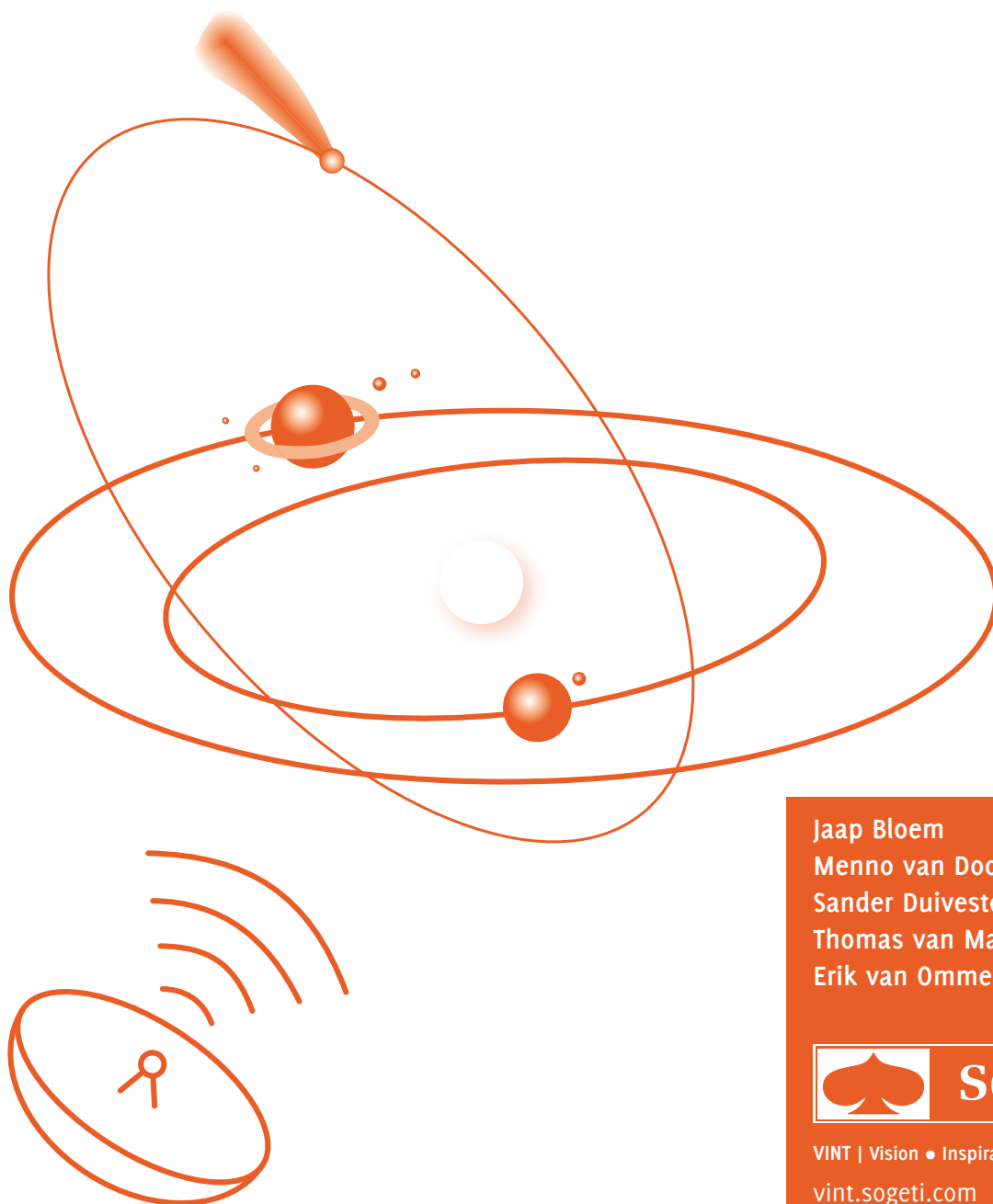VINT research report ① of 4
VINT research report ② of 4
VINT research report ③ of 4
VINT research report ④ of 4

Big Social

Predicting behavior with Big Data

Jaap Bloem
Menno van Doorn
Sander Duivestein
Thomas van Manen
Erik van Ommeren

SOGETI

VINT | Vision ● Inspiration ● Navigation ● Trends

vint.sogeti.com
vint@sogeti.nl

Table of contents

SOGETI

VINT | Vision • Inspiration • Navigation • Trends

The VINT Big Data research reports

Since 2005, when the 'Big Data' concept was launched—remarkably enough by O'Reilly Media, which had introduced Web 2.0 only a year earlier—Big Data has become an increasingly topical subject. In terms of technology development and business adoption, the Big Data field has undergone extremely rapid changes, and that is an understatement.

In *Creating clarity with Big Data*, the first of a total of four research reports, we offered an answer to questions about what Big Data actually is, where it differs from existing data classification, and how the transformative potential of Big Data can be estimated. The concrete adoption and plans of organizations are currently and primarily oriented toward the theme of *Big Social* we address here: basically the customer side, particularly inspired by the social network activity of Web 2.0. The data explosion is taking place all around us, but a major part of the discussion concerns the extent to which organizations should now plunge into Big Data. The answer is: only on the basis of a well-grounded policy. This certainly applies to privacy issues, which will be comprehensively covered in our third research report.

With four Big Data reports, VINT aims to create clarity by presenting experiences and visions in perspective: independently, and furnished with appropriate examples. But not all answers, by far, can be given. In fact: more questions will arise—about the roadmap that you wish to use for Big Data, for example. We have earmarked the fourth report for this task, dealing specifically with management and governance. And about the way in which you may be compelled to restructure your organization. And about the privacy issues that Big Data evokes, with regard to social analytics, for instance. And about the things new algorithms and systems will probably bring us.

The new data focus is a quest with many questions at the outset, while new ones will certainly arise during the journey. For this reason, we are only too pleased to exchange ideas and opinions with you: online at vint.sogeti.com/bigdata and, of course, in personal discussions. By actively participating in the discussion, you will help yourself and us to further refine all concepts relevant to Big Data, with the ultimate aim of gaining progressive insight into taking clear and responsible decisions.

In the context of inspiration, this report also presents seven issues about which we would be glad to hear your views. The downloadable PDF document on the website allows you to click on the relevant buttons. Subsequently, you are directly connected to the discussions in question.

Join the conversation

1 What's next in Big Data?
Nine observations with both feet on the ground

We deliberately begin this second research report on Big Data with both feet firmly on the ground. We present nine observations as the first sequel to our opening report *Creating clarity with Big Data*. As mentioned above, we have reserved our third report for the delicate matter of privacy — including a kind of Big-Brother anxiety — and the way to cope with this. This second report, *Big Social*, offers a multi-faceted orientation to the promising Big Data development with regard to Social Analytics and Social Media. Some of these are indeed very promising, others only promise a lot. The more confident you are in your own judgment, the more able you will be to give a personal assessment and act in accordance with your observations.

In this report we will focus primarily on the predictability of consumer behavior. On our Big Data weblog http://vint.sogeti.com/bigdata the range of Big Data topics covered will be much broader. The crucial question for now is simply *What's next in Big Data?* Many organizations are finding that they have been waiting too long for concrete solutions. These should be easy to implement, of course, and should amply repay the effort expended. Nervous eagerness and skepticism now mark the start of business practices that more and more will confidently build upon Big Data. Large Big Data projects and strategies are yet to come but the current emerging next practices that already can be discerned surely will find their place in daily operations everywhere.

In that context, please carefully consider the following observations, and see them primarily as central to the dynamic Big Data discussion that is currently in full swing:

1. Qualified best practices are under construction
2. Technology is making a breakthrough
3. Big Social will fulfill the promise of hypertargeting
4. Big Data runs the risk of becoming an out-of-control party
5. Clearly determine your actual needs
6. Behavior is often predictable without Big Data
7. Big Data builds upon traditional data centricity
8. Big Data ROI is simmering through
9. Social Analytics is the current Big Data pet topic

1 Qualified best practices are under construction

Anyone who has done any study on Big Data will surely know the promising report entitled *Big Data: The Next Frontier for Innovation, Competition, and Productivity* by the McKinsey Global Institute. It makes you think: so, something big is about to happen!

All the more remarkable, therefore, was the following admission by Michael Chui, one of the authors, at the MIT Sloan CIO Symposium in May 2012. Exactly one year after the publication of the *Next Frontier* report, Chui stated:

'There are no [Big Data] best practices.
I'd say there are emerging next practices.'

This seems to be at odds with the title of the report, but the coherence lies in the use of the word 'next'. Big Data will be capable of spawning benefits for innovation, competition and productivity, but convincing proof of this is not yet manifest. There is indeed a great deal about to happen, but much hard work must still be carried out in order to develop and implement 'emerging next practices'. With all the investments that have already been made, many organizations are not really waiting for such developments. Certainly not in the current economic malaise.

Organizations have begun to experiment, but it is as yet too early to give classic examples, best practices, that others can emulate. The Big Data domain is still undergoing too much change and it may even be likely that this will remain the case, in view of the claim of the favorable effect of Big Data on innovation and competitiveness. Both of course imply ongoing change and renewal.

2 Technology is making a breakthrough

In technological terms, a great deal is happening in the context of Big Data, such as the data-analysis language R. Another example is AMPlab at Berkeley University. With Big Data as its starting point, AMPlab orients itself to the combined forces of *Algorithms, Machines & People*. A Big Data milestone in the summer of 2012 was the appearance of the GraphChi software, developed at Carnegie Mellon University. This has enabled analyses on a common PC, where previously large computer clusters were occupied for hours performing such tasks. With a Twitter dataset from March 2010 as a benchmark, one single GraphChi PC turned out to be able to analyze this in 59 minutes. The previous occasion this was done, 1000 large computers spent 6.5 hours on the same task. The dataset in question is available free from the Infochimps.com website and contains 40 million users, more than 1.5 billion tweets, and 1.2 billion connections between users.

3 Big Social will fulfill the promise of hypertargeting

Regardless of how impressive this all may be, the big issue concerns the significance of it all. Not everyone is equally enthusiastic about Big Data; but many, including Jeff Dachis of the Dachis Group, hold the opinion that Big Data on social media forms the glorious hypertargeting future. Just think about it, says Dachis, hundreds of millions of people are busy on the social web, unguardedly sharing their whole lives with one another. It easily adds up to 500 billion dollars in brand engagement value.

At the beginning of 2012, Twitter had a total of 225 million accounts, and almost 200 million tweets were sent every day. In comparison: Facebook has more than 800 mil-

lion active users and LinkedIn has more than 135 million. This 'consumerization' of Big Data will only assume larger proportions in the future.

There is skepticism about the use of advertisements on Facebook in particular. Just before Facebook was launched on the stock market, General Motors slashed its advertising budget for the social network. But this same GM still spends three times that sum on engagement with people on Facebook. The major challenge is to measure the ROI of this action. For such tasks we now have advanced Social Analytics tools and new algorithms such as GraphChi.

4 Big Data runs the risk of becoming an out-of-control party

One of the applications of Social Analytics is to gather as much information as possible on the online behavior of people — Big Social Data — with the aim of predicting what they are going to do next and what they are going to buy. Peter Fader, a marketing professor at Wharton Business School and co-director of the Wharton Customer Analytics Initiative, inserts a few prominent question marks here. He compares the projected goldmine of Big Social Data to Customer Relationship Management, which made its breakthrough in the early 1990s. At first it was regarded as the Holy Grail, but nowadays a harder evaluation is given: it causes huge frustration and is much too expensive; in short, the IT party has run out of control. Fader is afraid that things will turn out the same way with the current Big Data hype.

5 Clearly determine your actual needs

Dragging the entire Twitter or Facebook 'fire hose' through some Social Analytics refinery is simply nonsensical, says Fader. If you wish to get involved in hypertargeting, you have to look at tweets at individual level and link them to the transactions that a person executes. But online and mobile do not jointly form the complete new world that the Big Social Data evangelists, in particular, would have us believe. Of course, more information can lead to new insights, but the question remains as to how many data is needed for this? How interesting is it, actually, to know where someone is shopping at any given moment and what he/she is looking at? And which information on this subject should we retain?

6 Behavior is often predictable without Big Data

Fader believes that the real golden age of 'predictive behavior' occurred about fifty years ago. At that time, consumer information was very scarce. In the 1960s, Lester Wunderman began what he called 'direct marketing'. That was genuine 'data science'. Everything that could be known about a customer was kept up to date. What the direct marketing pioneers eventually achieved was RFM: the relationship between *Recency*, *Frequency* and *Monetary value*. The effect of F upon M is evident. R was the great surprise: it is easy to convince people to repeat previous behavior, even if they only buy things sporadically. However, you have to reach them immediately. In the marketing business, everyone is familiar with RFM, but it often signifies little to e-commerce people. With lots of Big Data you will undoubtedly come to the same

conclusion, but that is a bit of a waste of all the time and effort expended. In this connection, a good eye-opener is the book *How to Measure Anything: Finding the Value of "Intangibles" in Business* by Douglas Hubbard, published in 2007. This is full of examples and tips to enable you to find out lots of things in a practical manner.

