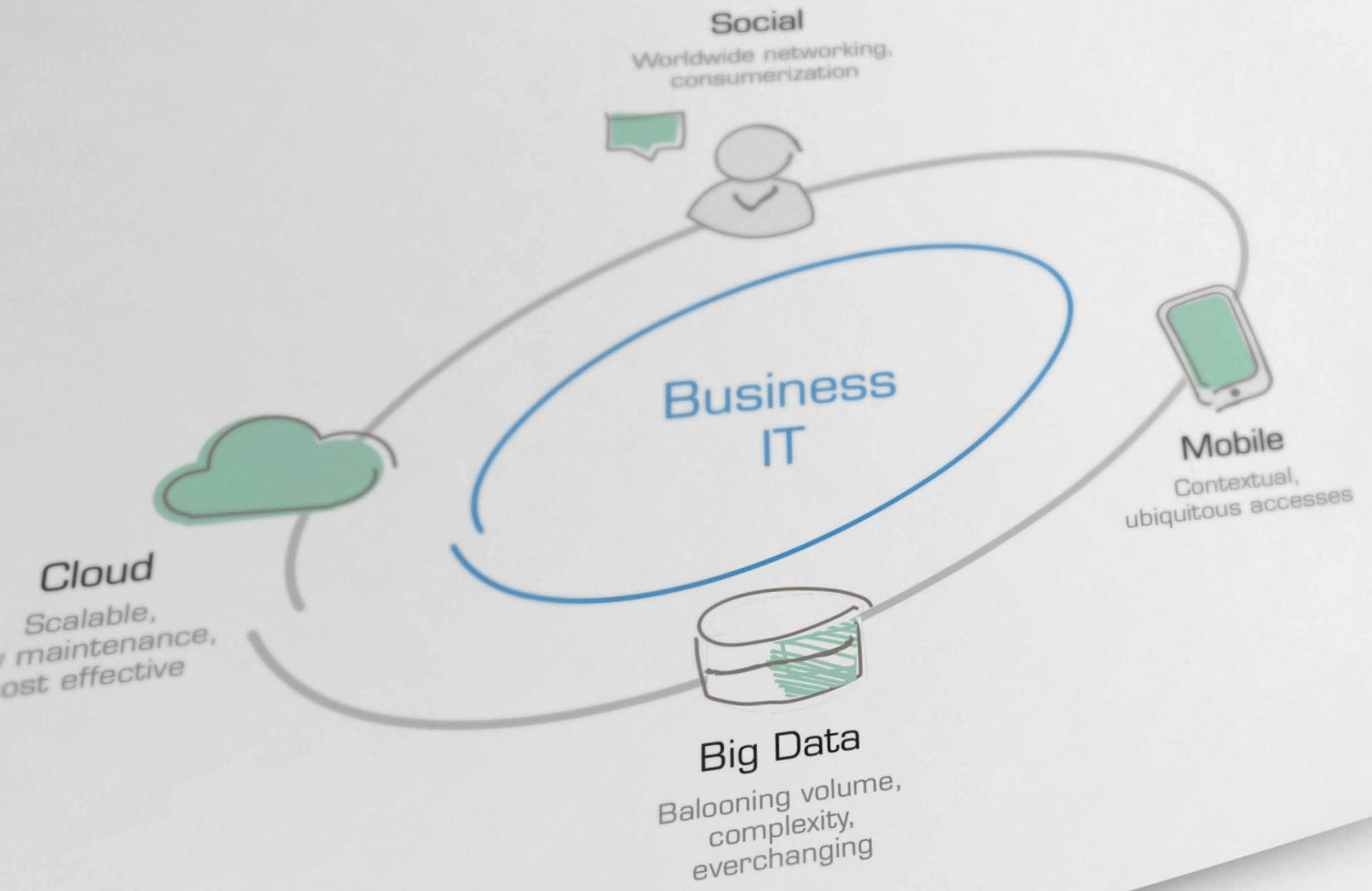
white **paper**



The Facts, Myths and Realities Behind Big Data

An Expert Overview

itelligence NTT DATA Business Solutions

Management Summary

When everything is connected to everything, data becomes really, really big … and keeping a clear vision becomes even more difficult than it already used to be. However, supported by the right technology, organizations can turn Big Data into a world of opportunities. But what exactly is Big Data?

And what about everything else that somehow seems to be related to Big Data: What is in-memory technology, SAP HANA, or Hadoop? But above all: What is or will be the impact of Big Data on business? This paper´s goal is to stay on top of developments. In addition to the personal perspective of our experts, this paper will also provide some clarification on current developments, give some facts, sort out some definitions, deconstruct some myths, and put away some red herrings as well.

About the Authors

Klaus-Christoph Müller

Berlin-based Klaus-Christoph Müller is Director Global Business Solutions at itelligence. Focusing on the development of industry solutions at itelligence, he has a long history in business development and strategic consultancy, as well as in solution development and presales. Starting in the automotive industry and working with Volkswagen and DaimlerChrysler Financial Services (now Daimler AG) in various expert and management positions, Klaus-Christoph has a long track record creating business value through innovative technologies.

Contact: klaus-christoph.mueller@itelligence.de

Christian Schmidt

Christian Schmidt is Global Director Business Analytics for itelligence in Munich. His work focuses on business analytics technologies, including SAP Business Objects™, SAP NetWeaver® Business Warehouse and SAP in-memory technologies (SAP HANA®). He gathered extensive industry experience working in various expert and management positions at Deloitte GmbH, Infineon Technologies AG and Mead Corporation (now MeadWestvaco). Christian currently works with customers implementing SAP HANA. He is also a well-respected speaker, presenting at industry events throughout Europe.

Contact: christian.schmidt@itelligence.de

Table of Contents

Today's Challenges

Challenges in Business

Two decades of tremendously increased internationalization have been matched by an accelerated rate of change for companies in almost all industries. This presents big corporations and also SMEs with multiple challenges – but also with a wealth of opportunities. They can now enter markets they never dared to dream of before. On the flip side, rising competition and market dynamics demand

- faster reaction times,
- accelerated innovation cycles,
- most flexible and adaptable business processes.

