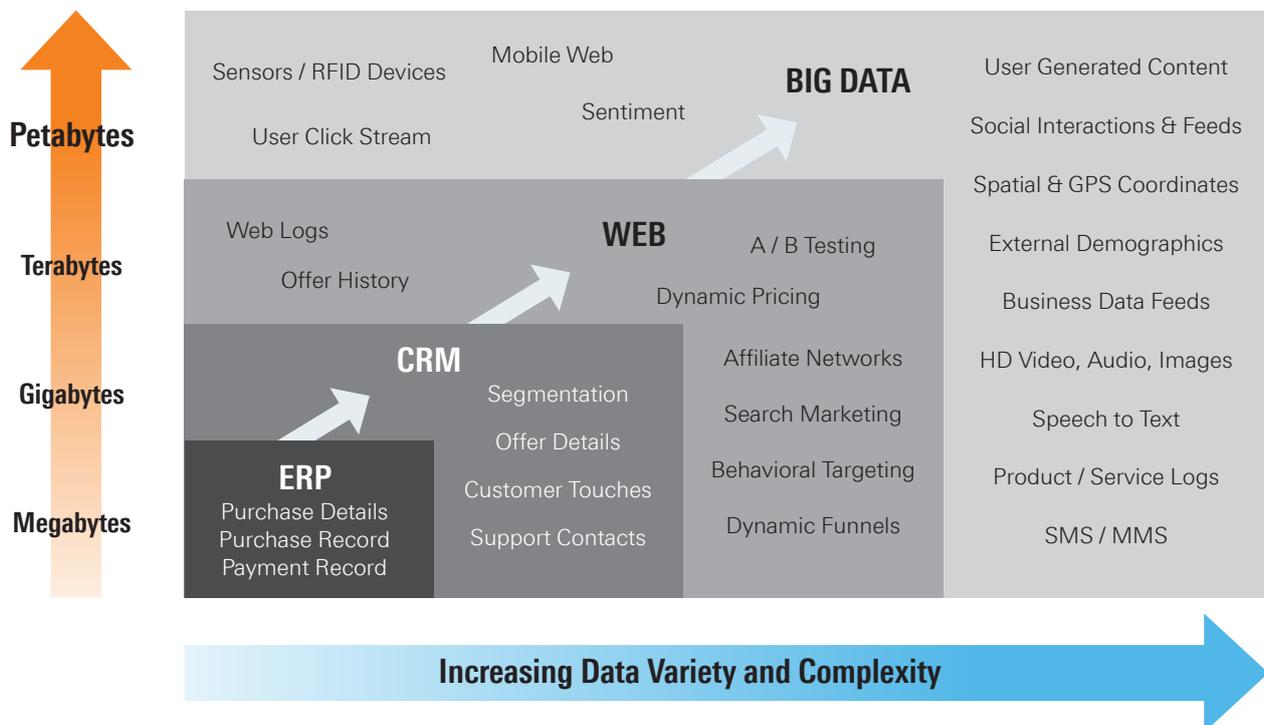
## PROF JOS VAN HILLEGERSBERG SHEDS LIGHT ON THE VALUE OF BIG DATA IN LOGISTICS

NOWADAYS, COMPANIES ARE ALMOST ENGULFED BY DATA. THE SECRET IS TO EXTRACT THOSE DATA THAT REPRESENT ADDED VALUE TO OUR ORGANISATION. DURING SUPPLY CHAIN INNOVATIONS, PROFESSOR JOS VAN HILLEGERSBERG TALKED ABOUT WHERE THE MOST IMPORTANT SOURCES OF BIG DATA CAN BE FOUND WITHIN SUPPLY CHAIN MANAGEMENT AND HOW WE CAN GLEAN ADDED VALUE FROM THEM FOR OUR BUSINESS. AS A PROFESSOR AT UNIVERSITY OF TWENTE AND CHAIRMAN OF THE PROGRAMME COMMITTEE OF TOP SECTOR LOGISTICS, HE STRESSES THAT BIG DATA FORMS AT THE VERY LEAST AN OPPORTUNITY WE CANNOT ALLOW TO PASS US BY.

The fact that big data mining is still rather tentative within supply chain management is certainly not an effect of the publicity this trend is receiving. "Much is written and said about big data. The number of suppliers offering big data solutions is growing too. However, I feel that companies are often unsure about how to initiate such projects and above all: what the added value for their organisation might be," says the professor.

He stresses that big data will continue to become bigger. He illustrates this statement through a telling figure, which compares the increasing variety and complexity of data with the volume of information we are presented with (see figure 1).