7 Big Data builds upon traditional data centricity

The majority of Big Data exercises take place mainly or even wholly in the existing data environment. If this environment is an advanced data warehouse with the corresponding tooling, substantial investment must have already been made here, large data sets will already be subject to examination, and organizations will not be particularly excited about investing money in Big Data and in solutions that are still currently under development, just for the sake of a few interesting experiments. Add the present-day economic predicament to this situation, and it will be apparent that the entire Big Data euphoria is not currently traveling under a lucky star. It cannot be denied that the data flow is increasing by leaps and bounds, but we have seen that coming for a long time now and it can be regarded as more of an evolutionary development.

8 Big Data ROI is simmering through

In the report *Big Data: The Next Frontier for Innovation, Competition, and Productivity* — which is still the directive publication *par excellence* — the McKinsey Global Institute indicates how easy or difficult it is to gather Big Data for each sector of the American economy, as well as what Big Data mining could contribute and what Big Data Maturity looks like, step by step. Despite partly fundamental reservations, this report nevertheless gives the impression that Big Data ROI can be gained within the foreseeable future. Eighteen months later, it still appears to be a little too early. It is even claimed that Big Data is not such an accurate term; Total Data Management or something similar might be better — is that what it's all about? The answer is: yes, this total data approach is essential and Big Data ROI is simmering through. We'll present some clear examples and directions in this research report.

9 Social Analytics is the current Big Data pet topic

Social Analytics, the station between Web Analytics and the so-called Next-Generation Analytics, is a powerful antidote to all too grumpy Big Data skepticism. Gartner sees the latter as the immediate future. Here again we are confronted by that treacherous word 'next' and thus with the discussion about relevance with respect to Social Analytics, and of whether the glass is half-full or half-empty. Sullivan McIntyre of Radian6, a part of Salesforce, sticks to the first option and emphasizes the following: *'It becomes increasingly possible to make guesses about future behavior.'* Paul Barrett of Teradata also puts things in a wider perspective when he states: *'We are still in the early, black-and-white-TV stage of Social Analytics.'* In fact, both viewpoints can be said to hold true since sobering realism and cautious expectations go hand in hand.

Join the conversation

Question 3
How is the value of Social Data represented on your company's balance sheet?

www.sogeti.com/vint/r2q3

Wanted! A multi-faceted orientation

The various discussion topics above and clear verdicts often being open require a multi-faceted orientation to be provided in this Big Social report. The main question now concerns the speed with which we will be able to switch from black-and-white to color, and then on to HD and 3D. In this research report, which we have simply named *Big Social*, we wish to supply you with sufficient confidence to be able to draw your own conclusions and to keep a watchful eye on developments. We do this online at http://vint.sogeti.com/bigdata, where we are pleased to consult with you about the very latest perceptions of VINT and a select group of esteemed connoisseurs in the field of Big Data.

In the eleven remaining sections of this report we look at the many facets of what we concisely call *Big Social.* First, we'll examine the rhythms of human activity, and then respectively: data explosion's potential, the 'Big Five' social sources, the marriage of Web and Social Analytics, the development of Next-Generation Analytics, the shift to data and algorithms, the distortion of the social media lens, the Big Data technology toolbox, the need of listening attentively, the strength of Big Social Data. Finally, as part of the summary we will touch upon the organization of privacy, the theme of our third Big Data report.

Although much more remains to be said, especially in these times of rapid Big Data development, these accounts and sketches are sufficient basis to form your own judgment, to draw conclusions, and to build innovative solutions in your own practice.

2 Rhythms of human activity

In June 2012, a remarkable 'emerging next Big Data practice' reached the news. Twenty-four hours in advance, it may already be clear where someone will be the next day, and sometimes the prediction has an accuracy of twenty meters (De Domenico, Lima and Musolesi, 2012). By involving the data of friends in the analysis of a separate people, a *Synchronized Rhythm of the City* can be ascertained. That insight lay dormant in a Big Data set, and originated from the smartphones of 200 participants. Without taking into account the behavior of friends, the accuracy of the location determination was only one kilometer. This improvement by a factor of 10 or 20 brought first prize to three British researchers in the Mobile Data Challenge competition run by the Nokia Research Center Lausanne.

To become familiar with human action, to predict something relatively simple such as the mobility patterns of people in a city for example, a great amount of personal data has to be processed. The iteration model needed for organizations to work on this comprises the following three-stage rocket, which, in this variant, is again related to the favorite social-mobile theme of location:

By examining an increasing amount of data from a growing number of sources, human actions are becoming increasingly predictable. The discipline concerned with this is often referred to as *Predictive Analytics*, and place and time are often favorite parameters, just as in the example presented above, where personal smartphone data was enriched with those of friends. In this case, the resulting algorithm produced an improvement that was at least ten times better. Although this is a research prototype, it is also a convincing demonstration of the power of more data from various sources in a real-time setting. In this form, that involves primarily *Variety* and *Velocity*.

For sure, there is no shortage of *Volume*. We leave behind digital traces, which often go further than place and time, wherever we go. Deliberately or unwittingly. When we go shopping, use a public transport pass, perform a search via Google, download an e-book, watch digital TV, listen to music via Spotify, are on the phone, send a mail, use a car navigation system, etcetera. Such data is available in abundance, to be read and utilized. Add the explosion of information that we share on social media such as Foursquare, Twitter, Zoover, Google+, Yelp, Pinterest, YouTube, Yammer and LinkedIn, and we have directly landed in a marketing heaven, where traditional focus-group research, surveys and sampling once constituted our limited resources in the framework of the trio Understand, Predict and Act. In our current Big Social era, that is now radically different.

What we want, from a marketing point of view, is to be able to zoom in and out on human behavior with an analysis instrument that is simultaneously a microscope and a telescope, and which also interprets a great deal of all the data too, so that we can make a timely and accurate offer. Or even better: can get paid immediately.

If a new product flops, we should not wait months to investigate the shopping sales figures. Signals from social media can give sufficient indication at an early date. This is called 'sentiment analysis', a concept that is well on its way to becoming the buzzword

of the year. After all, it is via social media that we consistently ventilate our opinions about everything and more. It is the raw material of tomorrow's predictive mechanisms. That is the promise and the prospect: we will be able to predict who is going to commit fraud, where the following burglary will take place, what customers will look for next week.

Already there are many new things to be known, provided we tap other than our traditional data sources. Social is the aspect that brings great opportunity, and with Big Data and new analytical methods and techniques, we can make full use of the opening offered. Hypertargeting and personalization have never been within such short arm's reach.

3   More data for better answers

Imagine that you have several explanations for a fact you have at your disposal. But, if you now have to choose, what exactly are you going to select? It's ten to one on that you will say exactly what an aggravated Sherlock Holmes also said 120 years ago:

> *'Data, data, data! I can't make bricks without clay!'*

In our Big Data age, this fundamental response from *The Adventure of Copper Beeches* (1892) is more relevant than ever. We need more data, preferably as many as possible. Making a choice on the basis of a lack of information is certainly a rather anemic solution and, according to Big Data evangelists, it is totally irresponsible. Guessing is extremely dangerous, in the context of chasing the truth, and also businesswise. This is particularly relevant if your competitors do not have to rely on guesswork because they have taken the trouble to perform more profound search.

Mystery-solvers — whether they are called Sherlock Holmes, Columbo, or Wallander — never know where they should seek their data. Detectives grope in the dark and that is very frustrating, certainly if there is pressure of time. Businesswise, there is often no further option: a choice must be made quickly. Fortunately, the Internet has made things much easier in this respect: Big Social Data can guarantee increasingly better insight. That is the major difference and, accordingly, following one's intuition is becoming increasingly synonymous with the path of least resistance nowadays. Alleged talents, such as our gut feeling, can easier leave us in the lurch, but digital data will not, and they are there simply for the taking, in conjunction with the necessary analytical methods and techniques. We can place them under the denominator of *Social Analytics.*

There are roughly two Social Analytics schools. The first (Roebuck, 2011) is fully oriented toward social media due to:

- the growth of social media
- the growth of social media analytics tools
- the growth of social businesses on the basis of these two.

The other school sees Social Analytics in a broader perspective and orients itself to the analysis and prediction of human action on the basis of diverse sources. Mary Wallace, Social Analytics Strategist at IBM, counts herself as a member of this group. We will follow this broader viewpoint.

The following account is a classic example of how, in this case, a department store chain could quite easily acquire rather intimate facts from purchase data. The turnover of the American department store Target rose from 44 billion to 67 within a period of 10 years because people were able to segment customer groups well and to make them purposeful and relevant offers. This was largely due to the efforts of data expert Andrew Pole, whose work it is to examine great quantities of data until a hard predictive value has been reached. For example, Pole scrutinized the sales of baby products in relation to changes in purchasing habits over the previous months — such as a lotion with less aroma because women tend to develop a more acute sense of smell during pregnancy, food supplements in the form of zinc or folic acid, and disinfecting sprays for the hands. Pole thus identified 25 products by means of which he could even determine the date the baby was due.

In Minneapolis, the local Target manager, who had no idea of how the offers were selected, could not explain to an angry father why his daughter was receiving discount vouchers for baby products. On closer inspection, it turned out that a data analysis of the store did get it right. This is all very smart and valuable for both parties, in principle, but customers must wish to participate and not feel that their privacy is being invaded. At the beginning of 2012, The New York Times Magazine published the much-discussed article on *How Companies Learn Your Secrets*, on the basis of this Target case and others. The cover immediately drew attention: *Hey! You're Having a Baby!*, which was constructed entirely of product

packaging. The pay-off was: *How Your Shopping Habits Reveal Even the Most Personal Information.*

## Hypertargeting as emerging Big Social trend

The Target practice is not just an interesting anecdote but rather an emerging Big Data trend. Big Social hypertargeting has even been industrialized by a company called MyBuys. It offers cross-channel personalization for online retailers and consumer brands. The company aims to drive engagement, conversions and increase revenue by capturing insights from individual behavior, then utilizing choice modeling algorithms to predict the products each consumer would most likely purchase. Underlying the MyBuys personalization engine is a Big Data repository of over 200 million consumer profiles and 100 terabytes of data, which the company uses to deliver real-time product recommendations.

This Big Data development corresponds to organizational change. Just like the emerging roles of Data Scientist and Chief Analytical Officer, hypertargeting involving Big Data underpins the importance of the relatively new Chief Customer Officer role. According to the CCO Council, a Chief Customer Officer is *'an executive who provides the comprehensive and authoritative view of the customer and creates corporate and customer strategy at the highest levels of the company to maximize customer acquisition, retention, and profitability.'*

4 Total Data Management: the 'Big Five' social sources

Numerous data may form the basis of behavior analyses, such as client cards, search terms on internet, purchases, and responses to discount vouchers. In addition to the traditional enterprise applications as a data source, there are currently at least four other data categories that nourish the 'emerging next Big Data practices' of Social Analytics in widest sense of the word 'social'. These are: mobile data and app data, search-engine data, sensor data, and semantic data (such as smart metering for example) and, of course, social media data.

Each of these 'Big Five' data sources has its own interesting characteristics. One is related to the way people perform searches on internet, another reveals the patterns behind purchasing. One type of data comes from one's own system, while another may come from an external system. Social Media Data concern the motives behind actions. We listen to these by means of 'listening platforms' and perform yet other analyses:
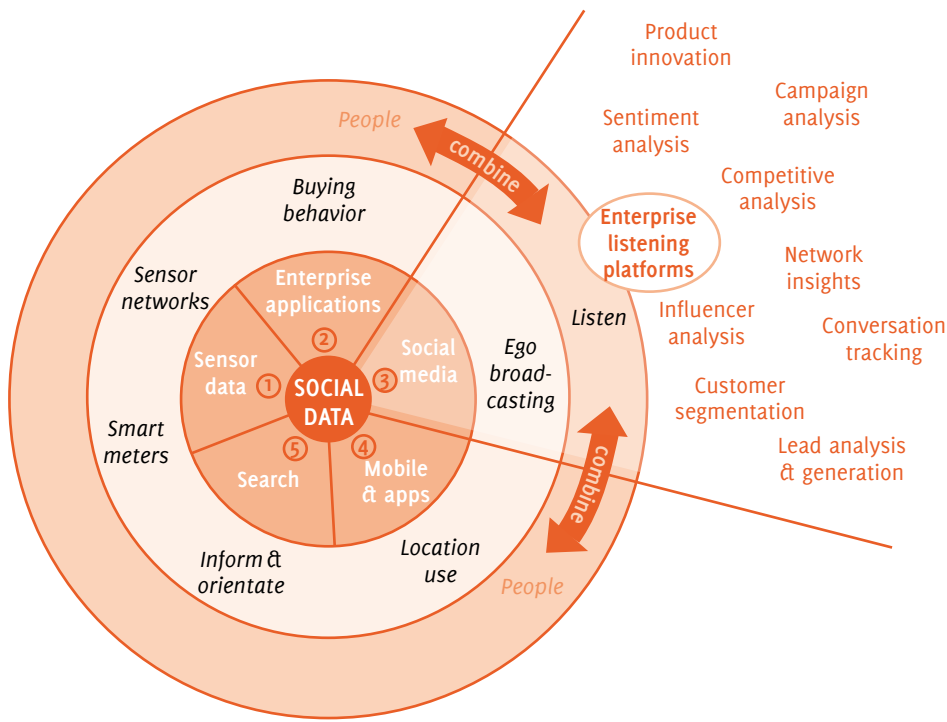
The 'Big Five' of Social Data, featuring the current Social Analytics practice for social media

It is essential to look at the whole picture and have a Total Data Management view, taking into account the strengths and weaknesses of data sets and the strength of smart combinations in the context of our three-stage rocket Understand, Predict & Act. But let us first examine each of the 'Big Five'.

