



Analytics Trends 2014

Rapidly Evolving World of Analytics



Contents

Foreward	1
The rise of the Chief Analytics (and Big Data) Officer	2
Machine learning finds a big data niche	4
A picture is worth a thousand numbers	6
Data products run amok	8
Is the data warehouse dead?	10
Data discover platforms: The new R&D lab?	12
Analytics drives entertainment	14
The talent crunch that wasn't	16
Contacts	18

Foreward

Few areas of business today are changing faster than where analytics is being used and how it is being used. Turn your head for a second and boom, you are falling to the back of the pack. From big data and visualisation to predictive modeling and more, analytics represents a rapidly evolving world of technologies and tools that few have time to keep up with. This makes it challenging to think about which trends really matter and which will prove short lived, which are hype and which will deliver tangible, timely business value.

We need to think about trends, if only to understand where the market might be moving and how fast. Despite all the uncertainty in the field, business leaders still have to make decisions and choices about the future.

We have taken a fresh approach to analytics trends, focusing on developments that appear to be important while also pointing out areas of disagreement, even within our own organisation.

Whether a particular trend will affect you immediately is hard to say. Yet thinking about the trend, and what it might mean for your organisation if it is true, may be time well spent.





The rise of the Chief Analytics (and Big Data) Officer

A few years ago, there were no Chief Analytics Officers (CAOs), no Chief Data Officers, no Chief Science Officers and no heads of big data. Today, there are many in the tens, if not the hundreds; that in itself is evidence of a trend. The more important question though, is whether C-level analytics and big data positions can help organisations do more with their resources. We do not have a lot of data on this yet, but the answer seems likely to be yes.

It is possible, of course, to fold big data and analytics into the responsibilities of existing C-level executives (a CIO, a CFO, a COO, or even a CEO). And while it is great that the topic of analytics has the attention of such senior executives, will it really get enough attention with all the other things that business unit leaders have to worry about?

There are many other examples of organisations that have successfully folded analytics into other C-level jobs, showing how combining analytics with other functions can work quite well.

Despite these successes, there is a strong logic to having a dedicated role for analytics. If this resource is to achieve its potential, it should have a strong advocate and overseer of analytical and big data resources (people, technologies, data, and so forth). If you are really serious about analytics and you want to employ them in a variety of functions and units around your organisation, this sort of job and title seems to be just the ticket.

There is also the Chief Data Officer (CDO) role, which is increasingly common in large banks. In principle, it is a fine idea to combine the responsibility for data management and governance with analytics. In practice however, many CDO incumbents seem to spend a great majority of their time on data management and not so much on analytics. Also, many of them do not have strong analytics backgrounds, which may lead them to gravitate toward data management topics.

However, it makes little sense to object. Any organisation that creates a senior management role with analytics in it, regardless of the specific title, has likely done a productive thing. The perfect analytics title should not be the enemy of the good one. The mere fact that the organisation is focusing on doing more with its data, seeking to generate valuable insights and working to make those insights come alive in the organisation, is a step in the right direction.

It is a natural, if temporary, role

When e-commerce began to rise around the turn of this century, many organisations created senior e-commerce roles to advocate for and guide the new capabilities they were building. It is probably a safe bet that their positions sped up their companies' initiatives in this important domain, even if the roles were eventually absorbed back into IT or marketing. The same can be said of CAOs and Big Data Officers today. Eventually these capabilities may well become pervasive but we should have someone to lead the use of them today and for the foreseeable future. Organisational structures always should be flexible and when we see an important new capability arriving on the scene, we should create a role to manage it.

Another C-level role may be overkill

Analytics and big data may be important enough to deserve their own role, but creating it will not necessarily help organisations succeed. First, there are plenty of other CXO positions that care about analytics; CIOs, CFOs and CMOs, to name three. If they have passion for the topic, why can't they lead its application across the business? A CAO has to report somewhere anyway. Secondly, we have seen these arguments for new C-level roles such as Chief Security Officers. A few companies established them, but little evidence exists that the new role was any more successful than a CIO or CTO overseeing the area. Adding a new C-level title almost indicates that the topic is faddish. Whatever happened, for example, to Chief Knowledge Officers? Why can't the Chief Strategy Officer spearhead analytics, given that strategy should be at the core of any analytics roadmap? The last thing analytics needs is the same fleeting popularity.



