

# CARLOS GUARDIA RIVAS

## IBM SPAIN

EXECUTIVE IT SPECIALIST - IBM SOFTWARE GROUP

---

### WHAT IS THE ROLE OF DATA WITHIN YOUR ACTIVITY AT IBM?

IBM is closely related to data and is a big data producer. For example, a few years ago IBM bought The Weather Company. This company generates climate data, compiles it and sells it to, especially, insurance companies and also to other private companies that make use of this kind of data. Even if IBM uses external data repositories, it is mainly a data producer.

As for the tools, at IBM we have a very complete portfolio of analytics tools, ranging from importing data from diverse sources, data integration in diverse formats, data cleaning, replication to other systems if needed... Our most famous analytics tool is Watson, which is basically a set of tools that allows you to do any kind of projection on the data: from visualization to prediction. In conclusion, we are a company that offers solutions in the main data-related areas. This is the field in which IBM is heavily investing.

### DOES IBM USE OPEN DATA FROM PUBLIC SOURCES?

We mainly use our own IBM-produced data and tools. We have a big variety of databases: structured, non-structured, analysis-oriented...

### CONSIDERING THAT MOST OF THE DATA IS GENERATED BY IBM, WHAT KIND OF STANDARDS DO YOU APPLY?

Our own customers demand us to adjust to specific standards when they use our tools. Therefore, we comply with all the typical standards in data management. If we didn't do so, we wouldn't sell anything! **We are always keeping an eye on the market's demand for specific standards.** For example, we have a petition system for our clients to request the application of a specific standard to a product.

### TALKING ABOUT THE CUSTOMERS' DEMANDS, WHAT ARE THE MOST RECURRING PROBLEMS THAT YOU ENCOUNTER?

One of the main concerns for our clients is the systems' scalability. Customers are continually developing new applications. And, in parallel, the volume of data is continually growing. We estimate that the total volume of data grows between 15 and 25% annually. And this growth in data leads to growing requirements for the systems, both from a storage perspective and regarding the concurrent transactions' capacity of the hardware. We are focused on offering global systems, of software and hardware, that are not affected by growth

restrictions: if your data grows, or the number of users and applications that make use of this data grows, you should be able to keep the same performance. For example, in the case of fintech companies, banks are facing more than a thousand connections per second coming from mobile phones... Scalability is a pressing issue.

Another concern has to do with the continued availability, because systems are becoming increasingly mission-critical. A couple of minutes-long blackout can mean huge losses. In order to face this challenge, the system needs to be designed in a way in which one part of the system can automatically take over the tasks performed by the part that is not working properly. High availability is absolutely important.

Moreover, data authenticity and security are also significant concerns. Illegal access to their data is a big concern for our clients and IBM is working to keep it secure.

#### AND HOW DO THESE CONCERNS LINKED TO AUTHENTICITY AND SECURITY RELATE TO THE TREND TOWARDS OPEN DATA?

It would be important to establish a way in which open data could be somehow validated by an entity. Data should be stored in controlled, validated repositories. Data that is not certified could lead to problems that would prevent this data from being useful. Maybe organizations like RDA could take up this role, proposing standards for open data repositories, for example. Public institutions could be in charge of these repositories, to keep data globally available. Standards should be global, that would be ideal.

#### IN THIS SCENARIO, DO YOU THINK THAT A GLOBAL OPEN DATA FORUM COULD BE POSSIBLE?

The challenge I could identify in such a forum would be to keep the repositories secure, certified, to prevent the illegal use of open data. I am thinking, for example, in health data, which is necessarily related to personal data.

I do believe that the idea of an open data forum, and an open data market, is possible. However, it is indispensable to establish how to profit from this data. Research data could be free, but at the same time very controlled. Also, it does make sense that companies pay something for the data they use to do business and make profit.

#### IN WHICH WAYS DO YOU THINK THE COLLABORATION BETWEEN THE PUBLIC SECTOR AND ACADEMIA, AND THE PRIVATE SECTOR COULD BE FOSTERED?

Both private companies and public organizations are data consumers, and they are both subject to European regulation. The finance and banking sector is an example. It could be useful to have opportunities where representatives of both sectors could get together and discuss the impact of the existent regulation in their fields, for

instance. RDA could organize informative sessions, talks or workshops to discuss these issues, and also receive feedback from them regarding the clients and citizens requirements and how the regulation has an impact on them.