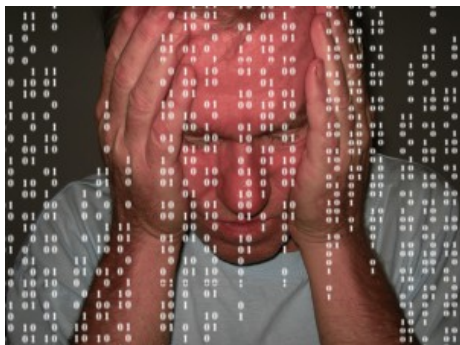


## Data Democratization and Data Ownership - Get it right!

The term "Big Data" started becoming "famous" some time around 2010, although it seemed to be coined [a few years earlier](#)... The real breakout took place back in 2012 [and never stopped ever since](#)... And here I am, at the very beginning of 2016 writing about something as obvious and as intrinsic to the Big Data revolution as **Data Democratization**.

But there are still so many companies that just don't get it... And they are so many opportunities just being wasted!

### What is Democratization of Data?



In a nutshell, it is the process and practice of expanding business information and the tools to analyze it out to a much broader audience than traditionally has had access.

Traditionally, big companies had gatekeepers of business intelligence and analytical tools... sometimes because of the necessity to control information, sometimes because of the incontestable advantage of sitting on top of all company data assets and deciding who sees what.

In addition, the toolset available for analyzing and consuming information was far away from easy to use, leaving the access to information in hands of teams specialized in *getting requirements and building reports*.

As your experience might have proven to you, this approach does not work for 2 main reasons:

- *it hardly scales* and for obvious reasons, the existing capacities usually tend to focus on satisfying the needs of only top managers, leaving the majority of employees left without the data they need for the strategic and more often operative decision making.
- as people tend to help themselves (over workarounds, trading off requirements, etc)... different departments end up with *multiple versions of the truth*.

### But who is the legitimate owner of the data?

David Loshin, in his book [Enterprise Knowledge Management: The Data Quality Approach](#) (Morgan Kaufmann, 2001), described what he called the *Paradigm of Ownership* not with the intent of establishing who the legitimate data owner should be, but to accent the complexity of ownership issues and to identify the list of parties laying a potential claim to data:

- **Enterprise:** all data that enters the enterprise or is created within the enterprise is completely owned by the enterprise.
- **Creator:** is basically the party entering data into a form of data base, logs, etc. Here we can consider for

example the responsible entities for sensors, IT systems generating logs, GPS, etc.

- **Compiler:** the entity that selects and compiles information from different information sources, such as a company Data Warehouse or a Data Lake.
- **Funder:** the user that commissions the data creation claims ownership... This is also the party buying external data to enrich own existing data sources.
- **Decoder:** in environments where information is "locked" inside particular encoded formats, the party that can unlock the information becomes an owner of that information (e.g.: DNA, information extracted from cryptic application logs, etc)
- **Packager:** the party that collects information for a particular use and adds value through formatting the information for a particular market or set of consumers.. Data curators and reporting services play usually this role.
- **Consumer:** data per se is only as valuable as the use companies make out of it. This "use" takes place in business departments when data is turned into insights to better inform business decisions. Those who end up unlocking the real value of the data might see themselves as "owners" -without them, data would remain useless-.
- **Reader as owner** - the value of any data that can be read is subsumed by the reader and, therefore, the reader gains value through adding that information to an information repository... It's in some way similar to the data Consumer.
- **Subject as owner** - the subject of the data claims ownership of that data, mostly in reaction to another party claiming ownership of the same data
- **Purchaser/Licensor as Owner** – the individual or organization that buys or licenses data may stake a claim to ownership... Funders are also related to this collective.

