



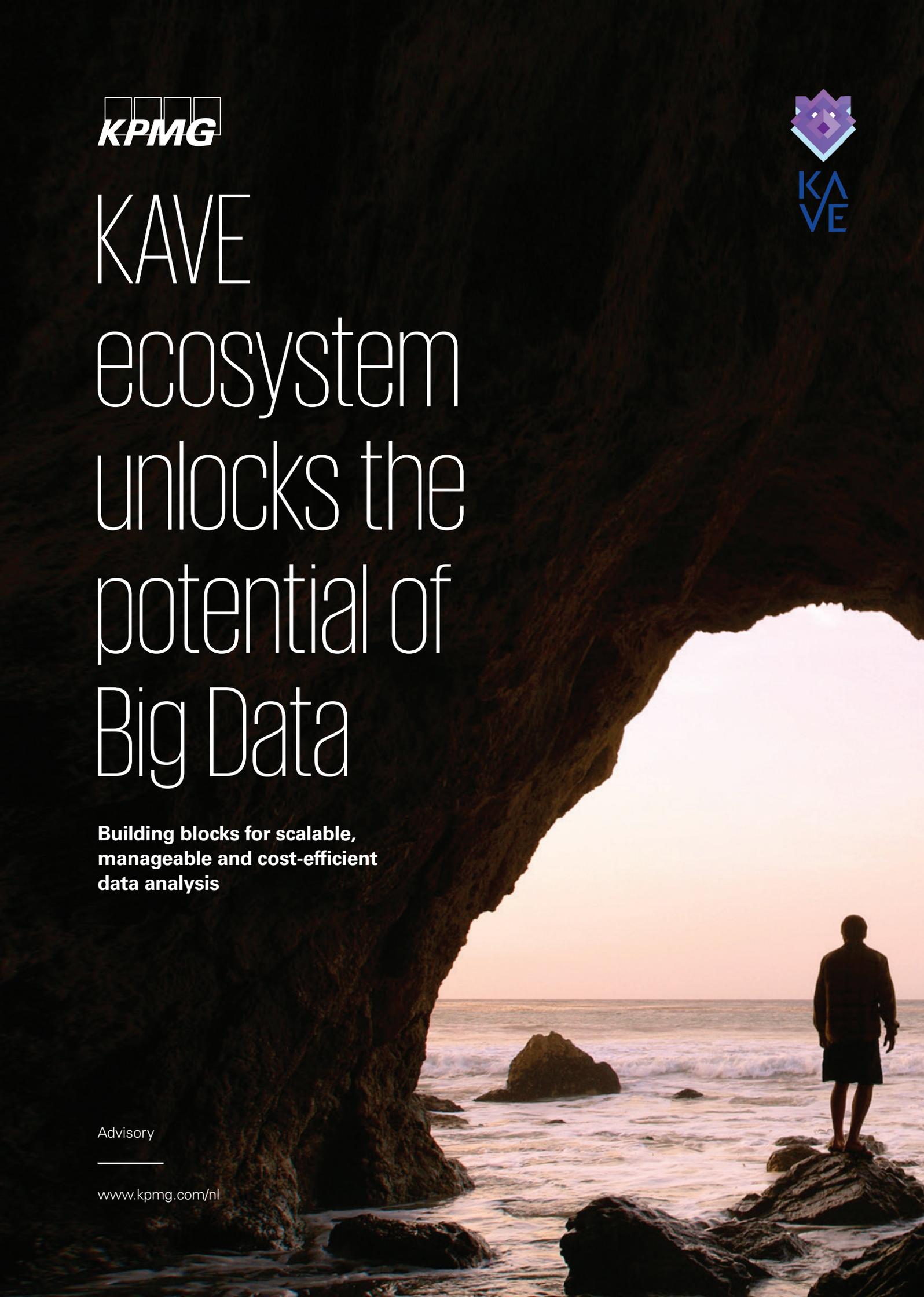
# KAVE ecosystem unlocks the potential of Big Data

**Building blocks for scalable,  
manageable and cost-efficient  
data analysis**

Advisory

---

[www.kpmg.com/nl](http://www.kpmg.com/nl)





Through complete use of your own data, you generate value; for people, for society, for customers, for businesses and governments. The first step lies in overcoming obstacles common to many organizations, to easily unlock the value of your data, external data, and to develop new applications. KAVE forms the heart of a new 'Big Data ecosystem', removing practical obstacles, allowing quick, easy and consistent Big Data analysis.