1 Sensor data

These are data from (network) sensors, such as smart meters, which record the energy consumption and energy production of each household and neighborhood. The network consists of appliances that use energy, cars for the storage of energy in batteries, and people. This sort of 'neighborhood analytics' is a component of new production systems: a kind of 'Social ERP'. But other data, too, such as the tracking of purchasing behavior in shops ('in-store analytics' and 'anonymous analytics'), as well as numerous data from human-machine interaction also fit into this category.

2 Enterprise Application Data

Enterprise Application Data is traditionally used to recognize social patterns, such as purchasing behavior. It sits in structured databases of systems for Customer Relationship Management (CRM), Supply Chain Management (SCM), Enterprise Resource Management (ERP) or belongs to the data on a company's own website ('owned media'). In this context, we refer for example to on-site Web Analytics: the pages that people visit, the options that people click on, which 'landing page' is best in terms of

leading to purchases, etcetera. There are also the so-called Cross Channel Attribution tools, which analyze great amounts of data from diverse sources (in-store, on-line and off-line sales). Enterprise Application Data are the building blocks of Business Intelligence, HRM applications, of production and commercial processes.

3 Social Media Data

This is all about data, often unstructured, that comes from individuals who are engaged in 'ego-broadcasting' on social media. This data is accessible to organizations. Corporate data from internal microblogs or company-based innovation platforms may also be an important source. So-called 'social listening tools' are applied in the analysis, enabling the subsequent steps of marketing and HRM.

4 Mobile Data

This includes data from mobile applications, such as the popular apps category. Flurry and other players supply tools for App Analytics and make use of the same sort of metrics as those applied in web use. Mobile data may form the basis of location-based services that are supplied in real-time. Social media on mobile devices often also transmit location data. Social and mobile data together can be said to be a marriage made in marketing heaven.

5 Search Data and External Internet Data (off-site)

Search data may come from search engines, such as Google Trends or Google Insights, from other suppliers that 'scrape' the search engines, from 'phoning home' software or ISPs. The data is used for trend analysis, Search Engine Optimization and Search Engine Marketing, by making use of keyword monitoring and services of Google Adwords, for example. Off-site web data offers better insight into the popularity of websites, into where the buzz is, or where comments are given. Such data may come from the logfiles of webservers or from page-tagging with Java scripts. This also is likely to occur of course on a company's own website: 'on-site internet data'.

Total Data Management

A Total Data Management view on the basis of these 'Big Five' is not only a vista on which organizations can work, it is also one of the prevailing trends we see among 'analysis vendors' such as Alteryx, comScore, Datasift, IBM, Infochimps and Salesforce. On Twitter, comScore characterizes itself vigorously as follows:

> *'comScore #measures the digital world. We manage #bigdata to bring you #mobile, #search, #video insights and more.'*

Data aggregator Alteryx appears to be quite exceptional:

> *'The only Business Intelligence company to offer packaged data from Tom Tom, Experian Marketing Services, Dun & Bradstreet and the 2010 us Census, and firmographics from the world's leading business data supplier in every license.*

This data, which spans spatial, demographics, household and firmographic market information, provides businesses with a deeper understanding of where, why and with who events occur. By analyzing this market data in combination with their own data, businesses are able to perform analysis that drives highly targeted and localized decision making, and improves the overall ROI of every piece of data.'

Salesforce does something similar with Social CRM, by making combinations of social media data and CRM; or the Datasift company, which offers a cloud solution to enrich enterprise data with social media data.

IBM is currently working in Israel on an app to combine enterprise data such as historic purchase behavior and personal customer preferences: an interesting Big Data application for stores and supermarkets. The app places all data in an augmented-reality layer to guide customers through the shop via special personal triggers and offers.

'Beyond what IBM's augmented reality app may offer retailers on a customer-by-customer basis, it has the potential to create a treasure trove of metadata regarding product sales trends, hot in-store selling spots, traffic patterns and inventory issues, all of which could be used to maximize revenue per square foot, a critical metric in retail.'

Google Now also mixes all kinds of social data, for individuals rather than companies, on the basis of our location, calendar, mail, searches, and video choices. What we need can be presented just-in-time, taking into account flight times and traffic jams.

At present, a convergence of data and tools is occurring all around us. This intertwining happened quite early, on the basis of our behavior on the Internet. For a good understanding of Social Analytics, it is important to know how this discipline evolved from Web Analytics. Then we will examine where it is going.

5 How Web and Social Analytics became entwined

Taking into consideration the behavior of customers, prospects and everyone who is involved in our brand and our products, attention is currently being devoted to the (Social) Web. In many cases, that is where most activity, representative activity or activity that generates most buzz occurs. Human behavior on the Internet has been the subject of analysis for a long time now. Early forms of Web Analytics were in existence as far back as twenty years ago. Social Analytics arose parallel to this, but much

later, in 2006. In 2010, Web Analytics and Social Analytics embraced one another definitively.

In the framework of 'emerging next practices', let us examine how that happened on the basis of a short timeline, spanning the last decade of the previous century and this first of this one. The data comes from the *History of Web and Social Analytics (1990-2010)* infographic by Webtrends and DK New Media. Accordingly, we first give two short descriptions of these two domains:
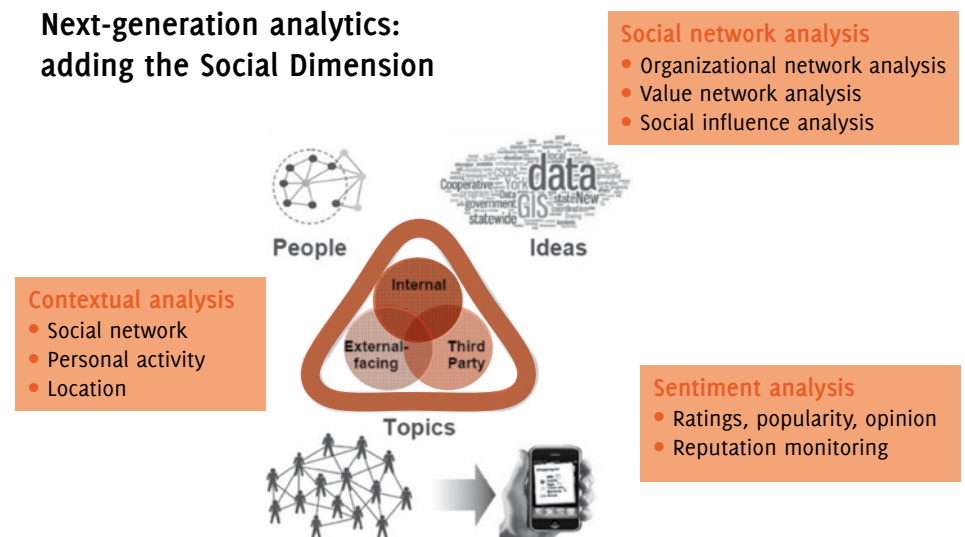
Web Analytics (1992–2010) according to Wikipedia:

> *'The measurement, collection, analysis and reporting of internet data for purposes of understanding and optimizing web usage.'*

Social Analytics (2006–2010) according to Gartner:

> *'Social analytics describes the process of measuring, analyzing and interpreting the results of interactions and associations among people, topics and ideas. [...] Social network [or media] analysis involves collecting data from multiple sources, identifying relationships, and evaluating the impact, quality or effectiveness of a relationship.'*
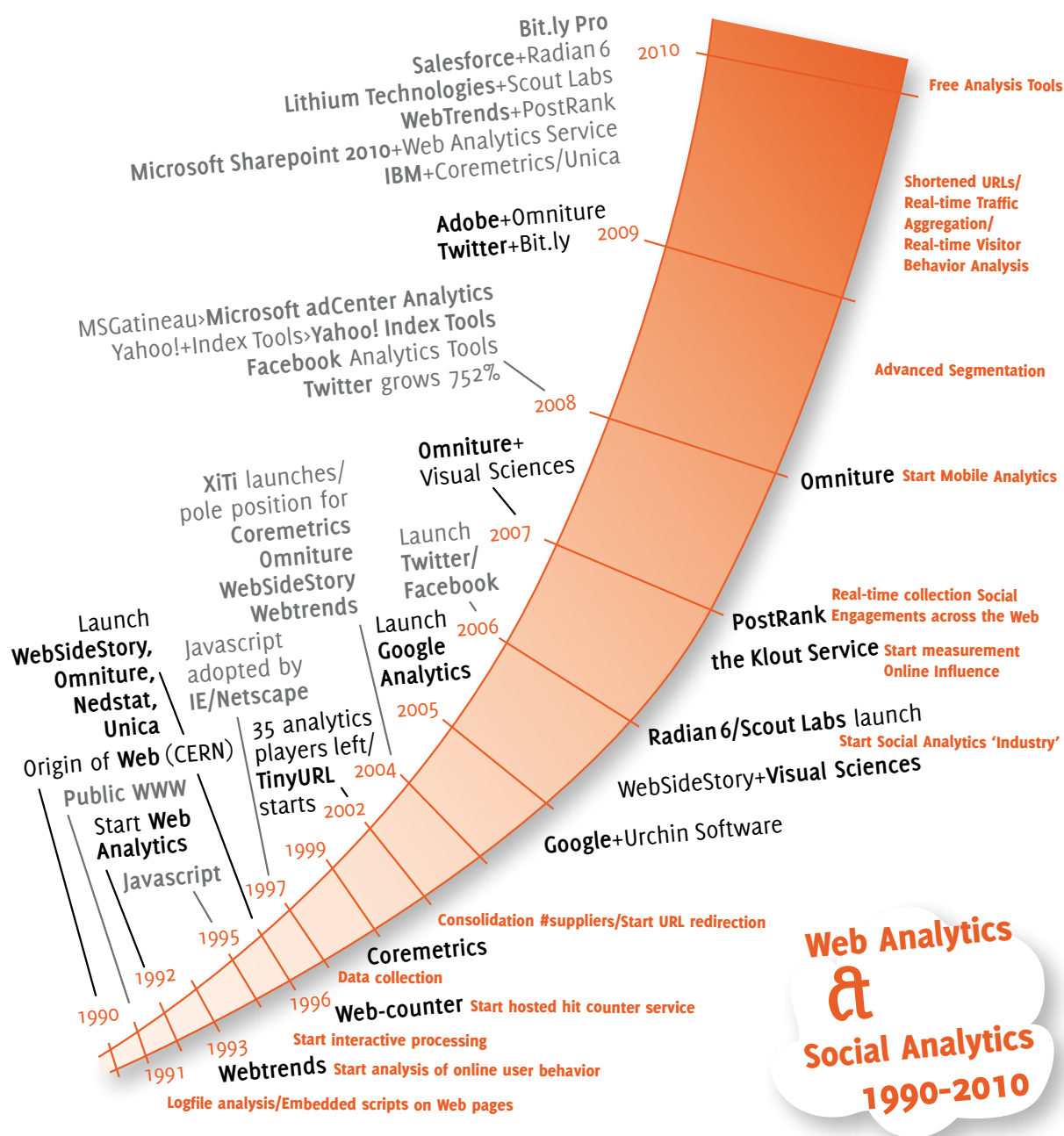

Next-generation analytics: adding the Social Dimension

Social network analysis
- Organizational network analysis
- Value network analysis
- Social influence analysis

Contextual analysis
- Social network
- Personal activity
- Location

Sentiment analysis
- Ratings, popularity, opinion
- Reputation monitoring

People — Ideas — Topics

Internal — External-facing — Third Party

Source: Gartner, http://www.scribd.com/doc/81893261/February-9-Top-10-Strategic-Tech-Dcearley

Social Analytics really arrived on the map in 2010, when Gartner included it in his top 10 of strategic technologies. Web and Social Analytics are now soulmates. The hectic dynamics depicted below in the illustration of *Web Analytics & Social Analytics 1990-2010* refer to companies that rise and merge (black), and 'emerging next

practices' (red), which eventually all become mainstream. This process is ongoing, of course, and the interesting thing is that, in these times of Big Data, Web and Social Analytics are evolving to a further stage, which Gartner terms Next-Generation Analytics. In the context of 'emerging next practices', section 7 of our research report looks toward a horizon beyond this concept of Next-Generation Analytics.