### BIG DATA = TRANSACTIONS + INTERACTIONS + OBSERVATIONS



Source: the graph above was created by Teradata

Figure 1

Prof Jos van Hillegersberg distinguishes between various data sources: “The major challenge is that more and more sources are emerging over the years. About ten years ago, the focus was on data from ERP systems, which were implemented in volume at that time. Those data are usually quite structured and manageable. The focus subsequently shifted more and more towards the customer, with plenty of CRM implementations as a result. This in turn created another layer of data around your ERP.”

The emergence of the Internet caused a terrific data boost. If only just from the large data pool formed by social media. “Consumers talk about companies and submit product reviews. As a company, you can extract very valuable data from that,” is the overall opinion. “In fact, we have noticed that more and more companies have also started carrying out their internal communication through social media channels, such as Yammer and Slack. The data that stem from this can be included in their analyses. Also, the continuously growing use of mobile apps bulks out the data stream.

**“THE MAJOR CHALLENGE IS THAT MORE AND MORE SOURCES ARE EMERGING OVER THE YEARS.”**

Prof Jos van Hillegersberg, University of Twente

Add to that the technological possibilities offered by Industry 4.0, which now means that large amounts of data are also being issued by sensors. Already there are more devices connected to the Internet than there are people who are active on the Internet, and this number will only grow.

“Tools that measure the behaviour of customers and employees can also generate valuable data. For instance, ‘gamification’ solutions, which are emerging in increasingly more environments,” adds Prof. van Hillegersberg. The firm Scania, for example, equips the on-board computers of truck drivers with a game environment, which should encourage drivers to drive more energy-efficient. This way, the driver can earn credits, which he can use to buy something nice. Product configurators - such as consumers who create a meal box on the Internet - can also provide interesting information about customer preferences.”

Last but not least, an increasing volume of open data is made available. Many authorities stimulate this trend. As a company, you could also consider opening up data sets, in order to contribute to efficiency improvements in the chain. Thanks to the onset of open data sets, we can make even better predictions within our supply chain.

Prof Jos van Hillegersberg: “Additionally, a considerable part of these data ends up in the cloud. The advantage is that it is easier to engage with and open up data in the cloud for other parties if you wish to carry out joint analyses. According to a study by AtScale on big data professionals, 14% percent already has all their data in the cloud, 13% most of their data and 26% some of their data. It is expected that in the near future 19% will have moved their data to the cloud, 23% will have moved most of their data and 30% will have moved some of their data. It follows that the volume of data in the cloud will rise sharply.”

It is positive that more and more data are within reach. But of course this fact does not make it easy to see the added value to our organisations. “What’s more, it’s not just the sheer amount of data you have to deal with; often these data are also continuously on the move. Think of streaming data, where a constant stream of new information is generated,” warns Prof Jos van Hillegersberg. “Also, the data are often very diverse in their nature: structured or unstructured, text, multimedia, GPS data, etcetera. The quality of all the information pouring in is not always reliable either. Even ERP systems often harbour ‘polluted data’. When you are working with external data sets, the risk of unreliable data is even bigger.”

## EXTERNAL RESEARCH INTO BIG DATA

The key question is whether it is worth diving into all those data for the benefit of creating added value for our companies. “I am convinced it is, as long as we approach it in the right way”, says Prof van Hillegersberg. “It may be a consolation that we are by no means the first to take the leap. By now, many companies have already started working with big data, albeit with varying degrees of success. We can learn from those projects.”

In order to gain an insight into the preliminary experiences with big data, Jos van Hillegersberg uses the Big Data Executive Survey 2017 as a reference. “For the purpose of that study, 50 CEOs of Fortune 1000 companies have shared what they do with big data. So, a relatively limited study, which nonetheless paints a clear picture of how large companies have advanced,” he says.

Almost half (48%) indicated in the survey that they have achieved measurable results from big data. Cited as the most important stumbling blocks for working with big data is the lack of a structured organisation (42.6%). The top also thinks that big data initiatives are held back, often by middle management (41%). Technological issues often are a hindering factor too (27.9). What’s more, the absence of a coherent data strategy (29.5%) and the lack of a shared vision surrounding big data (26.2%) are also mentioned as obstacles, for instance.

Cost savings feature prominently (49.2%) if we zoom in on the results that are achieved through big data. According to Prof van Hillegersberg, it may be interesting to gather data sets with the above point of view



in mind, in order to detect and deal with inefficiencies within this scope. In many cases, big data also offers access to new innovative solutions (44.31%) or provides the impulse for developing new solutions or services (36.1%). Companies are less successful when it comes to using big data for future-proofing their business, developing new skills within the organisation, and creating a data-driven culture. This leads us to conclude that culture change as a feeding ground for big data projects forms a particular challenge.