The amount of data available to us is growing exponentially. With that growth comes ever greater possibilities to create value through

data analysis. It is not a coincidence that 'Big Data' is currently such a topic of interest (even if there is also a fair amount of scepticism about it), since the accurate analysis of large amounts of data from various sources can help us to progress and innovate radically. The areas where Big Data may lead to significant advances vary from simple file reductions to optimising environmental systems in offices and tailor-made consumer promotions, even on to more-focused cancer tumour treatments. We are at the start of a great journey: the most valuable applications still have to appear.

# Data silos

---

Development of data analysis is often a slow process. One of the key reasons for this is the current structure in which parties operate. Or rather, the lack of an effective structure. This applies to internal organisations, but becomes particularly clear whenever an organisation tries to use external data or otherwise needs to form ad hoc partnerships with one or more parties. Over and over again such organizations need to reinvent the wheel with regard to the technology to be used, which analysis software and algorithms to use, how to provide guarantees with regard to personal privacy, and how to agree on the conditions governing the use of the data. To put it mildly, this is hardly

an incentive for businesses and governments to get the most out of data analysis. On the contrary, it acts as a disincentive. At a time when the information society seems to be maturing, there has to be a better way. KPMG's Analytics and Visualization Environment (KAVE) offers organisations high-grade technological and organisational building blocks to perform data analysis easily and under the appropriate conditions. But KAVE is more than that. KAVE also anticipates a (near) future when data analysis plays an increasingly central position in society. A society in which data forms the raw material of almost everything we do.

# Ecosystem thinking

---

An ideal world would have an ecosystem in which organizations share (controlled) insights into data. An ecosystem which has pre-considered the options as to how these insights are shared. In such an ideal environment, suitable parties would also have sufficient opportunity to build applications based on this data. For instance, one party will see opportunities for an application that better predicts traffic jams based on location data combined with meteorological data, another party will focus on models to predict when certain parts of a car need to be replaced: the ecosystem will provide the required mechanisms to create these insights as and when all telecom providers,

car manufacturers and the Dutch Meteorological Institute participate. In this ideal world, all suitable parties will be free to bring an idea to fruition with the data in the ecosystem – of course under the conditions imposed by the data owners. Such an ecosystem would also stimulate innovation: and when both the idea and the results are good, society would embrace it. If not, applications just fade away. In this way, the best applications emerge by themselves, just like the best apps in the current app-stores. The wisdom of the crowds separates the wheat from the chaff through reviews and ratings.

## Sharing insights as a source of value creation

The real potential of data analysis lies particularly in sharing insights into the data. Think, for example, of the leap forward that could be made in improving medical diagnoses and treatments if it became possible to share insights into (anonymised) data for the good of mankind. This also applies to organisations. A financial institution could increase the security of credit or debit card use through insight into location data: as soon as a person's credit card is used at a different location from the location of that person's smart phone, a possible fraudulent act may be prevented. So, the bank will need to be able to combine knowledge on the location of the phone (for example using data from a telecom provider) with the bank's information on the use of the card. Another example of creating insights from combining data is the supervision of medical claims. A very different type of control of healthcare costs could be created by hospitals and insurers sharing insights into their (anonymised) data. The fraudulent use (or misuse) of healthcare arrangements could be monitored effectively in almost real-time.



## Ecosystem elements

The themes discussed above resulted in the idea behind the development of KAVE. With KAVE, KPMG has begun a new way of thinking in which data analysis applications can be realised far more easily. A reality in which standardisation of data or technology is not necessary, where agreements have been made regarding the conditions of sharing both insights into data and data analysis applications. A reality which includes the highest level of data security, combined with complete transparency on how to deal with personal data. Insights into the increasing data flows and data analysis of governments, companies and individuals can be integrated and shared reliably, rapidly and easily due to KAVE. These conditions are supported by KAVE's essential characteristics.