Bit.ly Pro
Salesforce+Radian 6
Lithium Technologies+Scout Labs
WebTrends+PostRank
Microsoft Sharepoint 2010+Web Analytics Service
IBM+Coremetrics/Unica

2010

Free Analysis Tools

Adobe+Omniture
Twitter+Bit.ly

2009

Shortened URLs/
Real-time Traffic
Aggregation/
Real-time Visitor
Behavior Analysis

MSGatineau›Microsoft adCenter Analytics
Yahoo!+Index Tools›Yahoo! Index Tools
Facebook Analytics Tools
Twitter grows 752%

2008

Advanced Segmentation

Omniture+
Visual Sciences

Omniture Start Mobile Analytics

XiTi launches/
pole position for
Coremetrics
Omniture
WebSideStory
Webtrends

Launch
Twitter/
Facebook

2007

Launch
Google
Analytics

2006

PostRank Real-time collection Social
Engagements across the Web

the Klout Service Start measurement
Online Influence

Launch
WebSideStory,
Omniture,
Nedstat,
Unica

Javascript
adopted by
IE/Netscape

2005

Radian 6/Scout Labs launch
Start Social Analytics 'Industry'

WebSideStory+Visual Sciences

Origin of Web (CERN)

35 analytics
players left/
TinyURL
starts

2004

Google+Urchin Software

Public WWW

2002

Start Web
Analytics

1999

Javascript

1997

1995

Consolidation #suppliers/Start URL redirection

Coremetrics
Data collection

Web Analytics

&

Social Analytics

1990–2010

1992

1996 Web-counter Start hosted hit counter service

1990

1993 Start interactive processing

1991 Webtrends Start analysis of online user behavior

Logfile analysis/Embedded scripts on Web pages

6 Toward Next-Generation Analytics

Web Analytics and Social Analytics will be used together for some time to come, but the current focus is mainly on Social Analytics. If we were to pay tribute to someone who really deserves praise in this field, it would be Lars-Henrik Schmidt. Since the early nineties, this Dane has been engaged in the development of a descriptive philosophical perspective, which he calls Social Analytics. Schmidt envisaged a discipline generally geared to reporting on the 'trends of these times'.

We now extract such trends from the trending topics of Twitter. Every minute we can check what people are communicating about, by counting how often a word with a so-called 'hashtag' (#) is tweeted. If we want to know the trends in different countries, we can use Twirus, for example. Social media enables direct up-to-date reporting. Enthusiasm about the use of social media tools is therefore very understandable. Analyses are available at the drop of a hat, there are free tools at our disposal so we can start immediately, and at the same time we can zoom out for the Big Picture and in on the (potential) buyers of products and services.

But Social Analytics — particularly Social Network or Media Analysis — is a transitional phenomenon, or at least, the excessive emphasis on this discipline is. The Gartner top 10 of promising trends for 2011 contained two datamining related domains: Social Analytics and Next-Generation Analytics. Of course we are moving toward this next generation: an integral Big Data approach to so-called Total Data Management, presented as a fundamental focus in our first research report *Creating clarity with Big Data*.

In the last section, we already saw Gartner's definition of Social Analytics. Next-Generation Analytics is typified thus:

> *'It is becoming possible to run simulations or models to predict the future outcome, rather than to simply provide backward-looking data about past interactions, and to do these predictions in real-time to support each individual business action.'*

How this 'Next Generation' progresses further in more detail is presented in the report entitled *A Framework for Social Analytics* by Susan Etlinger and Charlene Li, which was published by the Altimeter Group in August 2011. Here, too, the excessive emphasis on Social is merely a 'passing phase' so to speak, for *'Social is One of many Signals — Data are King'*, as we read in summary terms at the end of the report. Altimeter sees the future of Social Media Measurement as follows:

> *'"Social analytics" or "social intelligence" will become an integral — and eventually indistinguishable — element of the enterprise's ability to sense, interpret, and recommend actions based on signals from the market. [...] One of the*

greatest impacts of the transition to the adaptive business is the advent of "Big Data"—the algorithmic increase in unstructured data that will stem from continuous interaction with customers, communities, and markets. [...] Clearly this future state is several evolutionary steps away from social media monitoring and will require an entirely new processing and analytics approach that is able to make sense of both the unstructured nature and the sheer volume of data.'

With this, Altimeter adopts the same position as Gartner with its division into Social Analytics and Next-Generation Analytics, namely: we are on the road from A to B, from Web Analytics and Social Analytics to *Total Data Analytics*, as we like to call it, making the connection to our vision of *Total Data Management*.

Note: For a complete account see the section "Eight key Big Social definitions" at the end of this report.

7   Data and algorithms instead of models

The proper interpretation and linkage of data already leads to better decision-making, more sales, fewer risks and cost reduction. But, keeping in mind the 'emerging next Big Data practices' of Michael Chui, we can anticipate a great quantity of further improvements and possibilities:

'We are still in the early, black-and-white-TV stage of Social Analytics.'

This is the view of Paul Barrett, Customer Management Director at Teradata. Let us examine what that literally means. Black-and-white was the TV period in which there were few channels and we required different antennas to receive them. Very often the only thing to see on television was 'snow'. We were troubled by 'atmospheric disturbance', and saw only 'snow' (or 'noise') if a strong wind had turned the aerial a little. It was far from being the ideal situation, with only gray tints to represent a colorful world. And, we could only see the same program at fixed times.

It would be an exaggeration to denunciate Big Social in a comparable way, because modern Social Analytics is in far better shape. But things could be better: a single antenna please, sharper picture, more details, more channels and sources, real time, various angles, more aggregation levels, pattern recognition and, above all: the ability to predict behavior. We need to know what people really want, serve them in a timely way, bind them to us, and build up relationships with and via them. This striking improvement is the commercial Big Data challenge for organizations.

If a dataset is large enough, and up to date, the empirical approach often works better than a formula. We could formulate a complex model to determine how any people will go down with flu, but investigating search results produces a faster and clearer picture. Gunther Eysenbach, a professor at the University of Toronto was the first to do so, in 2006. His conclusion was:

> '*The Internet has made measurable what was previously immeasurable: the distribution of health information in a population, tracking (in real time) health information trends over time, and identifying gaps between information supply and demand.*'

The same applies to many other things, such as the best pricing strategy for selling secondhand articles. We find that immediately in eBay data, which gives better insight into matters such as inflation and consumer confidence. In short, all kinds of answers are latent in large data sets, and we can uncover these without having to concern ourselves with models.

As far back as 2008, Chris Anderson of *Wired* magazine spoke provocatively about a Big Data vista, in which even theory-forming and the scientific method would become superfluous; but, of course, data cannot speak for itself. At most, empiricism and theory play leapfrog and, thanks to the data explosion, the emphasis currently lies on data and algorithms rather than on traditional models. This development has been ongoing for a number of years now; compare, for example, the statistical approach with that of machine learning:

> '*Statisticians emphasize probabilistic models for learning, and techniques for quantifying variation in the estimated model that results from variation in the learning sample. For many machine learners, the algorithm is the model, and emphasis is placed on developing interpretable yet flexible methods of learning in challenging context (computer vision, natural language).*'
>
> http://www.crm.umontreal.ca/Machine06/index_e.html

Scenarios on the horizon

In terms of predictive power and ambition, the latest developments based on Big Data and algorithms reach much further than traditional Web Analytics supplemented by dashboards to monitor Twitter and Facebook traffic and to give rapid response. Before dealing with this topic in the following sections, let us look over the horizon toward some interesting and realistic scenarios.

1 Drawing conclusions from apparently unrelated facts

We recognize this from online shopping, where we receive recommendations for items that we could be interested in. Sometimes the items are logically connected to our purchases: if we buy a digital camera, we are informed of an extra battery or memory card. But what if seemingly completely unrelated offers turn out be predic-

tive in some way? For instance, does the speed at which you cycle past the bakery say anything about the chance of you going to the cinema this evening? Or is the moment of the day in which you play a game of Angry Birds indicative of the fact that you might be interested in a more expensive bottle of wine when you go shopping later? Or perhaps a 'like' on Facebook says something about your general health? Such links are latent in Big Data: patterns that have remained unrecognized until now. It is not inconceivable that systems themselves will go looking for correlations and that they will present options to us, proactively.

2   Creating lifelike personas

In marketing, and increasingly in IT — for they are inextricably linked in the context of Business Technology — we often work with characters: concrete personal descriptions that are characteristic of typical customers. At present, it is still a human task to develop personas on the basis of research and insight: who they are, what they are called, what jobs they have, which preferences, etcetera. By devising scenarios for personas, valuable improvements that are good for the customer, company, process, chain etcetera can be recognized. By combining the patterns behind all the customers, better, richer and more meaningful personas can be developed.

3   Predicting on the basis of 'live' behavior

Much analysis is still oriented toward facts from the past, combined with actions that are taking place at this moment. But what if the real-time component is applied more intensively? What if my telephone passes on exact data about what I think, feel and want, on the basis of my location, my actions in the previous twenty minutes, the pages that I looked up on the Internet, the apps I used, the sounds from my surroundings, my agenda, and the activities of my Facebook friends in order to determine my state of mind on the fly. Which themes are buzzing around in my head, what am I feeling, which urgent needs do I have, consciously or unconsciously? Google Now is an interesting impulse toward a working version of this facility, and the influence of technology such as Google Glasses, which aims at continuous presence, could signify a giant step further in this direction.

4   From predicting to subtly influencing

When we have come so far that we can estimate, with any degree of certainty, someone's current frame of mind, the following possibility immediately arises: which minor and major impulses can we give someone to ensure that he or she enters a particular mental state, one in which he or she is quite happy, is open to experiment, and ready to dispense cash? Perhaps this involves music from a mobile telephone or activating a picture on Facebook, in order to set the right tone. Perhaps the route needs to be adjusted a little so that we ride through a tree-lined avenue? If enough people participate, a subtle variation will result in optimal influence. This may have a beneficial effect: to help people lead a simple, valuable and happy life; but the possible darker side of misuse and intrusive manipulation is no less realistic.

5 Really smart organizations

Nowadays, everything in an organization is digital: e-mail, telephone, content, finances, access doors, light, climate, presentations, training courses, you name it. Can we discover more patterns here? Perhaps only in the larger organizations initially, where the quantity of data is big enough to gain significant insight, but eventually it will be available as a service to every company, for instance via the Big Data algorithm set of providers like MyBuys. It will enable benchmarking in relation to successful companies, an optimization of processes on the basis of the way in which work is currently executed, an understanding of problems and opportunities, policies to cope with risks, and perhaps even proposals for the more human aspects, such as training courses, vacancies or evaluations.

6 Validation by means of variation

At present, Big Data is still primarily a matter of drawing upon data flows that already exist and attempting to formulate conclusions. But what would happen if the system itself were to go searching for new data, if it could try out things itself, by means of intelligent machine learning, in order to see what the effects are? For example, it could send a news bulletin to certain people to see if they do anything with it? Or perhaps send an sms with a fact from all Big Data to examine whether or not a threatening transgression of certain measurements can be avoided? This is already very commonplace in the world of web advertisements: the division of all visitors into groups, each of which receives a slightly different variant of an advertisement, even placed in a slightly different position on the page. By measuring which configuration has the most effect, the placement of advertisements can become increasingly efficient over time. If Big Social systems can undertake autonomous actions, variation can be embedded for the purpose of seeking and achieving maximum impact.

7 The end of unpredictability

This is most interesting, at least as a concept: if society, trade & industry and government authorities are all convinced of the importance of data, of the importance of searching for patterns and of better predictions about all kinds of topics, can we create a situation in which we can look forward, six months ahead for example, with reasonable certainty? Can we then anticipate movements on the markets, in stock markets and innovations, for example? Much of the present economy exists thanks to unpredictability: someone who is prepared to take risks in exchange for payment. If this risk diminishes in the near future because we know how people will react to something, what the risks are, and what the chance is of something happening, what will then form the core of our economy? As long as we cannot predict the weather, any prediction of society is probably out of your grasp. Although — how predictable are people in fact? Dirk Helbing and his colleagues have received 1 billion euros from the European Union for their *Living Earth Simulator* or *Future ict Knowledge Accelerator and Crisis Relief System.* The name says it all!