“Research by the Dutch Chamber of Commerce into 4,100 SMEs shows that smaller businesses are rather put off by all that data. No less than 49% cannot see the value of it for their company. They also

tend to think that they simply do not have the skills to start working with big data. So for the time being, it seems that it will be particularly the larger companies that will take the lead. But I am certain that the smaller ones will follow,” adds Prof van Hilleegersberg.

## RESULTS FROM OWN RESEARCH

In order to familiarise home-grown companies with big data and as a result boost its application, University of Twente and Dinalog (Dutch Institute for Advanced Logistics) cooperated on intensive research into its potential. To this end, Prof Jos van Hilleegersberg elaborates on three projects that his department at University of Twente have carried out together with several companies and government bodies.

The first study was aimed at attempting to adequately predict the arrival times of seagoing vessels through datamining AIS data. AIS is a navigation system that is used to quickly locate and identify vessels. For instance, the AIS transmitter on a ship periodically sends out statistical information (including name, dimensions and type of vessel), dynamic information (including position, course, speed), and route-specific data (including destination) to other vessels. “During this project, it was shown that the sensors on those routes transmitted a lot of incorrect and inadequate data. With the help of machine learning algorithms, we attempted to fill in the missing data, so that we could arrive at reliable route patterns and ditto predictions. This was a very significant point of attention,” explains Prof van Hilleegersberg. “There is much legislation on the level of AIS datamining permitted. As these types of applications do not breach the vessel’s privacy, it is allowed. But the legislation surrounding privacy is definitely something that big data users need to take into consideration when crossing the boundaries of their own organisation.”

## A SECOND STUDY FOCUSED ON THE PHARMACEUTICAL INDUSTRY

“A substantial amount of - often expensive - medication is shipped around the world. It is often important that certain temperatures are not exceeded. Such consignments are also considerably prone to theft. Within the framework of this study, we looked at the type of valuable information which could be gleaned from sensors in order to strengthen such chains. For example, information about where a certain pallet is located, whether the temperature is within the permitted margins and if it is still following the planned route. One of our partners in this project was Panalpina, which equipped its warehouses with the necessary antennae infrastructure for the purpose of the project,” Prof Jos van Hilleegersberg clarifies. “When we started this project, we focused on monitoring. For instance, it could be decided to stop further transit of a consignment if certain temperatures were exceeded. At a later stage we noticed that it was also very interesting to use data sets for retrospective analyses. For instance, this made clear which parties caused the biggest frustrations in the chain.”

The most recent study related to a Dutch refrigeration and deepfreeze warehouse that received a lot of international consignments. “A major shortcoming here was that the trucks did not always arrive at the arranged time, which disrupted operations within the warehouse. Big data was used as a basis for predictions on which trucks would be late or early at this company”, says the professor. “In this case we were able to use the open data, which included all traffic congestion information made available by the National Waterways and Highways Agency. These data were linked to data about weather, for instance, which in turn is also open data. We then combined the data with the movement data of the trucks. But contrary to what the company suspected, weather and traffic jams were only a small part of the variation.

Much larger impact came from aspects such as driving times and resting periods, contract agreements and even the driver’s level. In other words, big data is interesting, but do not simply disregard the role of contractual agreements and the people behind the activities.”

### THINKING PROCESS FIRST, IT SECOND

A large pitfall in big data implementations is that it is easy to be seduced by sophisticated software and sensors. This is why Prof van Hillegersberg advises not to dive into the technology immediately. The message is: obtain insight into your data, draw up a properly functioning model and only then start searching for suitable software. First of all, you should be thoroughly aware of what exactly you want to do with big data.

Prof Jos van Hillegersberg: “In essence, there are three objectives. As the research involving the Fortune 1000 companies already showed, data can be principally gathered to detect and tackle inefficiencies. Here, the focus is on ‘operational excellence’. The focus can also be aimed at optimisation, of your logistical network, for instance. This assigns a more strategic role to big data. A third objective of big data could be a better understanding of your customers, for better anticipation of and response to their wishes.

#### THREE OBJECTIVES

- DETECT INEFFICIENCIES
- FOCUS ON OPTIMISATION
- BETTER CUSTOMER UNDERSTANDING

For a better initialisation of big data, Prof van Hillegersberg has created a canvas. It can be used by companies when brainstorming about the value of big data for the organisation (see figure 2). “It is important to continuously question yourself about which problems specifically exist and how they could be solved using big data. If you think big, a business case should be able to justify the necessary investments. The canvas also provides food for thought about the challenges, the data sources you wish to use, and - not insignificantly - the organisational changes this will require.”