» Change is the process by which the future invades our lives.«

Alvin Toffler

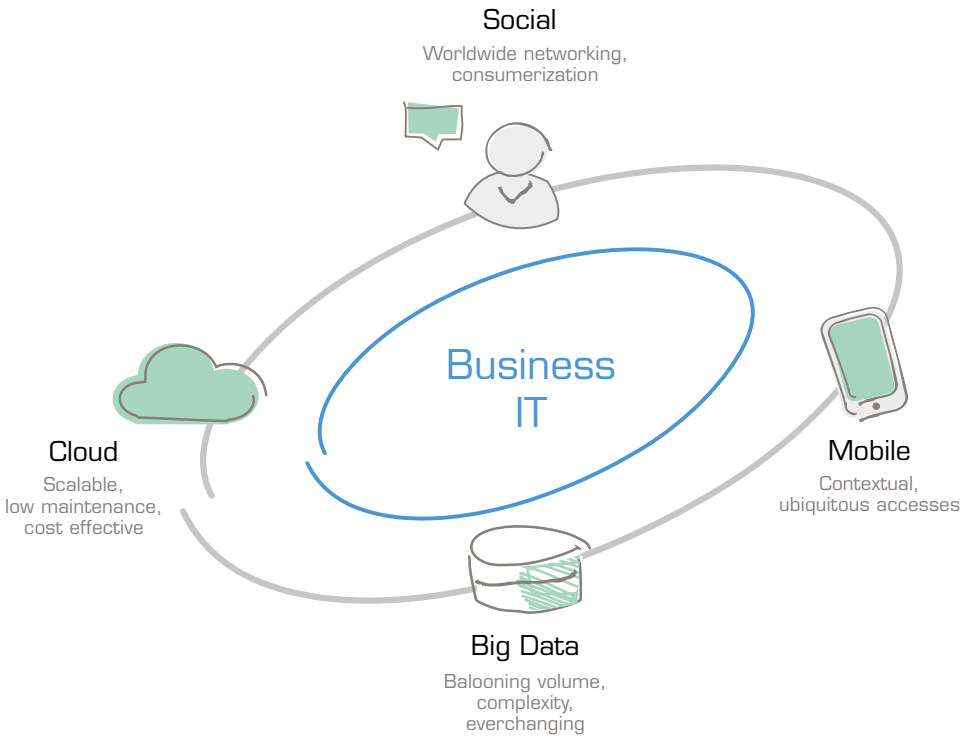
And as global competition and consumerization set the pace, relevant data is spreading and multiplying like virtual rabbits. The web, social media, and mobile devices expand all spatial and temporal limits. Why? Data in general and business-relevant data in particular are growing exponentially. In volume, the world's technological per-capita capacity to store information has almost doubled every 40 months since the 1980s. In 2012, 2.5 exabytes (EB) of data were created – every day.

Challenges in IT

The new global business reality is especially challenging for a company's IT systems and departments. Where there were relatively stable business environments, today there is constant change. New acquisitions and mergers, reorganizations, split-offs or new products require new processes and system reports. Thus, IT has to become much faster and way more flexible in order to keep pace with this ever-changing business environment.

One option to become more agile is getting away from classic waterfall projects. But it has to be clearly seen that consumerism has become an all-important topic, too. Big Data is only one of four big interdependent trends IT departments are facing at the same time in addition to mobile, social, and cloud (See Fig. 1).

Figure 1: How forces of innovation affect business IT



Considering these major shifts in technology, the only stable value in an ever-changing world seems to be budget constraints!

But there's also good news: storage is now cheaper than ever before! The storage costs of large data volumes have dropped precipitously over the last ten years. Also, data processing with RAM-modules and CPU is becoming faster and more cost effective (See Fig. 2).

This incredible price drop is accompanied by new in-memory technology such as SAP HANA and new analytical applications like Hadoop, which allow rapid, lower-cost conventional or predictive analysis of Big Data. But what's even more important: IT now has the opportunity to overcome one of its biggest pain points – the missing integration between transactional (or operational) data processing and analytics.

What is Big Data?

As it often happens with hot topics, everybody's talking about, many basic definitions and facts remain obscure. Instead a number of myths are given birth to.

So what is Big Data exactly? Basically "Big Data is a blanket term for any collection of data sets so large and complex that it becomes difficult to process using on-hand data management tools or traditional data processing applications". According to a more sophisticated definition by Gartner, Big Data can be characterized by the "3 Vs" – volume, velocity and variety.

Volume:

- Machine-generated data is produced in ever more massive quantities – in petabytes (PB) per day. A single jet engine generates 10 terabytes (TB) of tracking data every 30 minutes … and there are roughly 25,000 airline flights per day!
- Structured data needs high bandwidth and high-capacity storage.

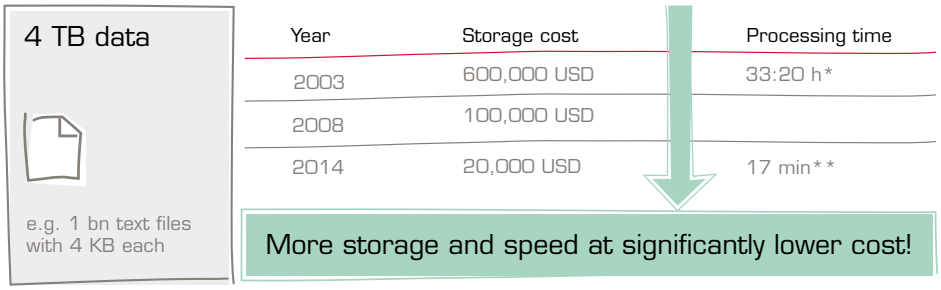
Velocity:

- Data streams on social media produce a constant influx of opinions and relationships that are valuable for customer relationship management.
- Even at 140 characters per tweet, the high velocity (and frequency) of Twitter data produces a volume of over 8 TB per day.

Variety:

- As new services are added, new sensors deployed, or new marketing campaigns executed, new data types are needed to capture the resulting information.
- The main challenge lies in identifying the value, the relevant information within the large volume of data, and then transforming and extracting that data for further analysis.

Figure 2: Significant cost savings in data storage and accelerated data processing



* Retrieving and uploading from HD storage
** Reading from RAM

» Big Data describes large volumes of high-velocity, complex, and variable data requiring advanced techniques and technologies to enable the analysis of information. «

Why has data volume increased dramatically over the last few years? Falling prices for infrastructure components like GPS transmitters, network connections and storage media are one reason. But what is more important is the fact that new ways of personal and business interaction through web, mobile, and social channels, and also mobility's exploding number of interaction points and interdependencies are creating more data entry and delivery points – a classic chicken and egg situation!

And if you are still asking yourself "how does this affect my business model?", well, maybe it doesn't – yet – but it will. Here are a couple of cases and numbers explaining why:

Let's start with an example from social media:
Facebook has more than 1.31 billion users sharing 1 million links and 3 million messages every 20 minutes. It collects an estimated 500 TB of data every day. "If you aren't taking advantage of Big Data, then you don't have Big Data, you have just a pile of data," said Jay Parikh, VP of infrastructure at Facebook. "Everything is interesting to us." (In comparison, the entire U.S. Library of Congress holds 20 TB of data!)