Much interesting research is taking place on the relationship between the subconscious and the conscious human brain. There seems to be consensus about the idea that the subconscious mind is mainly responsible for our behavior and that, to put it simply, identity and the 'ego' are spectators rather than helmsmen. It remains difficult to assess our own behavior. We may think, for example, that we find it important to live healthily, but many actions nevertheless indicate that health does not enjoy genuinely high priority. Reality is much more complex, of course, with interesting feedback between behavior, result and 'ego', but it is plainly evident that concrete behavior says more about who we are than about what we think. Big Social can play a useful role in giving people more insight: who are you, what kind of work best suits your qualities, and which relationships have the best chance of success? Hard data as a crowbar to finally break open the 'conditio humana'.

9 Hackers galore

No system whatsoever can ever be completely safe. The technology around e-mail spam and cyber attacks has demonstrated that people with less than ideal intentions also make use of IT for their own profit. Big Social tools can also give valuable help to criminals. All kind of private data can be used to guess passwords. Or think about an automatic system that can simulate a good friend in order to separate people from their money or to acquire compromising data. In view of the amount of spam in the world, it is clear that people no longer shrink from causing enormous nuisance to gain benefit. So what would a criminal do with a perfect social database?

8   Social media as lens distortion

Following the metaphor of zooming in and out with various kinds of data lenses, any distortion of the lens can be of major importance. Before discussing the tools that are intended to improve your vision and expand your horizons, we shall first draw your attention to several notorious distortions.

1 Not all Facebook accounts are real

Marketers are crazy about Facebook 'likes', even if it only to corroborate the success of their own actions. Such likes should, of course, come from real flesh-and-blood clients. The BBC study entitled *Who likes my virtual bagels* illustrates that an absurd product can be successful on Facebook thanks to the likes from fake accounts in Egypt. Almost 9 per cent of accounts are fake, which Facebook itself admits. So that amounts to around 83 million accounts. In 4.8 per cent of the cases it involves double accounts; 2.4 per cent do not belong to real people — the aim of Facebook — but to an organization or a domestic pet; and 1.5 per cent of the Facebook profiles are used for spam, for instance.

2   Twitter is also a distorted lens

Almost half of all Twitter accounts are allegedly fake also. There is a well-known joke: if all humans perished tomorrow but computers kept on running, trending Twitter topics would continue for years. There are companies that do separate the wheat from the chaff and maintain databases with only real live twitterians and tweeps.

3   Do you understand what is being said?

Automatic text analysis is not perfect. The quality of Google Translate amply demonstrates this. Tweets and comments may be ironic or may be full of slang. So how can we determine what the message is genuinely about and whether it is positive or negative? Nevertheless, technology is also advancing in this field and some market players actually claim that their computers and software really do understand what is being said.

4   What exactly are you measuring?

If only it were so simple, that the success of advertising could be predicted on the basis of what people tweet or post on Facebook. Much research has been performed on the effect of advertising, providing a completely opposing picture. For example, the *Remember the ad, forget about the product* effect: humorous adverts score highly but sales do not always rise. And, with reference to washing powders, we know that they do not score highly on likeability but none the less can be very effective. Thus, the number of likes on Facebook is not a reliable standard for judging whether or not to continue with a certain campaign.

5   How reliable is ego-broadcasting?

Most biographers warn about autobiographies. Probably rightly, because we tend to enhance our own performance in self-written histories. We already broached that subject in our book *Me the Media* — particularly regarding all those 'hyper-egos' that display a positive correlation with Narcissus in Greek mythology. What does that say about social media? How authentic or 'true' are those messages? Perhaps the unreal world will be experienced as the real world in the near future, a 'hyper-reality' as Umberto Eco among others calls it, a Disneyland 'that can give us more reality than nature can'. But even without these philosophical thoughts it is good to realize that data from social media must often be taken with a pinch of salt.

6   Can you hear what is not being said?

The situation outlined above provides enough reason to listen to what is not being said. Sherlock Holmes once solved a murder because a dog did not bark. All the things that are not tweeted, facebooked and pinterested can also be regarded as important information. Even if it is only because we know that some things are not readily included in autobiographies — despite all the openness and transparency on the Internet.

9 The toolbox is bursting at the seams

Despite and, of course, thanks to the above-mentioned critical remarks about predicting on the basis of online comments, the Social Analytics toolbox for companies has grown exponentially in the past few years. The following is only a small selection:

- Insight by Adobe
- Biz360 by Attensity
- Networked Insights
- Visible Technologies
- Scoutlabs by Lithium
- Radian 6 by Salesforce
- Cognos Consumer Insights by IBM
- The Azure/Hadoop solution by Microsoft
- A whole range of tools for Twitter and Facebook, such as Osfoora, Tweetdeck, HootSuite, MentionMap
- The mobile-analytic tools by Google and Flurry.

For a fairly current overview of diverse so-called 'listening platforms' please visit Wikispaces (http://socialmedia-listening.wikispaces.com/Tools) or the Social Media Monitoring Wiki by Ken Burbary, co-author of the book *Digital Marketing Analytics*, which will be issued by Pearson Higher Education Publishers in early 2013. In July 2012, Ideya Business Marketing & Consultancy counted the number of resources available and registered almost 250 social analytics tools, of which around 50 are free. (Read a part of the report at http://ideya.eu.com/publications.html.)

Some suppliers bundle a great number of social-analytics application areas to form a *suite* or a *hub*. To give an idea of important top domains, we now present the suite components of Visible.

The Social Media Hub by Visible

These tools aim to enable us to delve deeper into the thoughts, intentions and behavior of people, with the goal of stealing a march on our rivals.

Social Analytics can be applied in a range of activities, from pre-sales, such as *lead generation*, to after-sales and *customer support*. With Social Analytics, we can assess the sales apparatus: identify and reward important people in the organization on the basis of the *influencer scores* from the chatter microblog on Salesforce for example. We can also perform analyses on a higher level of abstraction, and measure *sentiment* about certain customer experiences. This is all aimed at inspiring the marketing apparatus or product development. We measure the 'rumor around the brand', 'buzz', 'word of mouth', or make comparisons between our brand and that of our competitors. We segment markets, divide customers into types (personas) or into locations whose coordinates are automatically sent via mobile apps. We can measure *whether or not an advertising campaign has been successful*, regardless of whether it has been geared to viral and social-media campaigns or to traditional TV advertising and published advertisements. We can examine who has which influence on social media and what these people can mean to our brand. And so on.

All such areas have their own terminology and acronyms. We are acquainted with the expression *Brand Protection*, and *Influencer Marketing* is the term used for spotting and influencing important people. This is all secondary to the possibilities that arose in the Google era (SEO, link-building, etcetera) and the enterprise application age of CRM, ERP and SCM. There are abundant possibilities, but that is not strange in view of the fact that *Social* represents life itself. The analysis of behavior is interesting for numerous corporate applications and that is why tooling is expanding in all directions. The acquisition of such tooling begins with the question about what exactly we wish to achieve with the tooling.

Join the conversation

Question 6
How do you navigate the largely unchartered terrain of becoming a Social Business?

www.sogeti.com/vint/r2q6

Innovation
Collaborating with customers to drive future products and services

Brand health
A measure of attitudes, conversation and behavior toward your brand

Customer experience
Improving your relationship with customers, and their experience with your brand

BUSINESS GOAL

Marketing optimization
Improving the effectiveness of marketing programs

Operational efficiency
Where and how your company reduces expenses

Revenue generation
Where and how your company generates revenue

Altimeter has developed a framework for Social Media Analytics, which defines six rather obvious domains to link business sectors such as marketing, innovation and operations to targets. It is, indeed, the old familiar *Management by Objectives*. In the report we see that the focus turns successively to listening, acquiring insight, determining metrics and, of course, taking action. In the various domains, predictive models form the basis of the activities to be developed.

10 Start by listening attentively

Sullivan McIntyre of Radian6 states that social media data play an important role in the step from reactive analysis to predictive analysis. If you entwine your social media with other systems, *'it becomes increasingly possible to make guesses about future behavior'.* He presents three criteria that these data must meet:

1 Are the data real-time?
On line, social data moves fast as lightning. The freshness of the data is crucial.

2 Are there interesting metadata?
If we receive thousands of posts and have to analyze everything manually, the moment has soon passed. A rich set of metadata enables us to respond to trends quickly

3 Are the data integrated?
Data must be able to be linked, as much as possible, to other relevant sources in order to be able to undertake the appropriate action.

To be able to apply Social Media Analytics, we must listen attentively to what is being said. We do so by means of *Enterprise Listening Platforms*. Our three-stage *Understand-Predict-Act* rocket presupposes that information is sent, received and interpreted. In this process, the influencer scores are key, as is a profile of the people being followed, a control center to keep the data up to date, a network analysis to determine connections, and sentiment analysis.

Listening platforms
Recorded Future, a company financed by Google and the FBI, is the name of a listening platform with a very specific mission namely to *Unlock the Predictive Power of the Web*. The company registers events that have not yet taken place. The organization uses blogs, websites and social media as input. It searches on 'next week' and 'next month', and delivers answers to questions such as 'Where is Obama going in August?' or 'Which beer festivals are held in October?' The launch of new products and technologies by competitors can also be monitored in this way, or perhaps conversations about a brand such as Coca Cola.
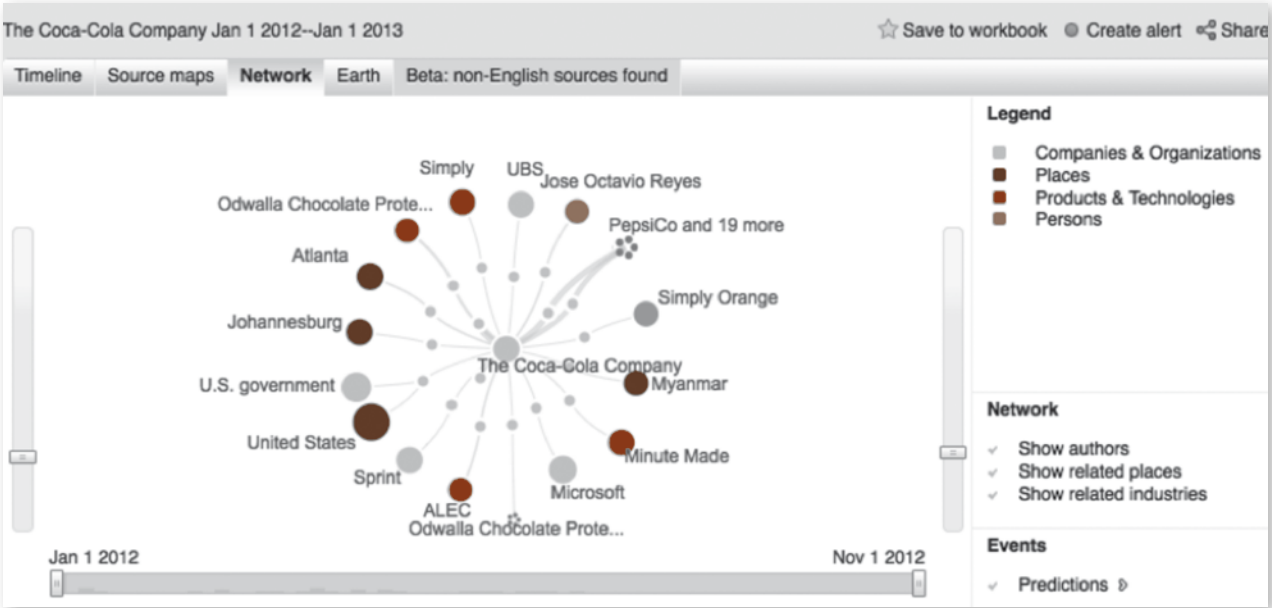
Influencer scores

There are also platforms that record the influence of people on social media; after all, we must maintain good relations with influencers. Twenty Feet has developed such a tool and calls it an 'ego-tracking service'. Ultimately someone is allocated a score, by the social network Klout, for example. Reports of influencers can be pretty detailed, such as this snapshot profile of Christopher Meinck made with Traackr.
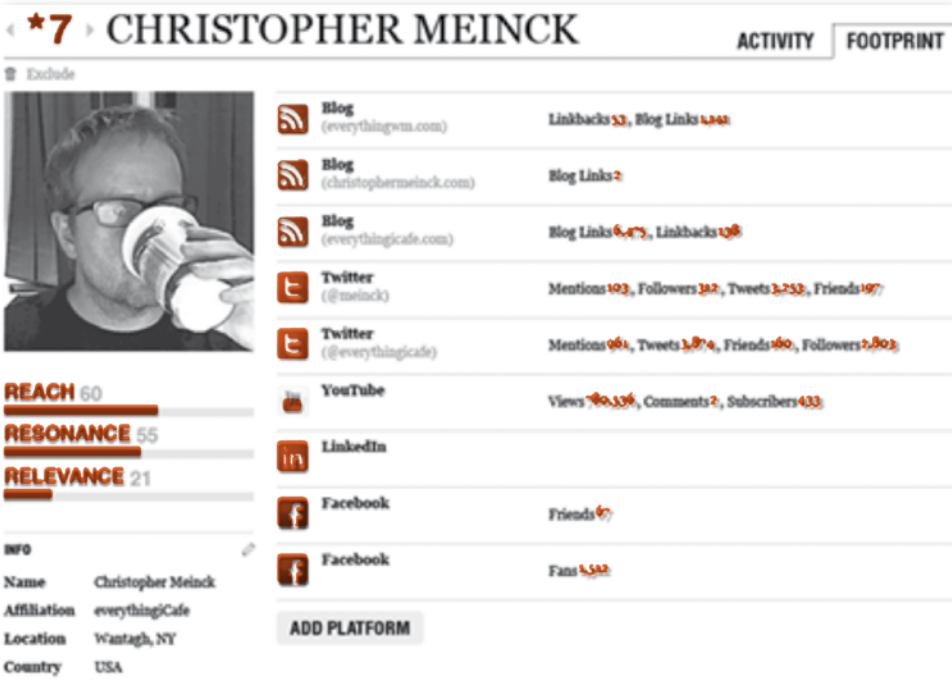
But much more can be known than merely whether or not someone is an influencer. Male or female, car driver or not, country of residence, hobbies, married, divorced, etcetera. This kind of information is stored on the servers of IBM, which has gathered all the Twitter accounts in the world and removed the 'bots' from them. A profile has been generated on the basis of the content of the tweets and the person's own profile. At the moment that, for instance, a media company wishes to know if a new film trailer will be enthusiastically received by a certain public, this type of Twitter profile can have a predictive effect. A few minutes after showing their trailer, an American media company was already informed that the intended target group was not enthusiastic, and could also discern those who were more positive.