So who is the legitimate owner? all of them? We don't have to and actually we can't answer this question here... but that's not the point either... what shall come across this *Paradigm of Ownership* is the fact that data democratization is going to make this discussion obsolete... Each and every aforementioned party has a very important role across the data life-cycle but none of them is justified unless the data is turned into action... So it is crucial to [make the data available to those in a position of exploiting the value in it.](#)

## The show stoppers

### 1) Internal resistance and willingness to open the data



Let me try my best to come up with a *sylllogism* "Aristotle style":

Knowledge is power  
You need data to infer knowledge  
Ergo Data is power

Now if you ask departments that have traditionally been in control of the data to grant access for everybody to the company data assets, you might face some reticence... They might not be that interested in opening up, as you can imagine, because:

- The *exclusivity* in terms of which department generates "knowledge" disappears... Moreover, the kind of knowledge a business unit is able to infer from the data is going to certainly be much more relevant for the business decision making due to the closeness to the point of action.
- The *quality* of their work is going to get exposed: data curation, standardized ETL processes, inconsistencies handling, security, data protection compliance practices, the presence of a good [Data Quality Manager](#), etc... Moving away from a blackbox approach might also reveal spots where the lack of quality manifests.

## 2) Security risks allowing access to potentially sensitive data

The more people get access to information, the higher the risk of leakage, misuse, etc... That seems clear... but this is not an excuse.

The awareness of how important is to preserve some guidelines while dealing with sensitive information is there (especially after so many information leakage scandals, episodes of customer data being stolen, celebrities' pictures, NSA, etc).

Companies needs to provide the proper training in security and data protection matters, the proper toolset to make transparent the compliance to the latest regulations and restrict information that's particularly sensitive to those who really need it.

But BE CAREFUL with the last statement... This "access restriction" card has been played *ad nauseam* by those concerned by the previous point... The question "*what makes a department eligible and what makes others not eligible to have access to particular data*" always help...

## 3) The risk of misinterpretation

If you have a look at your own company, you are most probably agree with me: not every employee is able to translate that information into useful practices... But on the other hand, almost every prospective employee you can speak to lately is going to name "analytical" as one of their strengths.

The problem is less the lack of skills to look at the data and more the interest of developing actionable plans around it (see [here](#)).

But in any case, companies should invest in the proper tools to remove complexity in dealing with data, in the right programs to get everybody up-to-speed to make data-driven happen and to increase the business intelligence capabilities there, where the business decisions are taken (for example by fostering a *pervasive BI strategy*).

## It is already happening and it is going to continue happening



Of course the *data democratization* is not something coming out of the blue... Actually it has been and it is still one of the topics many specialized sites has identified as one of the top big data trends (for both 2015 and the recently started 2016)

### Looking back to 2015

[...]

#### [No Ownership In Just One Department](#)

Data will become a commodity that is not just kept in one department alone and used purely by senior company leaders. 2015 is likely to see a democratization of data throughout the organization, meaning that more departments will become adept at using the insight that it can bring.

Rather than working towards a central strategy that is created by senior management, day-to-day activities will be based on data and the insights created from it.

### Looking forward to 2016

[...]

#### [Data In The Hands Of The Masses](#)

Data is no longer just something being discussed in boardrooms and laboratories at the highest levels [...].

## What if you don't embrace the data democratization paradigm?

Not democratizing your data access across your company is no longer a viable choice, because:

- The traditional approach *does not scale!*
- *Your competitors* are certainly making sure everybody got their data they need at their fingertips... if you don't your competitiveness is going to be undermined.
- A *data driven strategy* requires the maximum number of employees to infer actions out of the data to steer the business.
- *Pervasive Business Intelligence strategies*, intended to bring analytical and data science skills to the places where business steering decisions are made, require access to all relevant data sources. Data democratization is a mandatory step.
- What makes a difference when you extract insights from your data is the *contextual information* and the *business knowledge* your staff have been acquiring for years... and these two ingredients aren't known to be available in centralized intelligence departments.

## Takes away

Needless to say that the point I've been trying to make all across this post is *the need to democratized your data*... But in a nutshell, you might want to take these sub-points with you:

- If the data assets in your company is locked in the hands of gate-keeping departments and you don't have it where it is needed, you need to democratize your data!
- Data Ownership is a complex topic where several parties play different roles, but the discussion is futile if data is not employed to add business value.
- The main show stopper is the internal resistance to transition from the status quo, where data has traditionally been in the hands of gate-keepers.
- Companies need to invest in the proper training to develop data analytical skills and implement pervasive BI strategies
- To remain competitive is it incontestable to embrace the data democratization in your company... otherwise it is a matter of time to be outperformed by your competitors.

Also the sooner you start, the better! Help your company to break silos and start putting data there where it is needed.