Another example from a more conventional industry:
The "Internet of Things" lifts traditional machine-to-machine communication to the next level. You've maybe heard of the "connected car": Many vehicles are now equipped with Internet access and a wireless local area network (W-LAN). Cars can warn each other or help service partners or OEMs to detect problems before they arise. ABI Research's latest data on the Internet of Everything (IoE) shows more than 10 billion wirelessly connected devices in the market today; with over 30 billion expected by 2020.

Myth 1: "Big Data is New"

Big Data is not new. Companies have long been dealing with Big Data and performing pretty sophisticated analytics. Just think about the purchase proposals on eBay or Amazon. Considering this, Big Data (analytics) has already proven their value in many ways.

Practical Examples of Big Data

1. Medical research

Big Data offers a new way to get insights. Before Big Data, researchers created hypotheses about relations or impacts, which had to be verified through statistical methods. This was time-consuming and costly, so the number of variables had to be restricted within selected parameters. If you uncovered falsification, you had to start over.

With Big Data, researchers can take the opposite approach – working from data to hypothesis. Based on numerous available data sets from medical records, they can look for new relations.

2. Political research

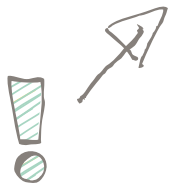
Before the U.S. presidential election of 2012, the University of Applied Sciences Ludwigshafen am Rhein harvested over 50 million feeds and posts from Twitter and Facebook in a showcase for SAP HANA in-memory database. Some 1.5 million posts per day – over 18 per second – were analyzed for meaningful opinions about the **candidates.** The winner predictions, based on this analysis of unstructured data, ranked way beyond expectations and surprisingly close to the actual results.

3. Retail

In 2011, retail giant Tesco opened a virtual store in the subway of the South Korean capital Seoul. There, commuters can shop on the go using their smartphones. QR codes on digital displays let customers order goods and have them delivered to their doorstep. Online sales increased by 130%. But Tesco also optimized internal processes with Big Data: 70 million refrigerator-related data points from units were fed into a dedicated data warehouse. They are analyzed to monitor performance, check if machines might need to be serviced and perform preventive maintenance to reduce energy costs.

4. Gaming industry

Tipp24 AG, a platform for placing bets on European lotteries, uses KXEN software (recently acquired by SAP) to analyze billions of transactions and hundreds of customer attributes. Tipp24 can develop predictive models to optimize targeting of customers and personalized marketing messages on the fly. This reduces the time to build predictive models by 90 percent.



Immense value of Big Data technology

Myth 2: “Big Data is a Single Technology”

All of the above examples demonstrate the immense value of Big Data technology – from knowledge management, business analytics and operational reporting to simulations and research.

What they do have in common are very intelligent, quite different solution approaches relying on structured and unstructured data from various sources. What they do not share is single Big Data technology – because there is no single Big Data technology. Big Data stands for the use of a combination of (old and new) technologies in order to gain insights while effectively managing data load and storage problems.

And what is also very important to understand: Big Data requires exceptional technologies to efficiently process those large quantities of disparate data, while providing insights and activity to end users within a minimum of time.

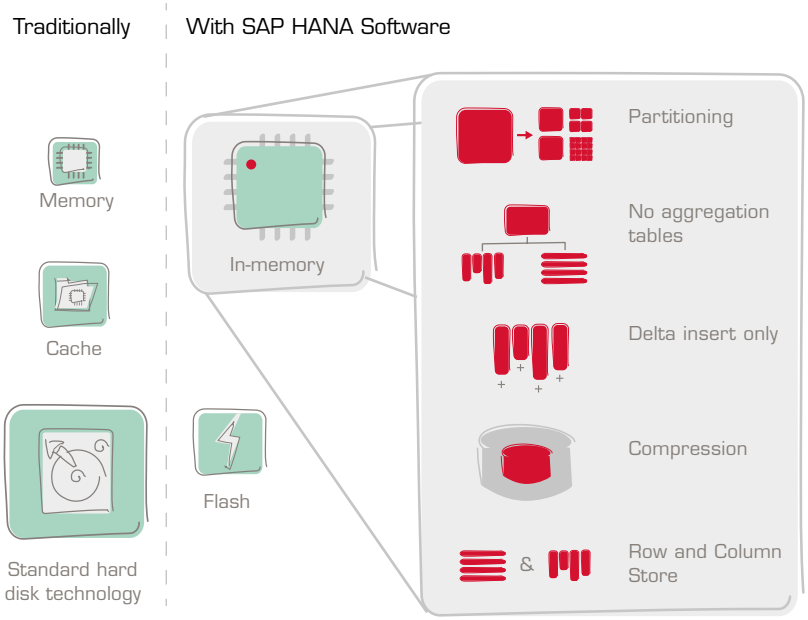
Now that we know the basic definitions of Big Data and the fog around two major myths has cleared, we can focus on two core Big Data technologies: in-memory technology and Hadoop.

What is In-memory Technology?

The traditional relational database model with row-based access was replaced by column-based access early in data-intensive applications such as Business Intelligence. But access to the data storage – the I/O rate – remained a bottleneck. Although modern processors can execute sophisticated algorithms rapidly, data exchange with the classical (relational) database did not speed up to the same extent.

Today, the massive performance increase of processors and the massive drop in prices for computing capacity has made direct storage of the data in the main memory ("in-memory") viable. In this way the time-consuming storage access can be completely avoided. And that's why in-memory technology has become a hot topic on many CIO's agenda these days.

Figure 3: Traditional vs. SAP HANA architecture



» In-memory computing will have a disruptive impact by radically reducing the total cost of ownership. «

A rather well-known example of a dedicated in-memory technology is SAP HANA, "an in-memory, column-oriented, relational database management system, developed and marketed by SAP". The capabilities include partitioning in database computing (complex operations take place within the database itself), the absence of aggregation tables (flexible data models without redundancies), real time replication and fast data loads, only delta insert, and data compression (up to 80%). But above all, in-memory row and column storage results in fast responses and a tremendous improvement in speed (See Fig. 3).

In-memory is particularly useful not only in business analysis or dedicated Big Data scenarios. It can also provide tremendous transactional processing speed. This, in turn, can pull instant insights from data lying idle due to performance restrictions of relational databases and the "I/O bottleneck". This increased speed makes real time analytics and added value in business operations possible. It can even lead to fundamental business model transformations.