The Social Media Listening Center

The picture below is one of a Salesforce Radian6 dashboard on various screens in a Social Media Listening Center. To consumers, it is very irritating when a certain service freezes or breaks down. Via their listening platforms, the webcare teams of organizations record such messages and undertake action on this basis.

KLM is one of the companies that work in this way. The next time that someone tweets, the previous conversation can also be directly accessed. These dashboards also provide overviews by region. People follow messages on a world atlas via language and country. In this way, sentiments about the last TV ad or about the treatment at the reception desk are presented on the dashboard in the control center.

Network analysis

MentionMap is an example of a *network analysis tool* for Twitter conversations. MentionMap displays tweets from people, and charts mutual relationships. As an example, we entered the Twitter-ID of Ben Lorica, Chief Data Scientist at O'Reilly Media. This produced the following MentionMap:

The conversations in which Lorica has participated led us to The Economist, Microsoft Research (MSFTResearch), Stanford University, Berkeley student Neil Conway, Rafeboogs and David Wilson, among others.

Sentiment analysis

FoodMood.in combines tweets with locations to measure the current mood with regard to food. The situation in the Netherlands is visualized below, with popular items such as pancakes, sushi, salad, as well as left-overs and carrots. Foodmood makes use of the sentiment classification of Stanford University, a 'trained' classifier that can cope with millions of tweets. Countries can be mutually compared, there is a list of the top-10 'happiest foods' of each country, and comparisons between the GDP of countries and the food scores are also possible.
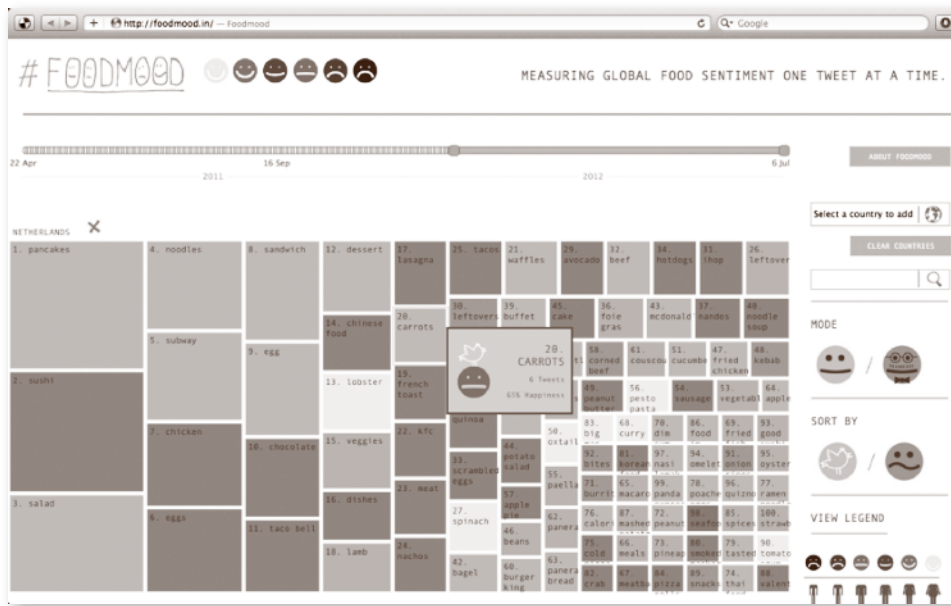
Besides general sentiment, we can also zoom in on individual tweets, so that the context becomes evident. An egg for breakfast, pancakes at a children's party, and chocolate on the sofa. We view the consumer in his or her domestic surroundings, as it were, and see how food is experienced.

In the UK, Bristol University measures emotions on Twitter and creates overviews of the situations:
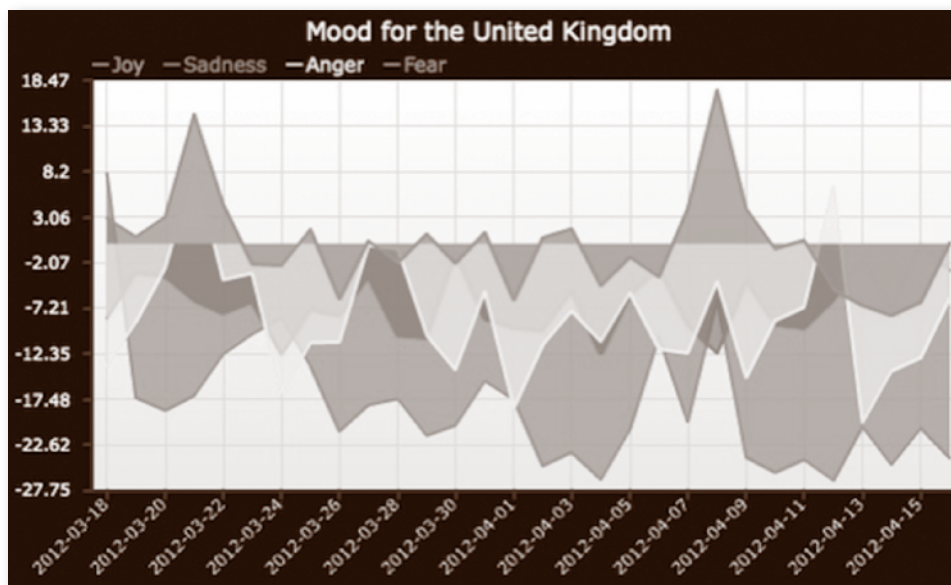
Fear and sadness were predominant during the summer riots of 2011, and also later when the government announced that public spending was going to be cut drastically. The School of Informatics and Computing in Indiana built a model to predict stock market fluctuations on the basis of this kind of mood analysis. It investigated whether or not public sentiment correlated with the value of the Dow Jones Industrial Average. After a detailed text analysis, considering words such as calm, alert, sure, vital, kind and happy, the conclusion was that, in 86.7 per cent of the cases, the daily rises and falls in the final value of the DJIA could be predicted fairly accurately.

In October 2011, the American Federal Reserve Bank announced that it would follow consumer confidence in much the same way, via Facebook, Twitter, blogs, YouTube, forums, Associated Press, CNN and the Wall Street Journal. The Global Pulse project of the United Nations makes use of sentiment analysis among the population to predict unemployment and suchlike. In this context, SAS found six indicators in expressions on social media, including items such as converting to a smaller car or postponing a holiday.

'SAS compared mood scores and conversation volume with official unemployment statistics to see if upticks in those topics were indicators of spikes in unemployment. The analysis revealed that increased chatter about cutting back on groceries, increasing use of public transportation and downgrading one's automobile could, indeed, predict an unemployment spike. After a spike, surges in social media conversations about such topics as canceled vacations, reduced healthcare spending, and foreclosures or evictions shed light on lagging economic effects. Such information could be invaluable for policymakers trying to mitigate negative effects of increased unemployment.'
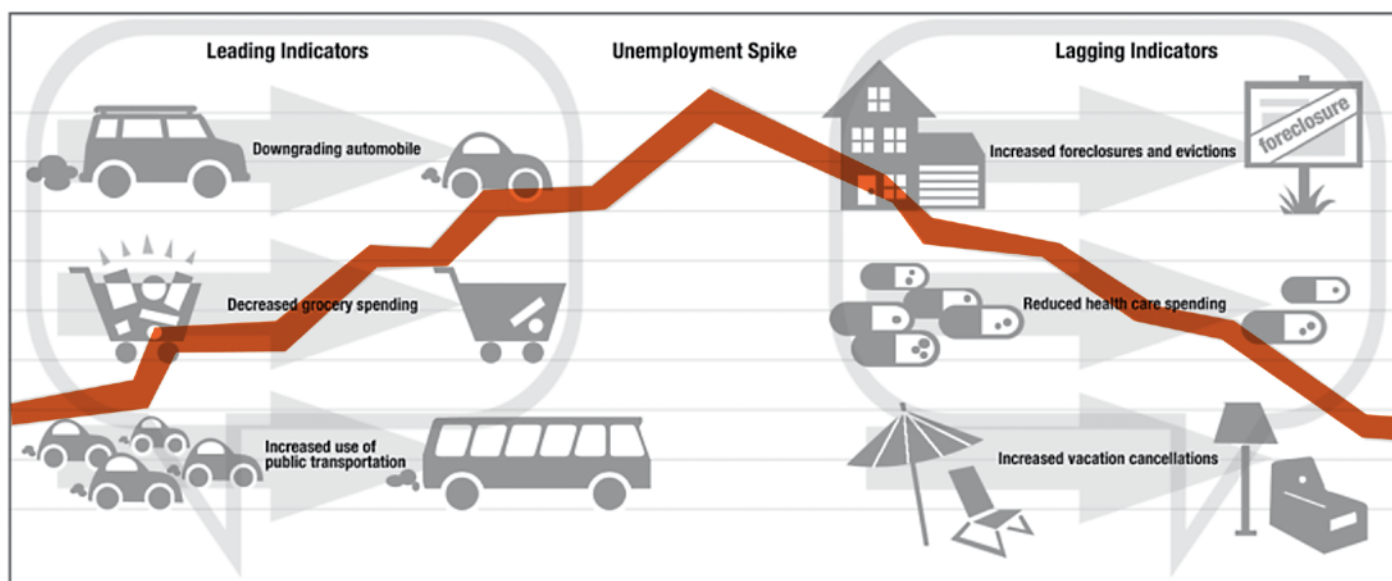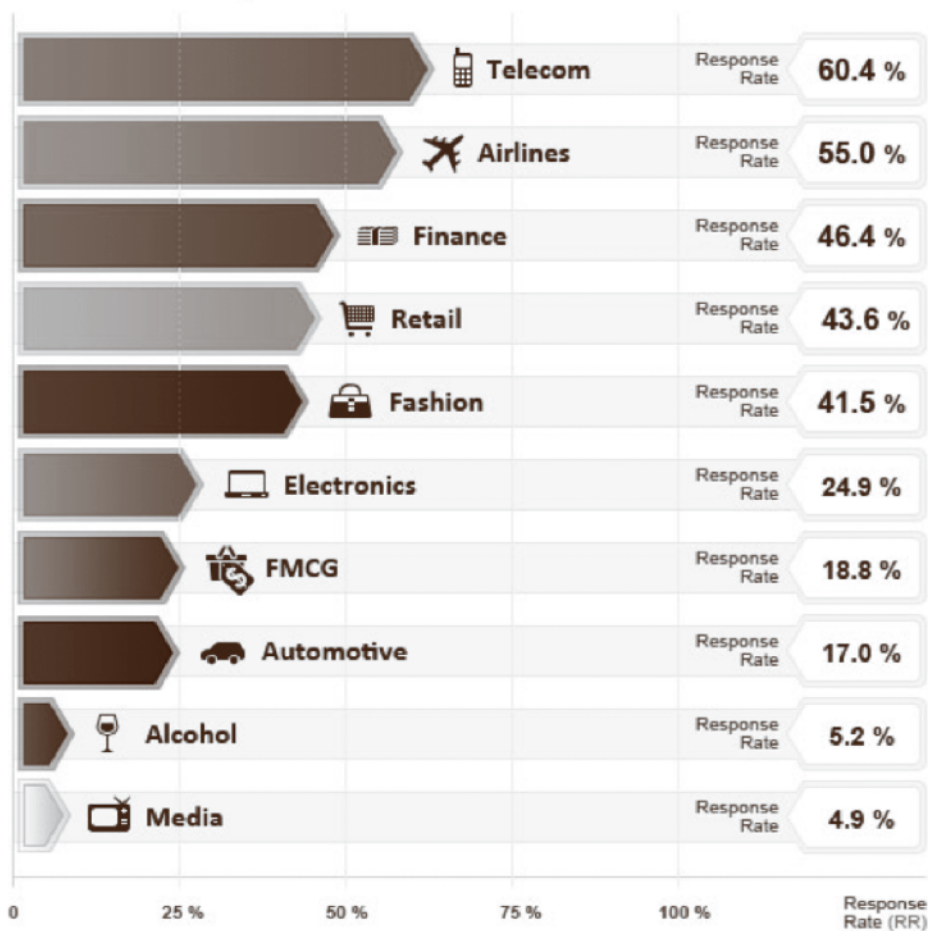

Analysis of social media using SAS shows increases in chatter about certain topics that are leading and lagging indicators of a spike in unemployment.

