



Editorial

PhD. Agapito Ledezma

Universidad Carlos III de Madrid, España

ResearcherID: K-3929-2014

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

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

La necesidad de explorar y explotar los datos

“Sin datos, sólo eres otra persona con una opinión” - W. Edwards Deming -

Hace poco más de cuarto de siglo, cuando aún estaba cursando mis estudios de ingeniería y el Internet estaba irrumpiendo en el ámbito académico en la región, solía incluir en la mayoría de mis trabajos académicos la frase tan socorrida, en aquel entonces y ahora, que viene a decir algo así como que *“quien tiene los datos, tiene el poder”*. En aquel momento, además de que la frase quedara muy bien al inicio de mis documentos, se podía vislumbrar que las llamadas nuevas tecnologías iban a revolucionar nuestro mundo y que los datos jugarían un papel fundamental en dicha revolución. Hoy en día, en pleno Siglo XXI, inmersos en la Cuarta Revolución Industrial, en la era del Internet de las Cosas, de las redes sociales y de la Inteligencia Artificial, los datos poseen un valor incalculable para el que los genera, pero más aún para quién los posee y los aprovecha.

En la actualidad, cada vez que realizamos una transacción bancaria, navegamos por Internet o le damos un *“like”* a una foto o un comentario en las redes sociales, estamos generando datos que, de una u otra forma, alguien está utilizando o podría utilizar. De allí que, crear marcos legales que regulen el uso de nuestros datos sea un tema primordial para cualquier Estado hoy en día. Los datos que generamos permiten que las empresas puedan sugerirte qué comprar, qué serie o película ver o incluso recordarte que tienes que ir al colegio a buscar a tu hijo.

Según un estudio reciente de la consultora tecnológica IDC¹, se prevé que la Datasfera Global, o lo que es lo mismo, todos los datos creados, capturados o replicados a nivel global en un año, alcance los 175 zettabytes (ZB) en el año 2025. Debemos tener en cuenta que un zettabyte equivale a un trillón de gigabytes. Para que resulte más fácil hacernos una idea de lo que representa esa cantidad de datos, si fuésemos capaces de almacenar la Datasfera Global completa en el año 2025, utilizando para ello discos ópticos DVD, tendríamos una pila de discos Blue-ray de una sola capa que daría la vuelta a la Tierra 222 veces.

Ligado al fuerte incremento en la producción, almacenamiento y procesamiento de datos que vivimos en la actualidad, encontramos iniciativas a nivel global orientadas a la promoción de los denominados Datos Abiertos. Los Datos Abiertos son aquellos datos *“que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, y que se encuentran sujetos, cuando más, al requerimiento de atribución y de compartirse de la misma manera en que aparecen”*². La filosofía detrás de los Datos Abiertos es promover, entre otras cosas, la participación ciudadana, la innovación, la transparencia, el autoempoderamiento y la investigación científica. Hoy en día, existe un número creciente de organizaciones, tanto públicas como privadas que, bajo la filosofía de los datos abiertos, ponen a disposición de la ciudadanía en general datos de interés de muy diversa índole. Instituciones y agencias estatales, universidades, empresas y otras organizaciones recolectan, procesan y almacenan grandes cantidades de datos relacionados no sólo con las personas, sino con todo tipo de actividades incluyendo fenómenos naturales y sociales. De allí que podamos encontrar datos abiertos relacionados con la industria, el medio ambiente, la economía, el deporte, el turismo o la salud.

La tendencia tecnológica actual, en la mayoría de los países del primer mundo, es la utilización de modelos guiados por los datos a la hora de dar soporte a la toma de decisiones. Ello implica no sólo generar y gestionar los datos de manera oportuna, sino que también hay que garantizar la calidad de éstos. Los modelos guiados por los datos son lo que los informáticos llamamos GIGO por las siglas en inglés de *“Garbage In Garbage Out”*. En otras palabras, si los datos no son buenos, tampoco lo son los modelos. Es por ello que la calidad de los datos se convierte en un factor fundamental en nuestros días.

La actual crisis sanitaria que se está viviendo a nivel global, causada por la pandemia asociada al virus COVID-19, es un claro ejemplo de la necesidad de, por un lado, políticas públicas orientadas a la gestión y explotación de los datos, así como a la promoción de los Datos Abiertos. La pandemia ha puesto de manifiesto las carencias que existen en muchos países en lo relacionado a la capacidad de generación, gestión y análisis de los datos con la finalidad de abordar un tema en específico. En estos momentos, existe una necesidad apremiante de convertir

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

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

los datos en acciones practicables que nos lleven a superar en el corto, medio y largo plazo los efectos causados por la actual pandemia. Países como Corea del Sur que, a través de su centro de prevención y control de enfermedades, ha puesto a disposición de la ciudadanía su información relacionada con el COVID-19, son ejemplos a seguir si queremos aprovechar el máximo potencial de los datos para afrontar un sinnúmero de problemas que afectan a los países de la región.

Retomando la frase, *“quien tiene los datos, tiene el poder”*, hago hincapié en que no sólo hay que tener los datos, hay que saber que hacer con ellos. Tenemos que ser capaces de convertir los datos en información, la información en conocimiento y ese conocimiento en acciones practicables. Políticas públicas que apuesten por los Datos Abiertos, la promoción de la investigación científica y programas académicos en el área de la Ciencia de Datos y áreas afines son, hoy en día, un imperativo para la región. Además, estas acciones deben ir acompañadas de legislaciones que, por un lado, garanticen la privacidad y la protección de los datos de todas las personas y por otro, tomen en cuenta las implicaciones éticas que se podrían derivar del uso masivo de datos y de las nuevas tecnologías.