Machine learning finds a big data niche

Machine learning is not a new idea. It has been around in theoretical form since the 1960s and in academic use since the 1970s and 1980s. Broadly speaking, machine learning is the ability of computers to learn from data. In analytics, it typically means the semi-automated development of predictive and prescriptive models that get better over time. The software learns how to better fit the data, separating meaningful signals from meaningless noise.

The challenge with machine learning in analytics is that it is a black box. Data scientists may understand how it works, but business people do not. In the past, managers were justifiably reluctant to make major decisions with no hypothesis or human explanation behind them. Indeed, many machine learning findings have simply been ignored by organisations. Today however, big data projects are often moving too quickly for traditional hypothesis driven analytics. In the area of digital marketing for example, companies that place digital adverts find they need to create thousands of new models a week, with the time to make a decision on where to place an advert measured in milliseconds. No set of human analysts could be that prolific with a traditional approach.

But that does not mean analysts and data scientists can just turn things over to black boxes. The leading machine learning environments still involve smart human overseers (people who specify the types of variables that can enter models, who adjust model parameters to get better fits, and who interpret the content of models for decision makers).

Machine learning can also increase the productivity of quantitative analysts. For example, one IT vendor is able to create 5 000 models a year to pinpoint the best focus for sales and marketing efforts, using only four analysts. Before machine learning, the organisation could manage only 150 models a year, with twice as many analysts.

The industrialised approach, sometimes called a model factory, is gaining steady adoption in certain sectors. And while we are still in the pre-Henry Ford stage as far as deep automation is concerned, the concept is likely to take off just as fast as conventional manufacturing did.

Automate this

This trend is spot on. It reflects a long developing technology that has reached critical mass. Furthermore, we are seeing considerable demand from client companies for people who are familiar with the approach.

They know that machine learning is not, and will likely never be, a completely automated way to develop models. We are also beginning to detect interest in new business models in this domain, such as machine learning as a service. Consulting organisations and cloud service providers are likely to be the early adopters.

Bring back the humans

This trend has its place, but its value and capabilities can be overblown. There is sizable garbage in, garbage out potential here. For example, just because someone knows how to operate the sausage maker does not mean tasty sausage will turn out. Companies still should rely on smart human analysts to hypothesise about relationships in the data and find models that support or overturn those hypotheses. It may be a slower process, but their companies may be much less likely to get into trouble. The reality of course is that this is not a black or white issue; there are shades of grey. Ongoing thought should go into deciding how humans can maintain control because they are ultimately accountable for the outcomes.



A picture is worth a thousand numbers

There is no question that **visualisation has become a critical capability for organisations of virtually every shape and size. Easy to use software makes complex data accessible and understandable for almost any business user. From discovery and visual exploration to pattern and relationship identification, today's visualisation tools easily affirm the adage that a picture is worth a thousand words, or in this case, numbers.**

This is especially true with big data, where visualisation may even be a necessary capability for driving insights. That is why visually oriented tools are rising in prominence for many big data applications. Users get to understand, explore, share and apply data efficiently and collaboratively often without the need for analytics professionals. And that is where the risk comes in. In their eagerness to dive into data, users may choose polished graphics over thorough data preparation, normalisation and rigorous analysis, glossing over important insights and analysis opportunities, potentially producing erroneous results.

Underlying the visualisation trend is a general movement toward management wanting greater involvement with analytics and data driven decision making. That is a positive and important trend in its own right and visualisation is a key enabler. Just

take care to not get caught up in all the sizzle. Visual analytics do not make sense for all decision situations, and some visuals simply do not add clarity to data.

Sometimes expressing results in plain old numbers makes sense where the glamour of visualisations creates mere noise. Do not overly patronise business people with visualisation as the new normal.

That said, visualisation has a legitimate place in almost any presentation of data. The benefits are real when humans need to make decisions about data:

Faster observations of trends and patterns. Often the essence of a story lies in larger patterns that occur at an aggregate level. Visualisation can help users see those patterns quickly.