At a glance: the benefits from in-memory technology

- Reduced operating expenses for applications (through data base/legacy applications offloading)
- Improved performance of transactional applications
- Growing horizontal scalability (scale up/down)
- Quicker response times for analytical applications
- Low latency application messaging (<1 microsecond)
- Dramatically shortened batch-processes execution time
- Self-service business intelligence and unconstrained data exploration in near real time
- Detection of correlations/patterns across millions of events in "the blink of an eye"
- Supporting Big Data (Big Data needs big memory)
- Hybrid Transactional Analytical Processing (HTAP) – running transactional and analytical applications in the same physical database

Myth 3: In-memory is "Just a Hype Spread by SAP"

Don't be misguided by self-proclaimed experts saying that in-memory is "just a hype spread by SAP" and "a new and unproven technology". More than 50 software vendors deliver in-memory technologies today, and in-memory computing has been available since the late 1990s for certain types of applications, especially analytics. Textual analytics and semantics, for example, have been around for years. At SAP, this started with SAP APO and the TREX search engine.

» By 2025 most IT organizations will follow SAP's path and run their entire infrastructure on in-memory databases. «

The only thing that is relatively new is the fact that in-memory technology is no longer confined to dedicated or specialized applications. Since 2013, SAP HANA has been able to replace existing relational databases completely, serving as a unified platform for the complete SAP Business Suite – as well as for SAP BW/BI/BO – and as an independent platform for new applications. By providing the SAP Business Suite on SAP HANA databases, SAP brings the full advantages of in-memory technology to transactional operations as well.

Gartner predicts that by 2025 most IT organizations will follow SAP's path and run their entire infrastructure on in-memory databases with persistence to flash. Tape and HDD will be dead.

What is Hadoop?

Hadoop is not a database. Hadoop – from Apache Software Foundation – is a Java-based software framework for scalable, decentralized software applications that supports easy handling and analyzing of vast data volumes.

As we said before, the three Vs – volume, velocity, and variety – represent Big Data's main challenges for disk-based relational databases. How does Hadoop handle the three Vs?

- **Volume:** Hadoop handles data volume by splitting data and data-processing between multiple "DataNodes." As data volume, or the processing workload on each individual DataNode increases, the data in a node can be split, adding more nodes.
- **Velocity:** Hadoop handles data velocity by avoiding – or at least postponing – the slowdowns that come from inserting data into a conventional relational database-management system.
- **Variety:** Hadoop handles data variety by storing data as Linux operating system files without checking or processing them first. Absolutely any type of data can be stored. There is no need to understand and define data structure beforehand.

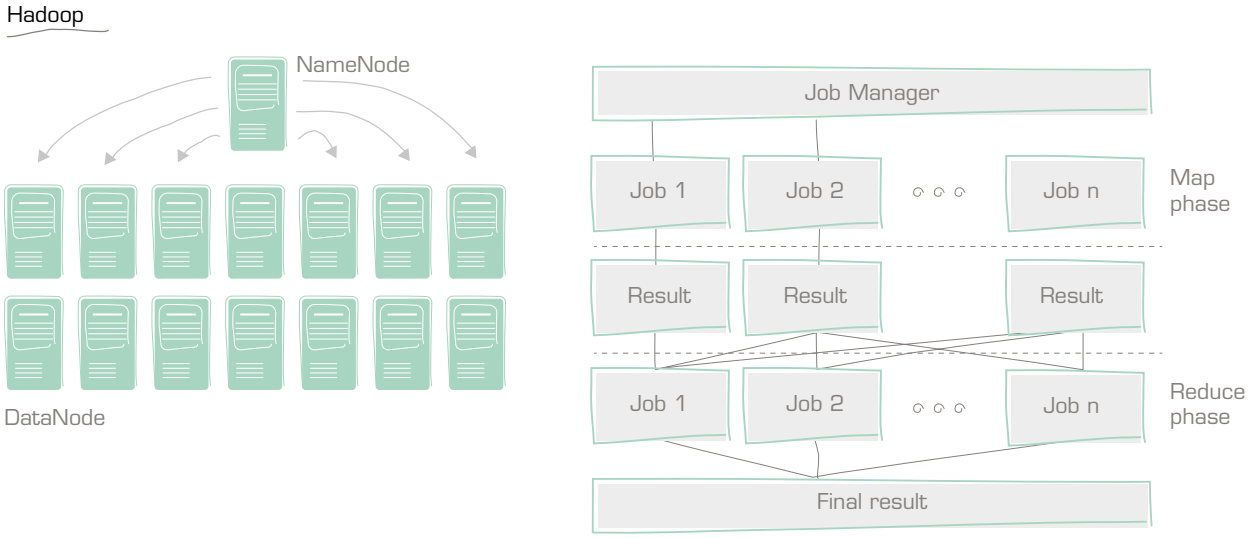
» Hadoop lets you store and access more voluminous, detailed data at lower cost. You can drill deeper and enter your business data in different ways. «

Hadoop runs on hundreds or thousands of commodity (low-cost) servers – these are the DataNodes. Each holds just a part of the data. By breaking a processing job down into thousands of smaller jobs running in parallel on individual machines, Hadoop can handle petabytes or more by adding more DataNode servers. Hadoop follows Google's Map Reduce Programming model to execute analytic and processing jobs in parallel (See Fig. 3). A key feature of the Hadoop software architecture is the way it separates how data is stored from the way it is processed.

Many overall principles of Hadoop look familiar. Remember stored procedures and batch? The effort in getting familiar with Hadoop will pay off quickly because the user interaction is simplified significantly: You get answers while still typing. Impressive – especially if you know the huge number of data sets involved.

Though Hadoop and similar approaches are in their early stages, they are well worth thinking about, as relational databases can't cope with Big Data. Furthermore, non-relational databases can deliver peak performance only when combined with in-memory technology. This leads us to the next questions: How does Hadoop relate to in-memory technology?

Figure 4: Cluster of commodity servers and the Hadoop Map Reduce Programming Model



Myth 4: Hadoop is an Alternative to In-memory Technology

Given the facts above, it's obvious that Hadoop is not an alternative or a competitive technology to in-memory computing. On the contrary: In-memory is the technical basis and an ideal platform for Hadoop.

Hadoop and in-memory technology enable large volumes of data from various sources to be easily integrated in real time – and Big Data analysis can be carried out without interference from existing business-analytics and data-warehouse implementations.

That is one of the reasons why SAP is joining forces with the Hadoop framework, resulting in a dedicated toolset for easy integration with existing SAP data warehouses (SAP NetWeaver BW). Several cooperation agreements also exist with Hadoop providers like Cloudera, Hitachi Data Systems, Hortonworks, HP and IBM – all based on the SAP HANA in-memory-database.