<p><b>ANALYTICS TECHNIQUES</b> (Explore, Explain, Predict)</p> <p>-</p> <p>-</p> <p>-</p> <p>-</p> <p>-</p>	<p><b>QUESTIONS</b> (No overview, no insight, no foresight)</p> <p>-</p>	<p><b>BENEFITS</b> (how is value created)</p> <p>-</p> <p>-</p> <p>-</p> <p>-</p> <p>-</p>
<p><b>DATA SOURCES</b> (Internal, External, Open)</p> <p>-</p> <p>-</p> <p>-</p> <p>-</p> <p>-</p>	<p><b>CHANGES NEEDED</b> (People, Process, Organization)</p> <p>-</p> <p>-</p> <p>-</p> <p>-</p> <p>-</p>	<p><b>CHALLENGES</b> (Data availability and quality, skills, risks, regulation, ...)</p> <p>-</p> <p>-</p> <p>-</p> <p>-</p> <p>-</p>
Version:	Organisation:	Name:

Source: Big data canvas J. van Hillegersberg ©

Figure 2  
The big data canvas

## WIDE RANGE OF SOFTWARE

One question Prof van Hillegersberg often faces is about the best software to take on big data. “A very tricky question. It’s impossible to just focus on a few. There are many tools for analysing big data and more are constantly added. We recently reviewed more than 80 and when we were finished, more tools were already emerging. What’s more, the best option will depend on your type of data, your organisation, the people working with the data, etc.,” he explains.

Additionally, you have to realise that processing big data into bitesize chunks involves several steps. “A first step is about ‘cleaning’ the data and preparing it for further use. Only then are they ready for further analysis. This step is often left out, despite it being the lion’s share of the work, especially if you are faced with an abundance of data, for example from sensors. Such tools increasingly contain artificial intelligence, so that oddities in your data are detected more quickly. Tools within this domain include Alteryx, Trifacta and Paxata. Another good open-source tool is openrefine.org,” adds Prof van Hillegersberg.

The next phase could be data analysis and visualisation. “This can be done with the help of an enormous amount of tools, ranging in price from free to exorbitant. Well-known players in this field are SAS, Qlik, R, SPSS, Tableau, Statistica, Knowledge Studio and Google FusionTables. A well-known open-source solution which could be useful here is knime.org.”



Finally, we could apply artificial intelligence and datamining to our data. “On the one hand, it has the advantage that you do not need as much statistical knowledge for this, but on the other hand there is the risk that it means the value of the analyses is not assessed as well. Solutions such as IBM Watson analytics and Wolfram Alpha have many uses, but they add a black box into the mix,” warns Prof van Hillegersberg. “This domain also offers several open-source tools, such as rapidminer.com, which allows you to make predictions based on historical data.”

The range of solutions is growing very rapidly within all these segments. What’s more, the tools are getting cheaper, faster and more powerful all the time. “While only large companies could afford such big data tools ten years ago, they are also very accessible to SMEs nowadays,” says Prof van Hillegersberg. “Social media mining used to be very difficult, for example. Now there are various tools that allow you to simply extract what is being said about your company on social media by using keywords. The tool by the Dutch company open.coosto.com also easily lets you start sentiment datamining, enabling you to zoom in on negative or positive news. This is free to experiment with, in fact. You only pay if you want to dig deeper.”

However you proceed though, Prof van Hillegersberg warns that a data-driven organisation is not built in one day. “It requires steps to be taken, from initial, opportunistic steps to an organisation where employees systematically examine processes based on data analyses. Of course, this necessitates a steep internal learning curve, which can be created in various ways. There are lots of communities, or you could sign up for online courses. It goes without saying that you do not become a genuine data scientist who is in the know about the ins and outs of big data overnight. However, I am convinced that many professionals within the organisation would benefit from becoming proficient in part of those skills. You could also let loose teams of students on your data, during a hackaton for instance. Such initiatives are becoming more common. There are also platforms such as Kaggle, supported by people who feast on data science, which create analyses for you. In other words, there are plenty of ways to experiment with big data.”

**“I AM CONVINCED THAT WE ARE STILL IN THE EARLY STAGES OF BIG DATA WITHIN SUPPLY CHAIN MANAGEMENT”**

Prof Jos van Hillegersberg, University of Twente

“I am convinced that we are still in the early stages of big data within supply chain management”, concludes Prof Jos van Hillegersberg. “This could create very interesting perspectives, especially if several parties are prepared to combine their data. Within University of Twente, we consider it our task to contribute to new models which support such initiatives, as well as to chart all challenges in big data implementations from its periphery.”