Better retention. For some people, visualisation makes it easier to understand and remember data. Also, visual tools can present entire concepts in stories, using real world elements.

Embedded feeds and user engagement. Advanced tools allow for real time tweaking of data and visualisation based on live feeds. Content can be edited and explored by both analysts and audiences, providing custom views for end users.

Of course these benefits have always been possible from visual analytics, but easier to use technologies are making them more accessible, even for big data. We will be seeing more bar charts, line graphs and heat maps for the foreseeable future.

Picture this

Visualisation is real and valuable. Without it, analytics adoption would likely be going nowhere fast. But like any technology, there is a right and wrong way to use it. If your decision makers want to use data and should you increase their understanding of data before they act, there may be no better tool. Of course this does not imply you always have to visualise. Computers do not care about visual analytics, so if your decision is to be automated without human oversight or intervention, do not waste time with visuals.

Not so fast

The overuse of visualisation can get in the way of clarity. Sometimes that is the result of making a complex story too simple and sometimes it's the opposite, making simple facts seem too complex. We have all experienced that sense of dread when turning the pages in a document, in which each page presents information in a different way than the page before. Who has the time and patience to read this kind of cumbersome material?

More important is the risk of putting so much energy into a presentation that the actual analysis gets short. If the underlying data and assumptions are not valid, the analysis will not measure up, no matter how visually interesting it may prove to be. With analytics talent in such short supply, particularly analysts who can creatively develop visual displays, this risk is clearly present.



Data products run amok

One unexpected legacy of big data is the development of data products and service offerings that are based on data and analytics. The purest manifestation of the data economy; they emerged from the online industry over the past decade. Just as they drove the development of new technologies such as MapReduce and Hadoop, they will likely lead to a new set of approaches to analytics and data management when they take root in a variety of industries and businesses.

For online companies such as Google, almost every offering is a data product. Another online business with a strong focus on data products is LinkedIn. That company's data products group has developed such offerings for customers as People You May Know, Groups You May Like, Jobs You May Be Interested In, and InMaps; these offerings have led to considerable gains in customers and retention levels.

A logical extension of data products in the online industry is to develop them in software companies. For example, an organisation of data scientists was acquired with the expressed intention of developing data products from the financial data it has accumulated through its tax and business software.

The development of data products and services however is taking place even in industrial and manufacturing organisations. One focus has been on embedding sensors in products and then analysing the data they produce for conditional and predictive maintenance. Organisations are also, for example, placing sensors in industrial products such as gas turbines, jet engines and medical imaging devices, and providing service to those products on the basis of sensor data analysis. The primary goal is to monitor and discover patterns in faults, errors, outages and alerts, and then develop predictive maintenance approaches to mitigate potential problems before they occur.

It is too early to say that data products and services will out earn or outperform more traditional products and services but in some industries and for some companies, they have significant potential. Many companies now have the possibility of gathering data from operations, analysing it and making the data and analysis available to customers. This type of data exhaust can turbocharge your business or your competitor's. Much, much thought should go into the ramifications of selling your data.

A straightforward extension of big data analytics

It makes perfect sense to examine and pursue the possibility of data products from online data, sensors and other new sources. This trend was mentioned more than 20 years ago, when Stan Davis and Bill Davidson wrote the book 2020 Vision. They argued that information businesses could capture data and resell it to customers. In the big data era, virtually every business can become an information business. To make this happen, companies need to form data products teams that combine analytics, technology and customer expertise. Many companies are already selling their data assets and many more organisations will likely explore this in further detail over the next few years.

Don't take your eye off the internal decision ball

There may be some opportunities for data products and services outside of the online industry but it is not a significant growth opportunity. Instead, companies should maintain their focus on analytics for internal decisions. There are still many unexploited opportunities for that type of work and the pay off is much more probable and quick. Before undertaking development of data products, perhaps companies should do a better job of using analytics to market and sell their existing products to existing customers.



Is the data warehouse dead?

The rise of big data has had many implications, one of which is to call into question the future of the Enterprise Data Warehouse (EDW). Many companies are considering turning to Hadoop clusters, which cost a fraction per terabyte to store and process data as compared to traditional warehouses. Hadoop is even cheaper than the data appliances of the past decade. In addition, Hadoop platforms do not just store data, they can also perform substantial processing tasks and even some analytics.