There cannot be any doubt that **compliance with privacy requirements** is our number one priority. This is ensured by enabling the involvement of Trusted Third Parties (TTPs). Personal data is converted into anonymised data and neither the TTP, the supplier nor the processor can remove this anonymisation. The result is that a developer or data scientist cannot trace data back to a person. The privacy principles, as stipulated by law, are enforced by the system's technology and set-up.

If the user clearly states that the personal data may be used for a certain purpose, or in the event of a legitimate interest such as a court order, the pseudonymised data can be re-associated, and the party which complies with all legal and ethical requirements will be able to link the insights to the appropriate person (this process is called re-identification). As this requires cooperation of all parties in the anonymisation chain, division of responsibilities guarantees compliance with rules and regulations.

The environment needs to be able to **process all common data formats and structures**. Until recently this was often an obstacle. But we now have state-of-the-art tools, methods and technology that can process almost all data. In a certain sense data has become like Lego: it always fits and you can build any model you like. Furthermore, data analysis is horizontally and therefore easily scalable: additional analysis capacity can be made available immediately by increasing the number of computers (and can be easily reduced in a cloud environment). In comparison: scaling-up of classic database systems often leads to technical and/or financial challenges.

It has to be possible to **make clear agreements regarding conditions** imposed by the data-processing party with regard to analysing its data and those imposed by the data-analysing party with regard to using its analysis. This also entails questions such as the fees for the use of the data (for example through subscription structures), the costs of using the environment, and the purposes of the analysis.

**The ecosystem needs to be open.** The power of an ecosystem, in which a large amount of data streams is processed, lies in the fact that there is open competition in drafting and developing applications. Application developers are free to stipulate conditions with regard to the use of their applications, and data-processing parties are free to stipulate conditions with regard to the applications they want to use for their data analysis. For example, analysis can be made available for free (e.g. for further development) to others. The comparison with Lego not only applies to the data but also to the applications. These are 'stackable', as is often the case with open source software. For example, an application for

predictive car maintenance could possibly also be reused in other sectors, because the algorithms are similar. And an application that leads to relevant results regarding monitoring and checking of invoicing energy consumption could possibly also be used for another energy company. Applying a review model – with users giving the application a rating and/or review themselves – will mean that the best applications emerge by themselves.

**Data analysis quality control.** Data analysis is more than number crunching. The quality of complex analysis in particular is not determined by the calculative power of the computer but by the data analyst. The difficulty is not finding

patterns. The real challenge is to understand or interpret, without bias, patterns in data, so that data analysis leads to meaningful – not disastrous – conclusions. The risk of misinterpretation must be addressed when analysing large amounts of data, because if you collect enough data, you will always find a striking correlation.

Therefore, if desired the quality of an analysis can be monitored with a peer review system. This is a crucial characteristic of the ecosystem to avoid incorrect conclusions being drawn. In addition to monitoring, if necessary, independent third parties can certify the analysis.

## Three options



### COMPANY KAVE

The ecosystem does not extend beyond the boundaries of the company's own organisation and only discloses internal data.



### SHARED KAVE

The ecosystem is open to a number of partners, with which agreements are made on the (conditions of the) usage.



### GENERIC KAVE

An open ecosystem which everyone can join. Conditions will also be made with regard to the use of analysis and data. However, the platform is open to everyone.

In the future even more layers will exist on top of this ecosystem to govern, control, and share insights.

At present, Big Data is not always seen as a good thing. People are worried about their privacy and assume that companies only want access to their data to make money. Governments and businesses face the challenge of showing the significant societal advantages Big Data offers.

This requires a trusted environment with room for open competition, together with full transparency on how the data and data analysis are dealt with. KAVE offers just such an environment.

# What can KAVE do for you tomorrow and into the future?

---

Organisations everywhere are exploring the possibilities of Big Data in new ways and from different angles. Some organisations have problems they want to solve using existing Big Data initiatives. Others want to put new ideas into practice. Every organisation faces different challenges, and KAVE helps to tackle those challenges. Some examples are retail businesses that can obtain unexpected insights

from the huge quantities of data they gather, in order to improve customer satisfaction; or hospitals that can, through collaboration with other parties, learn how to improve the outcomes of the medical care they provide.