Big Data's Impact on the Business IT of the Future

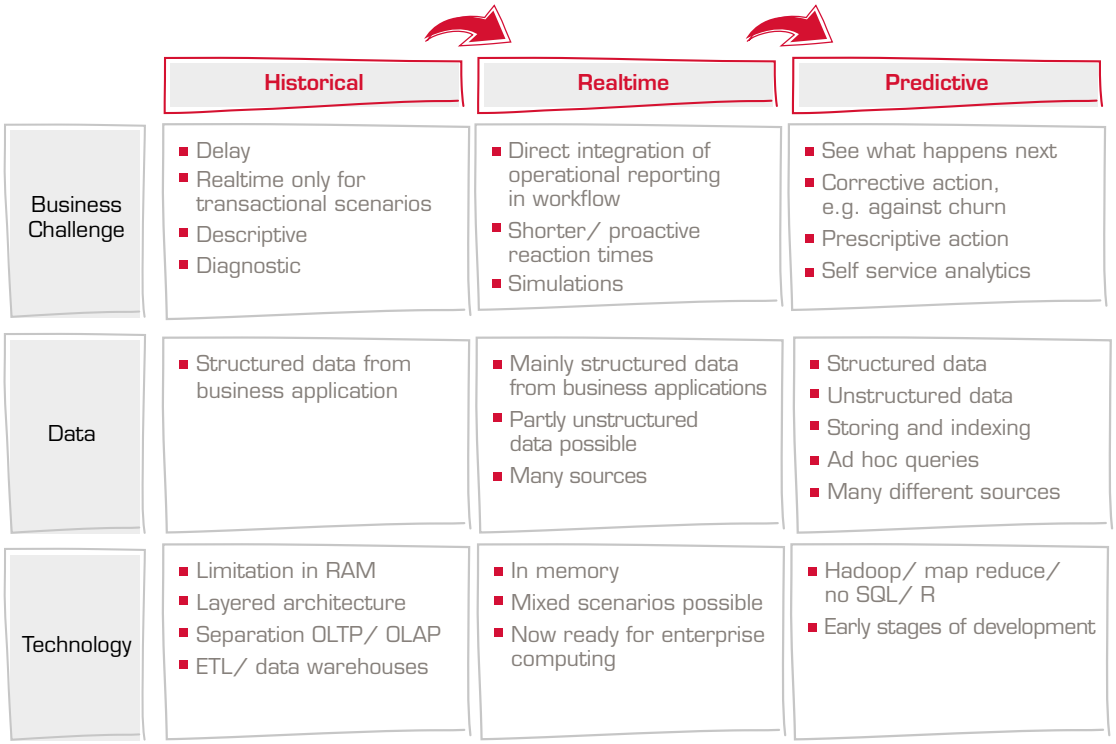
There is a tendency to make business decisions based on historical data, even though "real time business" was postulated years ago. But if you take a look under the bonnet, this postulate applied only to the transactional part of business systems. For performance reasons, the analytical part was uncoupled from the rest, both technically, because there were different systems, and logically, because analysis was always performed upon historical, aggregated data.

Improved performance through Big Data technologies, such as in-memory technology and Hadoop, allow for "real real time" analytics – so that companies can now focus on what's going to happen next and act strategically, rather than reacting based on purely historical data.

Formerly, many companies couldn't profit from valuable data even when available. Why? Since the beginning of the ERP era, companies have basically had two options for reporting and business intelligence. First, they could develop in-house software, which usually led to complex and rigid solutions. Business users couldn't change reporting structures and queries themselves. Predefined reports and delivered results tended to be outdated quickly and didn't allow for far-reaching-conclusions or rapid, well-informed decisions.

So over time, the second alternative – a data warehouse – became increasingly popular, because of its greater flexibility. But that flexibility came at a price: An additional system had to be established, and forming the reports still required expert extractors. These warehouses remain expert tools limited to trained people within certain lines of business.

Figure 5: The development of a company's analytical capabilities

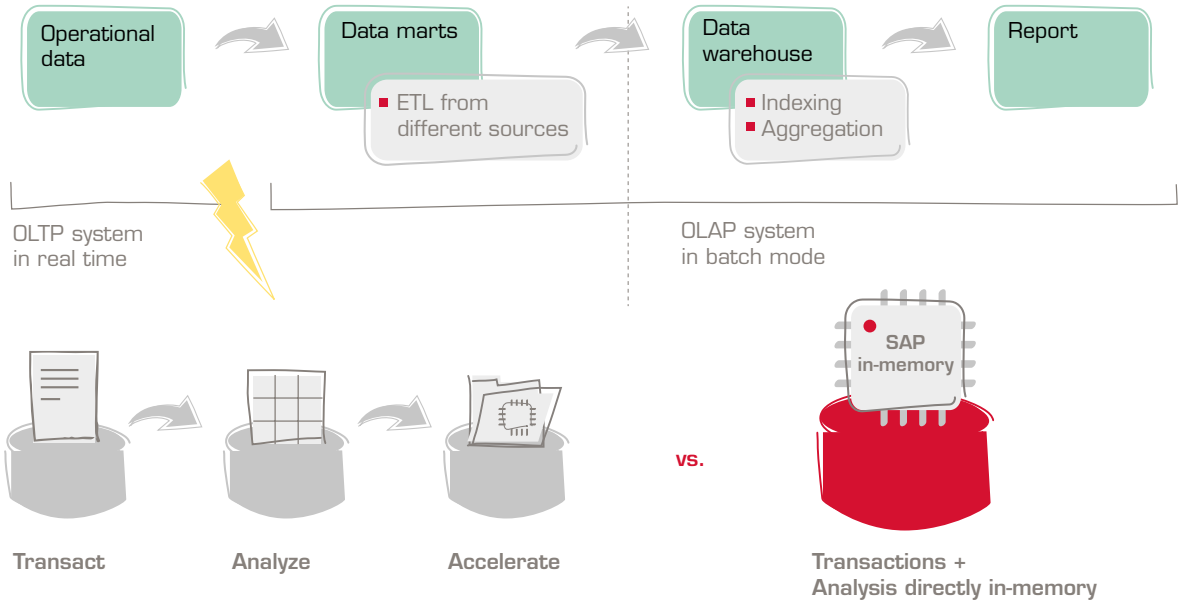


Even when appropriate reports are available, performance and response time are still an issue, along with concerns like mobile access for field sales or service. The lack in overall power resulted mainly from the split between OLTP and OLAP, brought on by the high price of storage and memory.

With in-memory technology, separation of transactions and analysis is no longer needed, nor are additional acceleration processes. Operational reporting can move back where it belongs: in the core operative system! This convergence of OLTP and OLAP workloads is called "Hybrid Transactional Analytical Processing" (HTAP). The term describes the technical ability to run applications and analytics in a single environment, while achieving notable improvements in performance and processing speed of the SAP HANA data platform – which should be seen as a "real" real time database platform.