Does this mean the end of the EDW? Well, the trend may not go that far. EDWs are still a very desirable way to store and process data for analytics. Plus, the structure of an EDW flows into numerous systems and functions of a company, both upstream and down. In addition, EDWs have the security, reliability, concurrent user support and overall manageability that large organisations need. They probably will not vanish anytime soon for production analytical applications. In fact, the rise of in-database analytics, in which analytical tasks such as scoring customers in a propensity model can be performed much more efficiently and has made EDWs even more popular for analytics.

However, alternatives to EDWs are increasing. For example, one major bank maintained a large EDW for production applications. However it also supported a series of data marts on smaller data appliances for applications that required somewhat less permanence and security. The bank had invested in a data discovery environment with a dedicated appliance for it. Furthermore, it had stood up a Hadoop cluster for unstructured data applications requiring little security and permanence. The bank is also investigating whether it should add graph databases for social network analysis and columnar databases for high speed analytics on numerical data.

Choice is a good thing of course, but with it comes complexity, confusion and cost. Organisations like this bank may need to develop well defined processes for deciding which data and applications go where, and under what circumstances they will need to be relocated to another platform. The value of such activities should be factored into the end-to-end roadmap of an organisations efforts and analytics journey.

With all this complexity, CIOs may yearn for the good old days when an EDW was the only alternative, but they will likely appreciate the lower prices per terabyte that their organisations may be paying.

RIP, EDW?

The main issue with this trend is that it does not go far enough. EDWs are a vanishing breed, although organisations may not turn over their installed bases quickly. You see the trend quite strongly in the online industry where companies still have large EDWs, but their data storage is now much larger in Hadoop clusters. As Hadoop-like products improve in function and maturity, we will likely see even more organisations adopting them. It is not surprising that you see EDW vendors putting their brands on Hadoop clusters and related technologies. They will need to do that, and do it well, in order to stay in business.

Long live the EDW

The big data societies may be very excited about Hadoop, but the EDW has served thousands of organisations very well and it is not likely to go away. It tends to be more expensive than Hadoop, but it has a much better fit with the data management and analytics needs of large, established organisations that can not bet their futures on unproven open source technologies. At some point we may see hybrid Hadoop/EDW environments, however in the short run, EDW is still king.



Data discover platforms: The new R&D lab?

Discovery has long been an essential aspect of research, development and product innovation in science oriented industries, however today it's no longer restricted to the chemistry or biology lab. Discovery is increasingly a function that involves data management and analytics as the primary focus. Leading organisations are adopting data discovery platforms; technology environments that make big data manipulation relatively easy and inexpensive.

In an effective discovery environment, more types of data can be considered, more variables and cases can be used in models and synthesised for new applications, and fast moving data can be analysed at speed. Once restricted to online and startup organisations, big data (and the products, services and large scale decisions it engenders) is now available to all organisations. Making the most of it requires a well structured discovery platform and process.

In a traditional analytics environment, discovery was less essential. Data exploration was slow, and few companies relied on analytics for decisions at scale. Online organisations had their own discovery

processes, but they relied on expensive and time consuming work by a few data scientists. In the current era, as analytics on all types of data becomes mainstream, understanding relationships in data has become an essential process of a critical business resource.

In the past, some organisations established analytical sandboxes, but discovery environments are different. First, the users interacting with these environments are not the analytics experts anymore; discovery environments are now the domain of the masses. Analytics aware end users are demanding access to these environments to interact with their data, look for patterns, or perform unique drill downs of the data that traditional business intelligence environments failed to deliver. Second, as these environments ingest more data, their architectures are becoming more complex, with Hadoop/ MapReduce datastores added to handle big data volumes, and specialised processing appliances being added to increase speed and performance. Today's data economy involves developing data products and data driven decisions at scale and in short order.

So far, few organisations have developed the types of stage gate data discovery processes that characterise many R&D organisations. But that is changing fast. If data is one of the most important off balance sheet resources that an organisation has, then determining how best to exploit it will likely become a permanent aspect of business.

