



Editorial

PhD. Agapito Ledezma

Universidad Carlos III de Madrid, Spain

ResearcherID: K-3929-2014

<https://orcid.org/0000-0002-0041-6829>

<https://doi.org/10.5377/torreon.v9i26.10252>

The need to explore and exploit data

“Without data, you are just another person with an opinion” - W. Edwards Deming -

A little more than a quarter of a century ago, when I was still studying engineering and the Internet was bursting into the academic world in the region, I used to include in most of my academic work the phrase that was so much in vogue, then and now, that comes to say something like, *“The one who has the data has the power”*. At that time, in addition to the fact that the phrase was very well placed at the beginning of my documents, one could see that the so-called new technologies were going to revolutionize our world and that data would play a fundamental role in that revolution. Today, in the 21st century, immersed in the Fourth Industrial Revolution, in the era of the Internet of Things, of social networks, and of Artificial Intelligence, data is invaluable for those who generate it, but even more so for those who own and use it.

Nowadays, every time we make a bank transaction, surf the Internet or give a “like” to a photo or a comment in social networks, we are generating data that, in one way or another, someone is using or could use. Therefore, creating legal frameworks to regulate the use of our data is a major issue for any State today. The data we generate allows companies to suggest what to buy, what series or movie to watch or even remind you that you have to go to school to pick up your child.

According to a recent study by the technology consulting firm IDC¹, the Global Datasphere, or what is the same, all the data created, captured, or replicated at a global level in one year, is expected to reach 175 zettabytes (ZB) in the year 2025. We must take into account that one zettabyte is equivalent to one trillion gigabytes. To make it easier to get an idea of what that amount of data represents, if we were able to store the entire Global Datasphere in the year 2025, using DVD optical discs, we would have a stack of single-layer Blue-ray discs that would go around the Earth 222 times.

Linked to the strong increase in the production, storage, and processing of data that we live today, we find initiatives at a global level oriented to the promotion of the so-called Open Data. Open Data are those data “*that can be used, reused and redistributed freely by any person, and that is subject, at most, to the requirement of attribution and sharing in the same way they appear*”². The philosophy behind Open Data is to promote, among other things, citizen participation, innovation, transparency, self-empowerment, and scientific research. Today, there is a growing number of organizations, both public and private, that, under the philosophy of open data, make available to the general public data of interest of many different kinds. State institutions and agencies, universities, companies, and other organizations collect, process, and store large amounts of data related not only to people but to all kinds of activities including natural and social phenomena. Hence, we can find open data related to industry, environment, economy, sports, tourism, or health.

The current technological trend, in most first world countries, is the use of data-driven models to support decision making. This implies not only generating and managing the data on time but also ensuring the quality of the data. Data-driven models are what we in computer science call GIGO, which stands for “*Garbage In Garbage Out*”. In other words, if the data is not good, neither are the models. That is why data quality becomes a key factor in our day and age.

The current global health crisis, caused by the pandemic associated with the COVID-19 virus, is a clear example of the need for, on the one hand, public policies aimed at the management and exploitation of data, as well as the promotion of Open Data. The pandemic has highlighted the shortcomings that exist in many countries in terms of the capacity to generate, manage, and analyze data to address a specific issue. At this time, there is a pressing need to convert data into actionable actions that will lead us to overcome in the short, medium, and long term the effects caused by the current pandemic. Countries such as South Korea, which through its center for disease prevention and control, has made its COVID-19-related information available to the public, are examples to follow if we want to take advantage of the full potential of the data to address a host of problems affecting the countries of the region.

1. <https://www.seagate.com/es/es/our-story/data-age-2025/>

2. <https://opendatahandbook.org/>

Returning to the phrase “*The one who has the data has the power*”, I emphasize that you must not only have the data, you must know what to do with it. We have to be able to turn data into information, information into knowledge, and that knowledge into action. Public policies that bet on Open Data, the promotion of scientific research, and academic programs in the area of Data Science and related areas are today an imperative for the region. Furthermore, these actions must be accompanied by legislation that, on the one hand, guarantees the privacy and protection of data of all persons and, on the other hand, takes into account the ethical implications that could be derived from the massive use of data and new technologies.