In a "real time" context the source system's data quality, governance and monitoring are more important than ever. Otherwise – as additional layers are eliminated with HTAP – SAP HANA will still deliver quickly, but only bad or incorrect information. Therefore data quality assurance has to change from an "after the fact, latency-driven" process into a "parallel, real time effort at the point of entry" process. Pretty much like the total quality management paradigm in lean manufacturing: Incorporate quality right from the beginning and avoid (costly) inspections and corrections later on.

Figure 6: In-memory technology like SAP HANA unifies and accelerates transactions and analyses



Deliver only what is required!

But the platform alone does not enable real time: Mobile access and state-of-the-art user interfaces are also a must-have to reduce complexity for the user.

The overall principle: Deliver only what is required – to the right place, to the right person, at the right time! *(A good additional read on this is the itelligence white paper "Making 'Changes in Order': The New Role of the CIO", especially the chapter "Why Usability Matters Most")*

Predictive Analysis – Because the Time is Right!

Predictive analysis, like any kind of prediction, is inherently uncertain. Big Data won't change that completely, but it will reduce knowledge gaps and therefore shrink the "terra incognita" of business significantly. Making hypotheses is the starting point, and the predictive analytics tools then help verify concerns, estimations and options, as well as providing evidence.

» Big Data will reduce knowledge gaps and shrink the "terra incognita" of business significantly. «

This approach relies on statistical methods such as linear or non-linear distribution models, regression tests, time series analysis, classification and clustering, etc. Most of these use cases are already available as open-source code. So-called R libraries for statistical calculation and graphics are used in most software for predictive analysis, for example, SAP Predictive Analysis. Open-source R delivers a very good basis for predictive analysis of large datasets. More than 3,500 add-ons are offered, which enable the creation of individual functionalities.

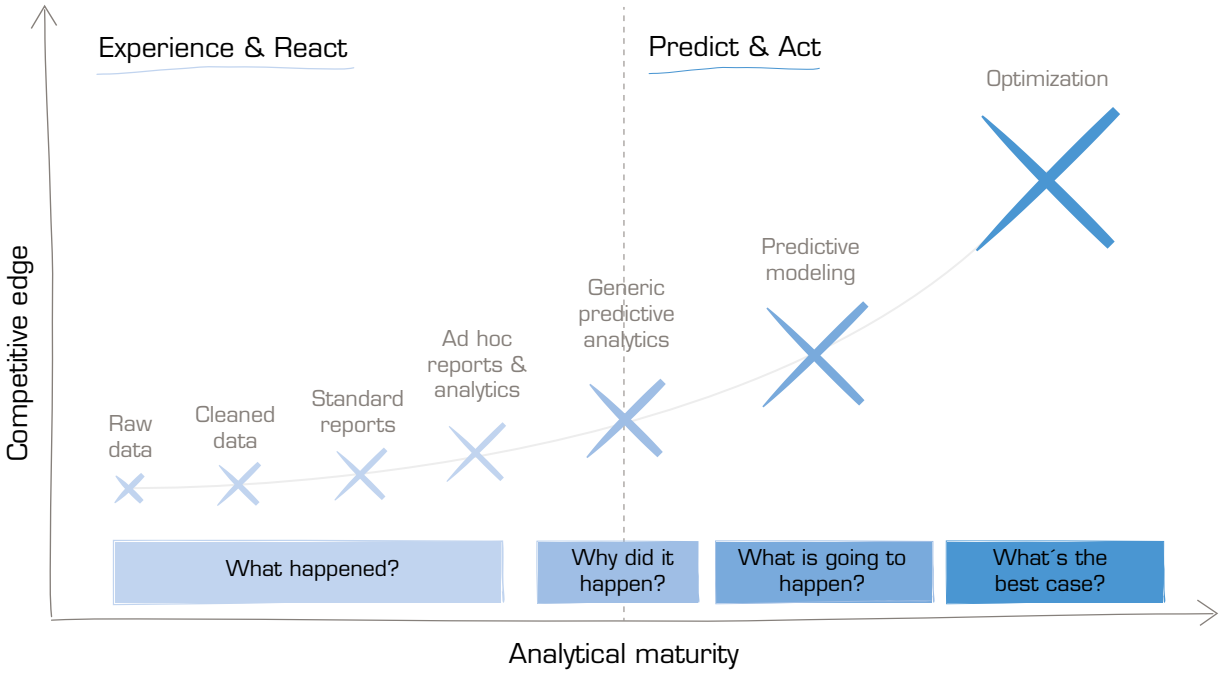
Such functionalities have been around for a long time, and statistical methods and tools have been on the curriculum of business schools and universities for years. Yet predictive didn't really take off, mainly because of missing collections of relevant data, difficult access to data and/or performance restrictions. Generally, only companies with very urgent business requirements were developing proprietary predictive applications.

Big Data technology now provides easy access to solutions that smash the biggest barrier to data aggregation – performance shortcomings and latency. But there is still a bit more to consider before we reach successful prediction:

- maximum usability for business users,
- easy and safe interactive, real time access to relevant data from different sources,
- specialized skill sets and expertise within your staff so they can ask the right questions and form the right hypotheses to move from agnostic analysis to predictive.

In this context, SAP InfiniteInsight® is a very interesting solution: It's not an expert stat workbench but a powerful solution for predictive analytics that can easily be used by business analysts – not just by data scientists. The software provides modelling automation and full access to the mountains of data that companies stock in their data-warehouse and CRM systems. You don't need manual preparations like pre-selection of data. The dual benefits for business are efficient data access for the average user, and protection from excluding valuable predictive information through misleading data pre-selection. The right proposals for customers can be made within seconds – right when needed.

Figure 7: The business value of predictive analytics



Myth 5: In-memory Computing is the End of Data Warehousing

Now that the separate operational data store for business analytics is obsolete, what will remain as a task for data warehousing in the future? Using in-memory technology like SAP HANA greatly reduces extraction, transformation, and loading (ETL) of operational data to and from a data warehouse. Data marts are no longer required. There are no aggregates, no indexes: The data is either directly accessible (in every detail!) or simply replicated on the spot ("sidecar scenario"). So is in-memory computing really making data warehousing obsolete?