**KAVE** makes three broad promises about how it can help businesses get the most out of Big Data.



**KAVE is a one-stop-shop for every Big Data issue**

---

## **'Is Big Data too complicated for my organisation because it involves a multitude of databases, systems and formats?'**

KAVE offers not only a qualitative toolkit to solve this, but also well-considered legal solutions and the best data scientists available.

## **'What does this mean for me?'**

It could enable you to unlock the value of Big Data for your organisation, independently of your existing structures and contracts and other existing preconditions.

## **KAVE in retail**

---

What (unexpected) insights can you obtain about how customers move through a store from smartphone data? We can obtain insights into what catches their attention;

we can observe and learn their shopping behaviour; we can learn how to better serve them. And we can also obtain other unexpected insights from this data.



## KAVE ensures that you remain in control

.....

### **'What if I am concerned that the reputation of my organisation will be in jeopardy if our Big Data initiatives go wrong?'**

KAVE makes it possible for you to remain in control of your data. If necessary, you can benefit from the advantages of Big Data without your data ever leaving your servers. We offer the most advanced solutions with regard to security and privacy and we understand how to set up the correct legal framework.

### **'What does this mean for me?'**

It means that you can explore Big Data initiatives in comfort, knowing that you will always retain control of your data.



## KAVE makes new combinations possible

.....

### **'Is there a potential to gain insights based on combining my internal data with data from outside of my organisation?'**

Performing combinations with disparate data sources is mostly still new territory for the corporate world, yet at the same time this forms the basis for the most promising aspects of Big Data. The attractive part of KAVE is that the insights gained in this way in particular can be shared without invading any person's privacy because our resources include reliable third parties combined with pseudonymisation techniques.

### **'What does this mean for me?'**

It means that you do not have to put limits on your innovative ideas because you now know that they can become reality through collaboration with other sources.

### KAVE in healthcare

.....

The more data healthcare specialists have on a patient – clinical data, preferably combined with data on lifestyle, nutrition, etc. – the more they are capable of providing tailor-made care with optimum impact, because they no longer need to take a scattergun approach to treatment. Instead, each situation can be tackled precisely and efficiently and every issue targeted. The question in this regard is: under

what conditions are we as patients prepared to share personal information about our health and lifestyle? KAVE offers solutions to ensure that these conditions, including privacy, are met; while sharing insights in order to improve the quality of healthcare by providing personal tailor-made care.

# Why KPMG?

---

## Professionals

More than 150 Data & Analytics professionals help organisations to create value through data every day. What is special about this group is that it contains a wide range of skills. We integrate expertise in the fields of audit, tax and advisory, and in doing so we offer a unique perspective on the challenges our customers face. The profiles of the professionals in this group vary from data scientists with doctorates who have unique experience at the European research institute CERN under their belts, to consultants who will manage your data analysis projects tightly and who will ensure that they are properly secured within your organisation.

## Strong partners

We have high-quality technology. We have a range of tools (developed in-house), we collaborate with strong partners such as Microsoft, we develop technological concepts ourselves (including location analytics solutions) and we have experience with relevant techniques for data analysis and visualisation. At the same time we understand that technology is only meaningful when it is intelligent and tailor-made.

## Contact

### Sander Klous

Director, Big Data & Analytics  
E: [klous.sander@kpmg.nl](mailto:klous.sander@kpmg.nl)



### Maarten Hoekstra

Senior Consultant, Big Data & Analytics  
E: [hoekstra.maarten@kpmg.nl](mailto:hoekstra.maarten@kpmg.nl)



### Rob Lambert

Manager, Big Data & Analytics  
E: [lambert.rob@kpmg.nl](mailto:lambert.rob@kpmg.nl)



## New business models

What do we do? We help organisations to experiment in order to explore new paths in Big Data and data analysis. We help them to discover new business models. But we also offer readymade solutions in order to properly secure their applications in this respect, with the certainty that these applications meet the highest demands with regard to quality, reliability and privacy. If required, we can use our own KAVE platform for this purpose.

In short; data is in our DNA. For decades, our business has been about generating high-quality insights from data and information – and therefore we can support our customers from the initial strategy-definition stage up to and including execution.

## Network

Finally, we believe in open [source] collaboration. We work closely with start-ups and other parties in order to unlock the maximum (technological) potential and learn from each other every day.