Make way for discovery platforms

Production discovery environments are a central concern among many of our most sophisticated clients. It does not make sense to address an important domain for analytics without first exploring the data to unlock the real insights, assessing the distribution and quality of the data, testing possible relationships, and developing some trial models. This is not the private domain of the data scientist, but an inclusive environment for both the amateurs and the experts to explore and discover. Traditionally, these environments were called sandboxes; places with limited rules and limited oversight. As discovery becomes a mainstream way of answering ad hoc questions, these environments evolve from being nice to have playgrounds, to a mission critical setting. Availability, version control and release management (all the dominion of the IT groups) become the order of the day, and the restrictions go up.

Few organisations need them

Discovery platforms are often a good idea for online organisations with a lot of unstructured data, but they may not be needed in many traditional industries. We have been able to develop analytics without these tools. For some, the discovery movement may be a new way to sell big data appliances. Clients are not necessarily asking for them, and this trend is not taking off within the great majority of companies. Even the sandbox idea raised hackles among sceptical managers and this sounds like the sandbox on steroids. With CEOs and CFOs demanding to see return on investment, organisations need to find the clear line of sight to show them that necessary value.



Analytics drives entertainment

The practice of analytics has prospered in well structured business domains such as pricing and supply chain optimisation. It has also been adopted by marketing organisations, which have begun to see the virtues of data and statistics alongside creative content. More recently one of the last supporters of pure creativity, the entertainment industry, has begun to explore shifting from human judgment and experience to using analytics to determine which movies, television programmes, plays and books customers want to experience.

The track record in the entertainment industry of predicting consumer interest has been poor. Many local and global films fail to make money, television programmes are quickly cancelled, and among the hundreds of thousands of books published each year, few sell more than 100 copies. Ironically, an industry with such a low batting average is a prime candidate for more analytical decision making.

In the film domain, some companies are beginning to use analytical criteria to decide which movies to produce. For example, some film producers and studios employ an analytics algorithm to decide which movies to make.

Other entertainment companies have begun to focus internal analytics groups on discrete business units, such as books and plays. For example, one entertainment giant is working with its books and theater units to predict which books and plays will generate significant profits.

Of course, creativity is still needed in any entertainment business and human judgment will not be totally replaced by analytics. However, there is a need for entertainment executives who can augment their creative judgment with some well chosen numbers. The most successful entertainment of the future will likely be creative, inspired, delightful and analytically sound.

More data could drive entertainment analytics

It seems inescapable that the entertainment industry may become much more analytical over time, primarily because almost every aspect of the industry can be monitored and predicted now. With the advent of streaming video and other content over the Internet, a trend that is changing virtually every entertainment business, we now have the ability to know what is appealing to many consumers. If we produce content that does not engage the audience, it is probably because we did not do our research and analytics. The entertainment industry executive who does not understand and consume analytics may be a dinosaur in a very few years.

More data, but less creativity?

It is certainly true that there is more data now about the entertainment options that consumers enjoy. However analysing that data more extensively is unlikely to lead to more creative and entertaining media content. In the film industry, the movies that are designed on the basis of audience reactions to early screenings are not noticeably more creative or successful than those created by an inspired director. Too much focus on entertainment analytics will likely lead to a lowest common denominator effect involving less creativity rather than more. Let us not forget that many of the world's greatest artists died as paupers. It was not until decades or centuries after their deaths that the brilliance of their work became widely appreciated. While good might be an immediate best seller, great might not.



The talent crunch that wasn't

Conventional wisdom says organisations are facing a large supply gap of data analytics talent, with significant shortages at all levels. Entry level positions are challenging to fill, and there is also a major drought at the most senior levels, though few such leaders are needed overall. Professionals who can deliver data backed insights that create business value, not just number crunchers, are especially hard to find.

Is the conventional wisdom correct, or are there other forces at play? Large organisations in traditional industries may be much less concerned about hiring lots of data scientists than are online organisations and startups. While they do need trained analysts, many are not so sure those analysts need to be hired all at once and turned loose without a cohesive strategy.

What's more, there is massive confusion about what a data scientist actually is. For some, a person who can manage spreadsheets and do basic reporting might qualify, at least in his or her own mind. For others, the data scientist definition speaks to a rare blend of statistical sophistication, data management skills and business acumen. To add to the classification confusion, an influx of new vendors with low cost visualisation applications has made data, reporting and business intelligence understandable and accessible to almost anyone.