Actually, it is not. Data warehouses will still be used – but their use and implementation will profoundly change. Data warehouses will be used mainly for agnostic tactical and strategic BI, such as comparing different scenarios with different likelihoods. But in terms of implementation, you will no longer need to store everything in a single database and then "denormalize" it. And, despite the valuable capabilities of in-memory technology, it no longer makes sense to permanently store all kinds of data in a single database – petabytes of unstructured data, social media data, web logs and other streaming data that might be relevant for decision making. Instead, companies can consider business orchestration models that allow users to answer these questions. SAP HANA combined with Hadoop, for example, can determine the appropriate source for relevant data.

Some Recommendations For Your Big Data Strategy

As we've shown, moving operational reporting back where it belongs and ending the inappropriate use of business intelligence for operational reporting can be taken for granted. The same applies to unleashing real analytics for agnostic and strategic scenarios. But what strategic conclusions can be drawn from this? Here are some recommendations …

Predictive Analysis Will Make a Big Difference

Predictive capabilities can meet an obvious need of many companies. For example, precise demand prediction will help avoid costly write-offs of overstock, which was previously unavoidable because of data non-availability and performance issues. In-memory technology can end this enduring displeasure. We can expect dramatically rising demand for predictive analytics, even from small and midsize companies.

The overall success of in-memory will also depend on a company's ability to further develop its employee skill set. Collecting data and providing the relevant information is just the first prerequisite for accurate predictions. The second is having the expertise and experience (!) to come to the right conclusions. It will take "deep analytical talent" – skilled analysts and managers who understand Big Data analysis well enough to make the right decisions.

Depending on individual pain points and business cases, most companies may do best to lay the foundations for bringing in real time capabilities before leaping directly into the future. But ultimately it's worth the investment on a real time platform, as this overview of benefits shows:

- **Improved efficiency:** Minimizing the number of systems and databases – and consolidating all processes on a single platform – greatly reduces complexity and costs (the full SAP Business Suite, for example, can run completely on one SAP HANA instance).
- **Fast, actionable reporting:** Formerly slow-running processes like material requirement planning (MRP), profitability analysis and other complex reports become "fast queries".
- **Advanced real time analytics:** Forecasting, simulation, what-if-scenarios, zero-latency drill downs become simple.
- **Continuous process innovation:** Goal-driven decision making improves both profitability and customer satisfaction.

Given this, we predict a major movement towards in-memory technology within the next two to three years.

SAP HANA is a Proven Platform

We recommend that from now on all new implementation projects should be based on SAP HANA. All future system enhancements and new products from the world market leader in business solutions will be based on SAP HANA. This modern, open standard technology supports multiple devices through HTML5 user interfaces, and all proven SAP functionality is ready to run on SAP HANA. Hosting the SAP HANA platform is already routine for many data centers today.

But, of course, there is not only one single strategy. Maybe the most important factor for decision makers is to ensure that the initial investment stays protected no matter what migration path and business scenario you choose. Safe passage to the new technology is assured, whether capturing existing system configurations or forming tailor-made migration tools.

If your competition has already started Big Data initiatives, you could quickly face severe disadvantages without comparable real time insights. And if not – why give your competitors the chance to start first?

General predictions are difficult to apply to different industries. Those under high margin pressure, like trade or retail, should gain the biggest competitive advantage, while their investment risk seems low. Early adopters of Big Data technologies include telecommunication providers (e.g. for churn prediction) and financial services (e.g. for fraud detection).

While studying use cases for real time and predictive scenarios, we learned that it often makes sense to cluster by functional areas or processes, rather than by industry. Many of our customers and contacts showed great interest in these functional areas:

- Financial performance, risk analysis, security, fraud detection, compliance
- Sales and marketing – customer tracking, sales channel optimization, customer value analysis, social media, customer sentiment analysis
- Supply and demand operational reporting, root cause analytics, strategic scenarios

However, these examples can only serve as starting points for individual assessments. What really matters are the specific situation and requirements of your company. They will determine whether you should a) focus on bringing the operational reporting back into the core system, or b) focus on gaining more insight by combining structured and unstructured information in a Big Data business intelligence project. One company might put more emphasis on getting the right information to the right person at the right time to speed up operational processes. Another might focus on identifying new business lines based on analytical insight.

The Business Case Determines Your Roadmap

Your individual business case determines the ideal starting points and the path to sustainable business value through Big Data. Rather than repeating well-known facts about how to calculate the business case, we prefer to provide helpful advice on how to get orientated.

1. Check for information already available

Where is the information in your company, and what might be its potential business benefit? Is there any really low-hanging fruit? You all know how it is when someone says, "Oh, if I had only known that before …" Sometimes simply providing access to valuable information is enough. You don't necessarily have to "think big" about Big Data. Studies show that most businesses making use of Big Data already had the information but lacked the right questions and the analytical support to provide answers to those questions. For example, global logistics firm DHL keeps track of its parcels at every stage, but they were unable to make further use of the data until they reworked their analytics systems.

2. **Find out how your preferred operation model impacts your business case**

   Are you going to have the full implementation on your premises, or is a "side-car model" with added cloud services the better option? As you can imagine, the answer to this basic question will greatly predetermine your efforts and costs.

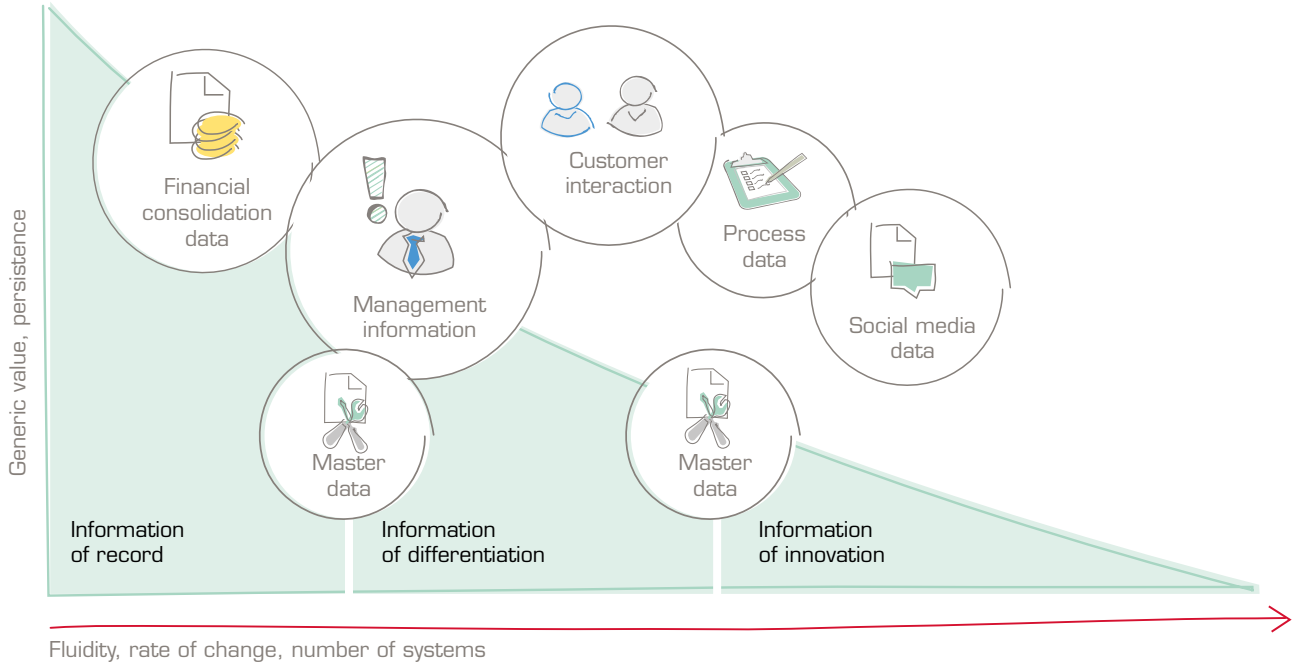
3. **Keep an eye on total cost of ownership**

