"Big data": a buzz word or something that makes sense for my company?

*From Data to Big Data*

Do you have any idea about the volume of data that we deal with – in a way or another – every day?

Altogether, we produce more information every minute than all the data which has been produced from the beginning of time until 2000... every minute!

Indeed every 60 seconds, we produce more than 100,000 tweets, we update more than 700,000 status on Facebook, we share more than 11,000,000 instant messages, we introduce more than 700,000 searches on Google, we send more than 170,000,000 emails, we upload more than 300 hours of video on YouTube, etc., etc.

And those figures have been collected some months ago, so it should be more today.

It seems incredible but that’s our present world, dealing with more and more information every day. And if we consider the development of IoT (Internet of Things), the quantity of data that will be produced by machines will exponentially increase the quantity yet produced.

In such a world, we can definitively talk about big data. In the last 15 years, IT has transformed a world of data to a world of big (enormous, gigantic...) data.
"Big data": a buzz word or something that makes sense for my company?

It seems far from my reality. Indeed, in my day to day life, I receive, read, analyse, keep for later and even delete lots of electronic documents (emails, presentations, word documents, spreadsheets, news ...) and physical documents (letters, invoices, contracts, orders, magazines, newspapers ...), all about various topics: commercial, finance, HR ... which is nothing compare to this ocean of big data.

My 100 emails, 5 spreadsheets, 3 presentations, 15 Word documents, 8 letters... that I receive in every average working day are no big data! No relevance! Really?

*From Big Data to Enterprise Data: how to use this?*

In my company, since the last 12 months, with 1.000 employees:

- We received 15,000 invoices and we sent 20,000;
- We wrote 1,500 commercial proposals;
- We produced 19,000 payslips;
- 1670 Curriculum Vitae have been received by our HR Dpt;
- ...

Considering everything, our IT Dpt. told me that we store about 12,000,000 documents every year.

That’s the reason why I talk about “Enterprise Data”. We cannot consider the words “Big Data” which apply to a really different world but even for a small company with 1.000 employees the data that is flowing in and out represents an important amount of information. And even more important,
"Big data": a buzz word or something that makes sense for my company?

in such data flows, how to identify, spare and retrieve the relevant information when necessary?

Then the questions are “how can I use the information that are produced and received in my company” and “what for”?

The “what for” is easy to answer; because, as it’s the case in the world of Big Data, the more you collect the relevant information out of the mass of data that is circulating in your company the better or faster you:

- Optimise and improve your business processes
- Predict customers behaviors
- Bring new services to your clients
- Make decision
- Control and monitor operations and business
- …

Based on a Data Analytics approach, the possible applications of Enterprise Data are endless.

We can present several types of big data that really can help organisations:

- **Prescriptive** – it reveals what actions should be taken.
- **Predictive** – An analysis of likely scenarios of what might happen.
- **Diagnostic** – It shows the past to determine what happened and why.
- **Descriptive** – What is happening now based on incoming data based on real-time dashboard reports.

And regarding the “how”… There are technologies out there today. The very ones used in the world of Big Data can be applied to your own world of
"Big data": a buzz word or something that makes sense for my company?

Enterprise Data: OCR, data entry, structure detection, automatic segmentation, data extraction, text mining, cleaning, stream analytics...
The main point is to define the goals and issues in order to know what to collect and analyse among all the existing information.

To illustrate let’s consider a concrete example: in which a company receives a lot of invoices every year. Invoices are really common documents with poor information, just used to control and to pay a product or service which has been ordered accordingly.

Provide a service of data extraction to collect the relevant information in order to fulfill the ERP / accounting solution is of course already interesting to do in order to reduce some operational costs.

But is that all? When you have lots of invoices out of which you extract information beyond accounting or workflow feeding data, meaning data relevant to other issues, then you can compare values from different providers, families, products or even between internal units of your company, link to other information areas like quality or project management and better define / complete your purchase policy, manage a pricing analysis and finally reduce your costs, improve quality and service.

That’s what we call Data Analytics meaning to collect, compare and analyse data.
"Big data": a buzz word or something that makes sense for my company?

My example is an easy one because I considered only 1 specific document: invoices. But we can find lots of other interesting applications on different types of documents flowing through different units of your company, as said the applications are endless and depend on your issues (increase the productivity, increase the quality, decrease the risks...)

Let's consider another example, some companies use predictive analytics for the entire sales process, analysing lead source, number of client meetings, types of communications, social media, documents, CRM data... This can be used to support sales, marketing, or for other types of complex forecasts.

The main key is about data extraction because data extraction for structured information (20%, i.e. databases) and unstructured information (80%, i.e.; emails, word documents...) from both physical and digital documents allows
"Big data": a buzz word or something that makes sense for my company?

organisation to collect the relevant data you need to take benefits out of your flowing information. It means:

- Express your goals
- Inventory the information you own and the one you can have access to
- Define the relevant information you need
- Set in place technologies and/or (external) services which will extract the relevant data (knowing that you need of course consistent data with the highest quality rate)
- Run algorithms which will compare, cross-check… data
- Analyse the results and put in place the efficiencies which have been highlighted
- And then? Express your new goals...

Of course, advisory companies and service providers can help you to point and bring the quick wins. And yes, if you deal with a certain amount of information, you have quick wins!

A last tip: the past runs to a certain extent our future, when you analyse your data from yesterday and today, you are able to predict, or at least to better forecast; it must be a real commitment for your company in the near future.

David Gray

www.numeneurope.com,
david.gray@numeneurope.com
+352 621 254 169