

CAROLINE GOULARD

CEO & Co-Founder, Dataveyes

Hello,

Thank you for having given me the opportunity to speak to you today. My name is Caroline and I am a co-founder of Dataveyes. Dataveyes is a company specialising in interactions between humans and data.

At Dataveyes, we design and build systems which allow humans to better understand data and work better, get better information and simply live better with it.

And the fact that today there are people like my team and me who specialise in this question of the relationship between humans and data shows that this relationship is not necessarily straightforward.

For the next few minutes, I'll give you a brief overview of these relationships between humans and data, related to this afternoon's issue.

Data is a subject which people talk about more and more, simply because of its massification in the past few years. It has become the preferred format for encapsulating the large quantities of information that circulate around us.

This massification of data has been seen as an opportunity by many people, including me.

- In the first place, an opportunity for businesses. The opportunity to improve their ability to make the right decisions, an opportunity to enrich their business models, to better meet the needs of their customers, but also an opportunity to modernise their processes and increase productivity
 - *e.g. one can cite here the example of Procter and Gamble, which has equipped 50,000 employees with a sales data visualisation system to enable them to make better-informed decisions*
- It is also an opportunity for public organisations. The opportunity to build better relations between citizens and politicians, thanks to the opening up of public data and the open data movement able to nourish everyday innovation, embody a momentum towards transparency and foster co-construction of public policies with citizens.
 - *Here we can cite as an example, New York City which has opened up significant data sets and created the conditions for innovative initiatives to emerge*
- It is also an opportunity for cities. The opportunity to become a smart-city, a city that meets the challenges of human development, which responds to tensions over natural resources, through optimised management of its networks through data analysis.
 - *e.g. the city of Barcelona has been ranked among the smartest cities in the world thanks to the implementation of advanced technologies that through the use of sensors better manage road traffic and urban lighting.*
- Finally, it is an opportunity for citizens, for the everyday life of citizens who may be able to have a more comfortable life thanks to a series of smart devices, assistant robots and tools that are going to simplify everyday life because they will be able to deduce a lot of things by themselves thanks to data.
 - *e.g. thermostats and smart meters which help householders to consume their electricity more efficiently and to make energy savings, thanks to data collected about the house.*
- It is also in a very general way, an opportunity to acquire greater knowledge of the world, and of oneself, because data can inform us of things that up to now could not be observed.

A world of opportunities therefore, which should be ours for the taking.

However, what we feel today, what we see in the media and what our fellow citizens say, is not this new world of data, full of promises, but are much more pessimistic observations on how data is currently used:

The recent events in America have given us several signs of this:

- A 1st symptom is that of the unexpected result of the American presidential election which has largely discredited opinion polls, the methods of predictive analyses and our ability to forecast thanks to data.
 - Indeed, there is a real problem if instead of helping us to better understand the world, data lead us away from the reality on the ground.
- A second symptom: Following this election, Facebook's algorithm has been widely criticised. It has been accused of fostering the spread of false information and thereby encouraging populism.
 - Indeed, there is a phenomenon that has been called the "filter bubble". This is where content proposed to web users is increasingly more finely personalised based on the personal data collected about these web users when they view information. These filter bubbles prevent web users from reading ideas other than their own by prioritising content intended to please them.
 - Here again, it would appear that the use made of data is not promoting better knowledge of the world, but on the contrary, being used to provide a narrower view.
- 3rd sign: What this election has also highlighted is our difficulty in making society, it is a finding that is also found in Europe and in France. Behind all of this are probably lots of sociological and philosophical causes, which do not really lie in my area of competence. But however a small link with this data massification phenomenon is to be seen:
 - Our digital habits accustom us to extreme personalisation, to targeting that is increasingly focused on the individual, and they also result in our increasing disintermediation.
 - Large communities no longer have much meaning. What meaning does a social class or even a nation have when statistics can be established at the scale of an individual? When one is able to quantify and observe an infinite diversity of human habits and reasons?
- Beyond the American election, there are many other revealing signs: as for example, the waves of protest of Facebook users when it updates the algorithm for presenting news in the users' feed, or where the same thing happens with Google and its ranking algorithm.
- These large digital giants use and monetise personal data, drawn from web users' behaviour, without necessarily providing the transparency, ethics and explanations needed for the end users not to regard them as intrusive and problematic processes with respect to their privacy and their right to be forgotten.
- Finally, ethical questions are being raised today, and increasingly strongly.
 - For example with the advent of self-driving cars: when an algorithm will have to judge what driving action to take in the event of an inevitable accident, and which lives to save as a priority. What will become of our moral beliefs, of our free will in these cases?

That is why there is a problem today: we are in the process of building a world that we understand less and less. It is creating a black hole between data and knowledge: our ability to implement smart technologies exceeds by far our ability to understand how the algorithms that feed them work.

And nobody wants to live surrounded by robots and algorithms that behave like black boxes. This prospect of a world in which we no longer understand the information that is exchanged in it is worrying.

That is why it is high time to react, and to change our approach to data, and to the systems that use large masses of data.

We need to change approach in at least three ways:

- First of all, there must be a change of approach regarding the question of meaning: we need to dedicate resources to giving meaning to data. It is a question of language: the language of data is often the language of code or mathematics. We need to find something else, something more heuristic. A language that speaks to humans. We need to establish better data-literacy, i.e. a better ability to speak and understand the language of data.
 - For example, data visualisation: making the information contained in the data understandable, making complicated phenomena understandable using other ways of interpreting information apart from pure logical or verbal rationality.
e.g. when we create a data visualisation platform that enables the general public in France to understand the assessment data for the quality of care in hospitals, which previously was very technocratic and complicated data.
- Then there must be a change in approach as regards the issue of uses: ensuring that new uses of the data are designed to increase the human dimension, to improve people's ability to be aware of the world, and to take informed decisions rather than build systems that take decisions for them
 - *This is the case for example when we produce maps that reflect in real time the state of a metro network, and the options that the city has to offer around it, to assist travellers with their journeys.*
 - This is true at the level of ordinary citizens but it is also true for businesses. The question of its use today is what explains why many companies are equipped with technologies to enhance their data without seeing effects on their productivity. This is a paradox that was already observed in companies at the end of the 1980s by Robert Solow under the name of productivity paradox, based on the finding that the computer technology was everywhere in businesses except in the productivity statistics. Because until it becomes part of everyday practice, investments in Big Data will not have any effect. What counts for us, for example, when we work with large groups, is to create tools that will enable employees to really harness this new source of information.
- Finally, there must be a change in approach with regard to governance and data ethics: this probably means overseeing and producing standards and this can only be done at the international level. And it is on this point, that I am turning to you, because you, in this room, are much better placed than me to make that happen.