11 The strength of Big Social Data

In this section, we conclude by aligning four concrete Big Social cases in a cross-section of various sectors: retail trade, banks, insurers, police, and product development in the computer industry. It may be too early to draw general conclusions about the sector in which Big Social could have most impact. If the activity of sectors on Facebook is an effective predictor of the impact, the range could look like this:

Most Socially Devoted Industries on Facebook

Industry		Response Rate
Telecom	Response Rate	60.4 %
Airlines	Response Rate	55.0 %
Finance	Response Rate	46.4 %
Retail	Response Rate	43.6 %
Fashion	Response Rate	41.5 %
Electronics	Response Rate	24.9 %
FMCG	Response Rate	18.8 %
Automotive	Response Rate	17.0 %
Alcohol	Response Rate	5.2 %
Media	Response Rate	4.9 %

0   25 %   50 %   75 %   100 %   Response Rate (RR)

Response Rate
The percentage of user wall posts that get responded by the company.

Data is taken from brand Facebook pages in these industries from the 1st of March until the 30th of May 2012.

socialbakers

Independent of industry partitioning, we advocate a focus on processes and goals, as Altimeter presented in its above-mentioned report. The cases in this section concern cost-savings, product innovation, marketing and sales. Just as with the Target case in

section 3, we see that a purposeful shift toward Total Data Management currently and immediately bears fruit — low-hanging fruit — when the approach is 'social'. The case of Walmart Labs shows that preferences displayed on social media online and offline can be directly converted to recommendations. The case of the British company Wonga illustrates that Big Data can make more bank products profitable. For insurers it is important to prevent fraud and to limit outlay for health costs. Finally, we examine the role of Big Social in combating crime.

1 Walmart Labs & marketing innovation

Walmart Labs, the social data R&D department of Walmart, originated with the acquisition of Kosmix, a social media analysis firm that is primarily known for its Twitter filter tool, TweetBeat. The most important achievement of the Lab has been the development of a tool that performs semantic analyses on Twitter, Facebook and Foursquare. In this way, the so-called *Social Genome* can be charted, consisting of rich profiles of customers, topics, products, locations, and events.
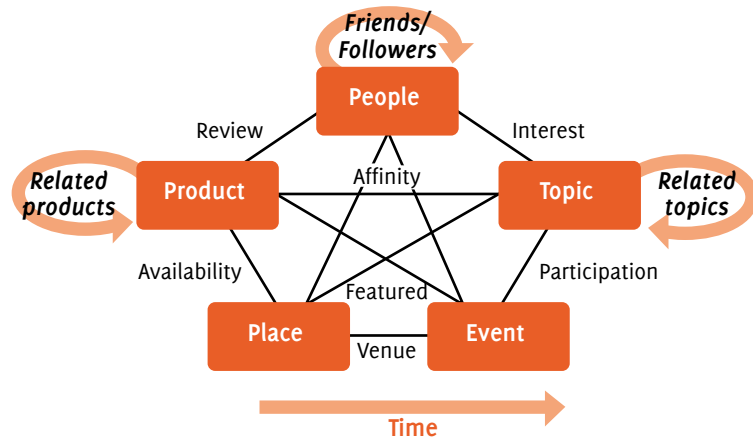
This concerns the interpretation of a network on the basis of relationships: a person is interested in a topic, a person attends an event, an event is related to a topic, an organization is associated with a product, etcetera. In this process, Walmart makes use of public data on the web, its own data, and data from social media.

The first results are now evident. By means of data taken from Facebook and Twitter, better recommendations can be given on the website and in shops, using the new Walmart app Shoppycat, for example, which provides gift suggestions to friends when someone's birthday is approaching.

2 Borrowing money with Big Data

Wonga, an English slang word for money, is a start-up that offers small, short-term loans without human intervention. On the basis of Big Data, Wonga cultivated this market and developed a profitable niche that had become unprofitable for the larger

Join the conversation

banks. People who need money often try to hide their sad financial state, but hard data does not lie, according to Wonga founder Errol Damelin.

It began with the experiment entitled *SameDayCash* and the beta-period ran as expected. For every successful loan, another defaulted. Nothing exceptional, but important facts could be gathered: about people and their behavior. SameDayCash fed the current Wonga algorithm. In the first year of operations, 100,000 loans were issued, with a collective value of 20 million British pounds. The algorithm examines all kinds of information sources, including social media. The system makes use of a basic set of 30 information points that are enriched with thousands of other data points. The algorithm exposes anomalies in the relationships. Because the loans are short term, there is a constant influx of new data. Financially this approach is feasible because it involves only small amounts.

3 Fraud detection as a Big Social killer app for insurance
Fraud costs the American insurer Property & Casualty around 30 billion dollars a year: approximately 10 per cent of all claims for damages. By linking historical data to demographic profiles, the chance of fraud can be estimated. A person's network says much more. If three Facebook friends have already been caught, an extra check on a person may turn out to be quite smart. Thus, with richer profiles, Predictive and Social Analytics help reduce the most important expenditures of insurers, such as fraud and risk analysis. Text analysis also plays a major role in scrutinizing damage claim forms.

Traditionally, risk analyses may target social-demographic data, driving behavior, credit information and the like. New data points contribute to better risk analyses. Life insurance companies adjust their prices on the basis of medical data that suggest a healthy lifestyle. Many Facebook likes for extreme sports on a profile could perhaps lead to higher premiums in the future.

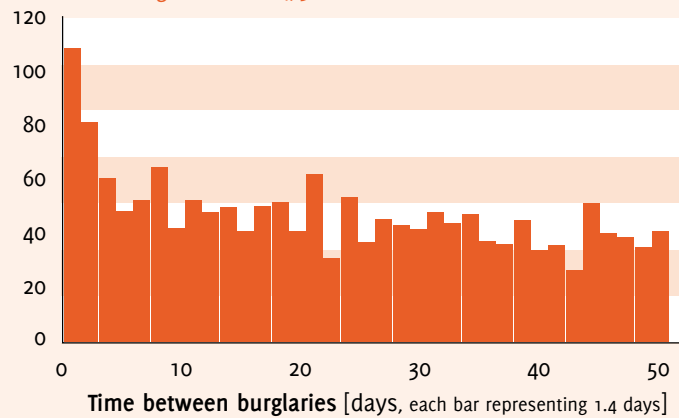### 4 Combating crime with the Predictive Police
In the Netherlands, police officers go on duty with a smartphone in order to be able to pick up signals in the neighborhood from social media. In this way, they can show their faces before something serious happens in the schoolyard, for example. This kind of initiative is supported and shared by *Police 2.0*, a community that pursues the Big Social transition in their field.

In Santa Cruz, California, the police make use of an algorithm that helps prevent crime. The software was developed by two mathematicians, an anthropologist and a criminologist from PredPol, a 'predictive police' company, based on a model that predicts the aftershocks of an earthquake. Practice indicates that a criminal often returns to a spot where he or she has previously been active. *Aftercrimes* follow the same pattern as the *aftershocks* of an earthquake.

Number of earthquake pairs separated by $\leq$ 110 km
in Southern California in 2004/5

Time between shocks [days]
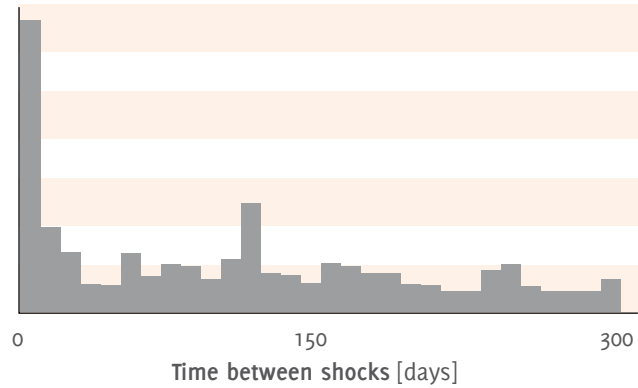
Number of burglary pairs separated by $\leq$ 200 m
in Los Angeles in 2004/5

Time between burglaries [days, each bar representing 1.4 days]

For some time now, various police departments have been making use of software by CompStat in order to predict crimes. These computer programs had only data at their disposal from at least a week old. The PredPol algorithms include real-time data instead. Taking into account location, time and type of crime, the software is capable of defining *prediction boxes* with an accuracy of 500 square meters.

In the first sixth months of the pilot project, the number of crimes fell by a quarter. It works quite simply: '*The suspect shows up in the area where he likes to go. They see black-and-white talking to citizens — and that's enough to disrupt the activity.*'

PredPol is an inspiring case but all kinds of traces on social media remain a valuable source too, for instance when riots occur like in some British cities in August 2011. Another good practice is to match police data with governmental sources. Even without algorithms a lot can be gained from smarter digital detective work: both in cyberspace and in the real world.

12 Summary and the organization of privacy

The adoption of plans in organizations for Big Data currently and predominantly covers the theme of Big Social: the customer side, inspired in particular by the social network activity of Web 2.0. But, if we take the concept of 'social' in a broader sense, an increasing amount of Big Data potential is released. This is more or less the route we have followed since the early nineties: first with Web Analytics, then with Social Analytics and now with Next-Generation Analytics. In this age of Big Data, further development is progressing toward Total Data Analytics and Total Data Management.

An important part of the discussion revolves around the extent to which organizations should embrace Big Social Data. The answer is: only on the basis of a well-grounded policy. Smart entrepreneurship in the growing dataflow is the key to capturing the raisins from the pie, so to speak. The question as to whether or not an organization initially is working with real Big Data (sets) is actually irrelevant. Scaling up will occur organically, and a good number of privacy issues are closely attached to this situation. We shall deal with these comprehensively in our third research report. Number four will be devoted to a Big Data roadmap, perceived and approached from various angles.

Modern Social Analytics applications enable organizations to understand the rhythms of human activity, to attach predictions to them, and to plan and implement corresponding actions: Understand, Predict & Act. The possibilities of personalization and hypertargeting are steadily increasing, and the toolbox is bursting at the seams. But do customers want that? It gives many of us a somewhat uncomfortable feeling to realize what commercial organizations know about individuals and groups. The organization of privacy and the guarantee of our personal integrity is perhaps therefore the domain *par excellence* to which attention should be paid. Big Data, Big Social and Big Brother are not worlds apart — certainly not in our human perception.

The beginning of this second research report was devoted to the issue of *What's Next in Big Data?* Many organizations are happy to be mere spectators at the moment, because there are no qualified best practices as yet and, this being the case, the financial risk could be considerable. But customers have a completely different mindset. Primarily they are alarmed by, for example, prospects of rising insurance premiums because they have presented themselves on the Internet a bit too enthusiastically, participating recklessly in certain leisure time activities, or showing themselves to be great fans of cigarettes and beer, to name just a few lifestyle choices. Regardless of what organizations may think about Big Data and Big Social, customers' Big Brother fear will force them to deal seriously with the situation, to adopt standpoints, and to express these vigorously.

Technology is advancing rapidly, we can make ever-better predictions, and we can step effortlessly from Web and Social Analytics on to Next-Generation Analytics. The accent is increasingly being placed on data and algorithms rather than on models. In short: the commercial power of Big Social Data is undeniable and is growing. At the very least, this entails increasing guarantees and responsibilities where themes such as privacy, personal integrity and, above all, perception and sentiment are involved. This is perhaps the very first observation for organizations to make with both feet firmly on the ground.

At present, Big Data is a dynamic and trending discussion theme. In the field of technology, of organization, of ROI, etcetera. For this reason, we are eager to keep in contact with organizations and individuals with regard to all 'next practices' that are currently being developed and evaluated: online at http://vint.sogeti.com/bigdata and, of course, in personal conversations.

Eight key Big Social definitions

As we have seen in this report, the realm of what we conveniently call Big Social contains different analytics and analysis flavors. For that reason this separate section is dedicated to the five main topics in this field. Eight definitions are provided in a logical order: one for Predictive Analytics, one for Sentiment Analysis, one for Web Analytics, four for Social Analytics, and one for Next-Generation Analytics.