   Total cost of ownership is very different from total cost of acquisition. Both are equally important, but the picture can be misleading. Look at the comparison between disk and RAM prices. At first glance, RAM is far more expensive than disk, but considering performance-per-second, RAM is far more economical. Also important: lower running costs due to fewer systems, lower requirements for power, cooling, floor space, resources, etc.

4. **And last: Remember that it is always easier to pile up costs than to evaluate benefits**

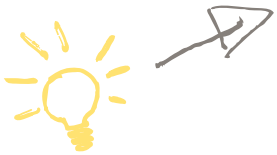
   Although there are some very impressive examples of Big Data analysis out there, the business case is often still unproven. But this shouldn't stop you from thinking forward and helping your company stay ahead of the competition.

Figure 8: Not all information requires a Big Data approach



In general, information of record is much less volatile and therefore could be stored (or better consolidated) in a single system. It can be very helpful here to recall the three V´s determining Big Data: volume, velocity and variety. But even if these criteria are being met, you still need proof of value.

As a rule of thumb, try to classify your data like this:



Classify your data

Structured / already collected	»	Make it happen now!
Structured / not yet collected	»	Make it happen soon!
Unstructured / not available yet	»	Wait and see! (If you see the chance for it, then just do it!)

At least the first two scenarios are well covered by in-memory technology. SAP HANA, as technology leader in in-memory computing, is ready to use. You can improve your business processes by either replacing relational databases completely or putting them alongside conventional databases to speed up particular functionality. No matter what you decide, a full range of additional tools and enablers is available for secure and fast access to in-memory technology. So be prepared and at the top of your game: 2014 is about to become the breakthrough year for in-memory technology.

Being able to accurately predict the future will no longer be wishful thinking, but a key skill for many IT professionals – a skill that will deliver great business value for their companies.

Read more …

… about future-oriented strategies and software solutions in our series of white papers. If you want to know more about Big Data please contact the authors or visit us online at: **www.itelligencegroup.com**

References

Hilbert, Martin; López, Priscila (2011). "The World's Technological Capacity to Store, Communicate, and Compute Information". Science 332 (6025): 60–65

http://en.wikipedia.org/wiki/Exabyte

"IBM What is big data? — Bringing big data to the enterprise". http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html. Retrieved 2014-05-06

http://en.wikipedia.org/wiki/Big_data

TechAmerica Foundation, Demystifying big data: A Practical Guide to Transforming the Business of Government. www.techamerica.org/Docs/fileManager.cfm?f=techamerica-bigdatareport-final.pdf

http://www.statisticbrain.com/facebook-statistics Retrieved 2014-05-12

http://gigaom.com/2012/08/22/facebook-is-collecting-your-data-500-terabytes-a-day

http://en.wikipedia.org/wiki/Internet_of_Things

https://www.abiresearch.com/press/more-than-30-billion-devices-will-wirelessly-conne Retrieved 2014-05-12

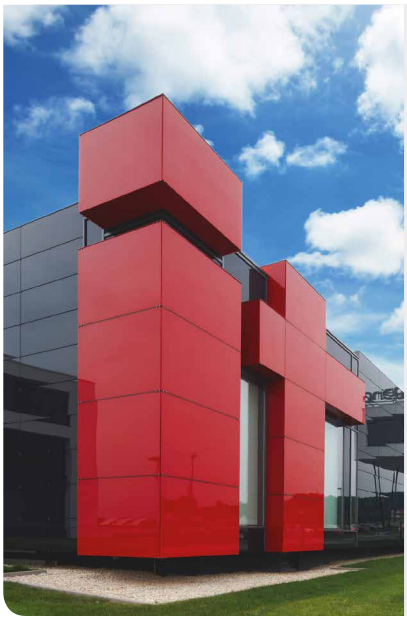
http://en.wikipedia.org/wiki/SAP_HANA

https://www.gartner.com/doc/2657815/hybrid-transactionanalytical-processing-foster-opportunities

http://www.sap.com/bin/sapcom/en_us/downloadasset.2013-08-aug-23-23.Architecture%20Matters-Thought%20Leadership%20Whitepaper%20on%20Enterprise%20Data%20Warehousing%20by%20Claudia%20Imhoff-pdf.html

www.mckinsey.com/features/big_data

http://www.forbes.com/sites/gilpress/2013/12/12/16-1-billion-big-data-market-2014-predictions-from-idc-and-iia/



About itelligence

itelligence is one of the leading international full-service providers for SAP® solutions, employing more than 3,800 highly qualified employees in 22 countries. Our SAP expertise is unrivalled and has been developed over the last 25 years, supplying specialized integrated business solutions for over 5,000 customers worldwide. Our service portfolio includes SAP Strategy, SAP Licensing, Implementation and Process Consulting, Outsourcing and Application Management Services. We are pioneers in the development of industry-specific solutions and our experts work hand in hand with SAP on new scenarios and ground-breaking solutions for business analytics, in-memory computing and SAP HANA®, cloud, and mobility. In 2013, itelligence generated total sales of EUR 457.1 million.

www.itelligencegroup.com

Austria
Belgium
Canada
China
Czech Republic
Denmark
France
Germany
Hungary
India
Malaysia
Netherlands
Norway
Poland
Russia
Slovak Republic
Spain
Switzerland
Turkey
Ukraine
United Kingdom
United States of America

itelligence NTT DATA Business Solutions

Imprint/Contact
itelligence AG ▪ Königsbreede 1 ▪ D-33605 Bielefeld ▪ www.itelligencegroup.com