In response to growing demand, colleges and universities are adding programs (degrees, majors and courses) in analytics and data science, which will broaden the mix of analytics related pedigrees. We could soon be drowning in individuals with data scientist titles, further blurring distinctions about the skills they bring to the party and how they integrate with the broader team.

Part statistician, part predictive analyst, part business analyst and part trusted advisor, these professionals have to:

- Understand business
- Be able to analyse social and unstructured data
- Design and test predictive models
- Know maths and statistics
- Cross the lines between social sciences, business and mathematics
- Tell stories using data

That is a tall order, maybe even a mythical creature. It is quite possible that data scientist as currently defined requires more attributes than most individuals should be expected to have. Some companies are resolving the gap by creating teams of diverse skill sets. They are mixing and matching professionals to deliver a balanced response to business analytics questions, creating high performing teams that can deliver business value.

When you are recruiting analytical people, be clear what your needs are. Just saying you need a data scientist is like advertising a slot for a smart person who is good with numbers.

The talent crunch is real

Companies are struggling at almost every level of analytics recruiting today. The team approach makes sense, but it assumes there are enough people to hire in a wide range of analytics related roles.

That is not the case. Even as universities create many more analysts, it will be years before they enter the labor force and become effective. Great data scientists also rely on their past experience and getting these new graduates to that point will take some time.

Don't buy the hype

Some of the current talent crunch is a function of hoarding, not real demand. In response to predictions about impending shortages of qualified analysts, organisations scrambled to recruit talent beyond what they actually needed. This led to experienced people being asked to carry out activities like straightforward reporting that could have been done with lower level talent. It also led to lower level talent doing busy work, for example cleaning data, which is better done by machines.

In addition, startups are hiring like crazy. Millions are being invested in identifying the high value use cases, building the magical model to predict the future and creating the subscription service to ingest data and extract insights they can sell.

This will likely all shake out over the next several years, with market forces and employees themselves driving the outcomes. Third party solutions will cover many of the most valuable use cases and internal teams will be relegated to more mundane work, if not cast adrift.

Contacts



Dave Kennedy

Managing Director, Risk Advisory Africa

Deloitte South Africa

Tel: +27 11 806 5340

Email: dkennedy@deloitte.co.za



Derek Schraader

Africa Leader: Data Analytics

Director, Risk Advisory (Johannesburg and Cape Town)

Deloitte South Africa

Tel: +27 79 499 9046

Email: dschraader@deloitte.co.za



Werner Swanepoel

Director, Risk Advisory (Johannesburg)

Deloitte South Africa

Tel: +27 11 209 6664

Email: wswanepoel@deloitte.co.za



Ashleigh Theophanides

Director, Consulting (Johannesburg)

Deloitte South Africa

Tel: +27 11 806 5340

Email: atheophanides@deloitte.co.za



FC de Vos

Senior Manager, Consulting (Johannesburg)

Deloitte South Africa

Tel: +27 11 209 8263

Email: fcdevos@deloitte.co.za

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited (DTTL), a UK private company limited by guarantee, and its network of member firms, each of which is a legally separate and independent entity. Please see www.deloitte.com/about for a detailed description of the legal structure of Deloitte Touche Tohmatsu Limited and its member firms.

Deloitte provides audit, tax, consulting and financial advisory services to public and private clients spanning multiple industries. With a globally connected network of member firms in more than 150 countries, Deloitte brings world-class capabilities and high-quality service to clients, delivering the insights they need to address their most complex business challenges. Deloitte has in the region of 200 000 professionals, all committed to becoming the standard of excellence.

This communication contains general information only, and none of Deloitte Touche Tohmatsu Limited, its member firms, or their related entities (collectively, the "Deloitte Network") is, by means of this publication, rendering professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser. No entity in the Deloitte Network shall be responsible for any loss whatsoever sustained by any person who relies on this communication.

© 2014 Deloitte & Touche. All rights reserved. Member of Deloitte Touche Tohmatsu Limited

Designed and produced by Creative Solutions at Deloitte, Johannesburg. (807626/sue)