The focus these days lies on Social Analytics. We start with the original and broadest angle by Lars-Henrik Schmidt, followed by the social media dominance of Kevin Roebuck, move on to the standard Gartner definition, and conclude with the converging social information and systems view by Mary Wallace. It is this converging way of looking at Big Social that underlies this report.

The last definition concerns Next-Generation Analytics. It is open enough to accommodate all future Big Data development that is already explicitly appreciated in the definition of Predictive Analytics.

Predictive Analytics (Wikipedia)

'Predictive analytics encompasses a variety of statistical techniques from modeling, machine learning, data mining and game theory that analyze current and historical facts to make predictions about future events. […] Technology and Big Data influences on Predictive Analytics: […] The volume, variety and velocity of Big Data have introduced challenges across the board for capture, storage, search, sharing, analysis, and visualization. Examples of big data sources include web logs, RFID and sensor data, social networks, Internet search indexing, call detail records, military surveillance, and complex data in astronomic, biogeochemical, genomics, and atmospheric sciences. […] It is now feasible to collect, analyze, and mine massive amounts of structured and unstructured data for new insights. Today, exploring Big Data and using predictive analytics is within reach of more organizations than ever before.'

Sentiment Analysis (Mejova, 2009)

'As a response to the growing availability of informal, opinionated texts like blog posts and product review websites, a field of Sentiment Analysis has sprung up in the past decade to address the question What do people feel about a certain topic? Bringing together researchers in computer science, computational linguistics, data mining, psychology, and even sociology, Sentiment Analysis expands the traditional fact-based text analysis to enable opinion-oriented information systems. Sentiment Analysis is closely related to (or can be considered a part of) computational linguistics, natural language processing, and text mining. Proceeding from the study of affective state (psychology) and judgment (appraisal theory), this field seeks to answer ques-

tions long studied in other areas of discourse using new tools provided by data mining and computational linguistics. Sentiment Analysis has many names. It's often referred to as subjectivity analysis, opinion mining, and appraisal extraction, with some connections to affective computing (computer recognition and expression of emotion). […] These are usually single words, phrases, or sentences. […] Sentiment that appears in text comes in two flavors: explicit where the subjective sentence directly expresses an opinion ("It's a beautiful day"), and implicit where the text implies an opinion ("The earphone broke in two days"). Most of the work done so far focuses on the first kind of sentiment, since it is the easier one to analyze.'

Web Analytics (Wikipedia)
'The measurement, collection, analysis and reporting of internet data for purposes of understanding and optimizing web usage.'

Social Analytics (Wikipedia)
'Social Analytics is a philosophical perspective developed since the early 1980s by the Danish idea historian and philosopher Lars-Henrik Schmidt. The theoretical object of the perspective is socius, a kind of "commonness" that is neither a universal account nor a communality shared by every member of a body. […] It might be said that the perspective attempts to articulate the contentions between philosophy and sociology. The practise of Social Analytics is to report on tendencies of the times.'

Social Analytics (Roebuck, 2011)
'Social Analytics refers to the tracking of various media content such as blogs, wikis, micro-blogs, social networking sites, video/photo sharing websites, forums, message boards, and user-generated content in general as a way for marketers to determine the volume and sentiment around a brand or topic in social media.'

Social Analytics (Gartner, 2010)
'Social analytics describes the process of measuring, analyzing and interpreting the results of interactions and associations among people, topics and ideas. These interactions may occur on social software applications used in the workplace, in internally or externally facing communities or on the social web. Social analytics is an umbrella term that includes a number of specialized analysis techniques such as social filtering, social-network analysis, sentiment analysis and social-media analytics. Social network analysis tools are useful for examining social structure and interdependencies as well as the work patterns of individuals, groups or organizations. Social network analysis involves collecting data from multiple sources, identifying relationships, and evaluating the impact, quality or effectiveness of a relationship.'

Social Analytics (Wallace, 2011)

'If we look at the academic definition of social analytics "the process of measuring, analyzing and interpreting the results of interactions and associations among people, concepts, and facts" and apply this more broadly to the business, then a couple of things happen. Firstly we start to be able to harvest actionable social insights from existing enterprise applications, secondly we create a bridge that allows us to marry legacy business solutions with the new generation of social business platform, and thirdly we significantly increase the ROI we can realize from our social investment.'

Next-Generation Analytics (Gartner, 2010)

'It is becoming possible to run simulations or models to predict the future outcome, rather than to simply provide backward-looking data about past interactions, and to do these predictions in real-time to support each individual business action.'

Literature

AMP Lab: UC Berkeley Algorithms, Machines and People Lab, http://amplab.cs.berkeley.edu/

Anderson, C. (2008): "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete", http://www.wired.com/science/discoveries/magazine/16-07/pb_theory

Barrett, P. (2012): "10 questions CMOs need to ask about social media", http://www.asterdata.com/blog/2012/04/26/10-questions-cmos-need-to-ask-about-social-media

Cellan-Jones. R. (2012): "Facebook advertising: Who likes my virtual bagels?", http://www.bbc.co.uk/news/technology-18812126

Center for Economics and Business Research (2012): *Data equity: Unlocking the value of big data*, http://www.sas.com/offices/europe/uk/downloads/data-equity-cebr.pdf

Chief Customer Officer Council (2012): "The Role of the CCO", http://www.ccocouncil.org/site/the-role-of-the-cco.aspx

Conan-Doyle, A. (1892): "The Adventure of the Copper Beeches", http://sherlockholmes.wikia.com/wiki/Story_Text:_The_Adventure_of_the_Copper_Beeches

Dachis, J. (2012): "Big Data Is The Future Of Marketing", http://www.businessinsider.com/big-data-is-the-future-of-marketing-2012-7

Domenico, M. De, A. Lima and M. Musolesi (2012): *Interdependence and Predictability of Human Mobility and Social Interactions*, 2012. (Read also M. C. González, C. A. Hidalgo, Barabási: *Understanding Individual Human Mobility Patterns*.)

Duhigg, C. (2012): "How Companies Learn Your Secrets", http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html

Economist, The, Insurance data (2012): "Very personal finance. Marketing information offers insurers another way to analyse risk", http://www.economist.com/node/21556263

Economist, The, Special Report: International Banking (2012): "Crunching the numbers. Banks know a lot about their customers. That information may be valuable in more ways than one", http://www.economist.com/node/21554743

Etlinger, S. (2011): "Research Report: A Framework for Social Analytics", http://susanetlinger.wordpress.com/2011/08/10/research-report-a-framework-for-social-analytics, http://www.altimetergroup.com/research/reports/a-framework-for-social-analytics

Eysenbach, G. (2006): "Gunther Eysenbach coins the term 'Infodemiology' and wins AMIA award", http://www.ehealthinnovation.org/?q=node/89

Fader, P. (2012): *Customer Centricity: Focus on the Right Customers for Competitive Advantage*, http://executiveeducation.wharton.upenn.edu/resources/upload/Wharton-ExecEd-Customer-Centricity-excerpt.pdf (excerpt)

FuturICT: http://www.futurict.eu

Gartner (2010): "Gartner Identifies the Top 10 Strategic Technologies for 2011", http://www.gartner.com/it/page.jsp?id=1454221

Gartner (2010): "Next-Generation Analytics: Adding the Social Dimension", http://www.scribd.com/doc/81893261/February-9-Top-10-Strategic-Tech-Dcearley

Global Pulse: http://www.unglobalpulse.org/

Gomes, L. (2012): "Is There Big Money in Big Data?", http://www.technologyreview.com/news/427786/is-there-big-money-in-big-data/

GraphLab: GraphChi, http://graphlab.org/graphchi/

Guazelli, A. (2012), Predicting the future … in four parts: [1] What is Predictive Analytics? [2] Predictive modeling techniques [3] Create a predictive solution [4] Put a predictive solution to work http://www.predictive-analytics.info/2012/07/predicting-future-in-four-parts.html

Hardy, Q. (2012): "Big Data for the Poor", http://bits.blogs.nytimes.com/2012/07/05/big-data-for-the-poor

Hausmann, V. et al. (2012): "Developing a Framework for Web Analytics", http://uni-koblenz-landau.academia.edu/PetraSchubert/Papers/1862203/Developing_a_Framework_for_Web_Analytics

Hickins, M. (2012): "Taking Small Steps to Big Data", http://blogs.wsj.com/cio/2012/05/24/taking-small-steps-to-big-data/

History of Social Media from 550 BC to 2010: http://www.webanalyticsworld.net/wp-content/uploads/blogger/10-socialMediaTL_05.png

Hubbard, D. (2007): "How to Measure Anything: Finding the Value of 'Intangibles' in Business, http://www.hubbardresearch.com/wp-content/uploads/2011/08/TAC-How-To-Measure-Anything.pdf,http://howtomeasureanything.com/

Ideya (2012): *Social Media Monitoring Tools and Services Report 2012*, http://ideya.eu.com/images/SMMTools ReportExcerpts 09072012Final.pdf

McIntyre, S. (2012): "From Reactive to Predictive Analytics", http://www.radian6.com/blog/2011/08/from-reactive-to-predictive-analytics/

McKinsey Global Institute (2011): *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation

MyBuys (2012): "MyBuys Named the Leader in Personalization for Third Year in a Row", http://finance.yahoo.com/news/mybuys-named-leader-personalization-third-110000414.html

Nerny, C. (2012): "Point Smartphone, Get Data: IBM to Unveil Augmented Reality App for Retailers", http://data-informed.com/point-smartphone-get-data-ibm-to-unveil-augmented-reality-app-for-retail

Nokia Mobile Data Challenge, http://research.nokia.com/page/12000

Police 2.0: The Possibilities of the Digital Revolution for the Dutch Police Force (in Dutch), http://criminaliteitswijzer.ning.com/

Recorded Future: Unlock The Predictive Power Of The Web, https://www.recordedfuture.com/

Rezab, J. (2012): "70% of Fans Are Being Ignored By Companies — Now what?", http://www.socialbakers.com/blog/655-70-of-fans-are-being-ignored-by-companies-now-what

Roebuck, K. (2011): *Social Analytics: High-impact Emerging Technology — What You Need to Know: Definitions, Adoptions, Impact, Benefits, Maturity, Vendors*

SAS (2012): "Claims Fraud: Prevent fraud before claims are paid", http://www.sas.com/industry/ins/fraud.html

SAS, UN (2012), "Can a country's online 'mood' predict unemployment spikes?", http://www.sas.com/news/preleases/un-sma.html

Shaw, W. (2012): "Cash machine: Could Wonga transform personal finance?", http://www.wired.co.uk/magazine/archive/2011/06/features/wonga

Talbot, D. (2012): "A Phone That Knows Where You're Going", http://www.technologyreview.com/news/428441/a-phone-that-knows-where-youre-going/

Ungerleider, N. (2011): "The Federal Reserve Plans To Monitor Face-
book, Twitter, Google News", http://www.fastcompany.com/1786730/
federal-reserve-plans-monitor-facebook-twitter-google-news

VINT (2008): *Me the Media: Rise of the Conversation Society*, http://www.methemedia.com/

VINT (2012): "Creating clarity with Big Data", http://blog.vint.sogeti.com/?p=4559

Wallace, M. (2011): "Social Analytics is more than just Social Media…", http://allthingsanalytics.
com/2011/11/04/social-analytics-is-more-than-just-social

Wallace, M. (2011): "What's in a name?", http://allthingsanalytics.com/2011/10/13/
whats-in-a-name

Walmart Labs: "Social Genome", http://www.walmartlabs.com/social/social-genome/

Webtrends/DK New Media (2011): "History of Web and Social Analytics", http://www.market-
ingtechblog.com/infographic-history-web-social-analytics/

Wiki: Social Media Listening, Monitoring, Measuring, and Management Tools, http://social-
media-listening.wikispaces.com/Tools

Wiki: A Wiki of Social Media Monitoring Solutions, http://wiki.kenburbary.com/
social-meda-monitoring-wiki

About Sogeti

Sogeti is a leading provider of professional technology services, specializing in Application Management, Infrastructure Management, High-Tech Engineering and Testing. Working closely with its clients, Sogeti enables them to leverage technological innovation and achieve maximum results. Sogeti brings together more than 20,000 professionals in 15 countries and is present in over 100 locations in Europe, the US and India. For more information please visit www.sogeti.com.

About VINT

It is an arduous undertaking to attempt to keep up with all developments in the it field. State-of-the-art it opportunities are often very remote from the workings of core business. Sources that provide a deeper understanding, a pragmatic approach, and potential uses for these developments are few and far between. VINT, the Sogeti Trend Lab, provides a meaningful interpretation of the connection between business processes and new developments in it. In every VINT publication, a balance is struck between factual description and the intended utilization. VINT uses this approach to inspire organizations to consider and use new technology.

Big Social

Participate in our Big Data discussion at www.sogeti.com/vint/bigdata/questions

SOGETI

VINT | Vision ● Inspiration ● Navigation ● Trends