

Los perfiles profesionales para trabajar en *Big Data* y *Open Data*

Marta Ortiz-de-Urbina Criado, *Universidad Rey Juan Carlos*

Alberto Abella García, *FIWARE Foundation*

La información es hoy en día uno de los recursos más estratégicos para las empresas. El presente y el futuro de muchas empresas está asociado a su capacidad para recopilar, gestionar y analizar grandes cantidades de datos volcados en la red –*Big data*– (Juan, 2016). Estos datos permiten conocer el mercado, los hábitos y necesidades de los clientes y de la sociedad, en general, y tomar decisiones estratégicas. Si bien hay mucha información disponible gracias Internet y al desarrollo de la tecnología, lo que se encuentran las empresas es que todavía hay escasez de personas especializadas en el tratamiento de los datos y de la información, en especial, en el campo de los datos abiertos (*open data*).

En 2017, COTEC publicó un informe sobre la generación de talento *big data* en España. En este informe se destaca que el valor estimado de la economía del dato en Europa puede llegar a alcanzar el 4,7% del PIB, en 2020 y, en España, crece un 30% cada año. Por otra parte, señala que el 65% de las empresas pueden convertirse en no competitivas si no adoptan *big data*.

En España, el fenómeno del *big data* está en pleno proceso de desarrollo, pero faltan profesionales del *big data*, como comenta Javier Lahoz, Head of Data Engineering en Orange España y director del Máster en Big Data de ICEMD. La falta de oferta hace que sea una profesión bien pagada (Data Center Market, 2019). Por ello, el reto actual está en la generación de talento desde el sistema educativo y desde las empresas y las administraciones públicas.

Por otra parte, dentro de la economía del dato, se está desarrollando el movimiento *open data* que, desde la publicación en abierto de datos, a través principalmente de portales de datos abiertos, permite la reutilización de datos y su uso para la creación de nuevos productos y servicios (Abella, Ortiz-de-Urbina-Criado y De-Pablos-Heredero, 2019). En este caso, también es un reto la formación y captación de profesionales especializados en *open data*.

Con estos nuevos profesionales, los departamentos de recursos humanos tienen un papel relevante en la definición de los puestos de trabajo (Data Center Market, 2019). Entre las cualidades que tiene que tener un experto en *big data* están: capacidad para adaptarse al cambio, ser comunicativo y resolutivo, tener conocimientos de programación y capacidad para interpretar datos, así como dominar las áreas de matemáticas, estadística e informática (Juan, 2016).

Juan (2016) comenta que los roles más demandados en las empresas son: *Chief Data Officer* (CDO), *Business Data Analyst* (analista de datos), *Data Scientist*, *Data Engineer*, *Data Steward*, *Citizen Data Scientist* y *Data Artist*. Moneo (2016) añade algunos perfiles asociados a la industria de los datos, para trabajar en el manejo, interpretación y comercialización de los datos como *Data Evangelist*, *Data Manager*, *Data Protection Officer* y *Data Journalist*.

Chief Data Officer (CDO) es el que se encarga de liderar y gestionar datos y la analítica asociada al negocio, es el responsable de todos los equipos especializados en *big data* y tiene que asegurarse de la empresa sea *data driven*. Es un perfil muy demandado y su salario, en España, puede oscilar entre 40.000 y los 90.000 euros brutos al año (Juan, 2016) y, en grandes empresas, podría ser superior a los 100.000 euros.

Data Manager o gestor de datos es el responsable del diseño, mantenimiento y reutilización de los datos de una organización. Entre sus funciones está la búsqueda de información almacenada y la realización de estudios periódicos del contenido de las bases de datos, para que luego pueda ser reutilizada. Entre sus funciones están, además de definir protocolos y estándares para garantizar la calidad e interoperabilidad de los datos, el diseño de herramientas que sirvan para capturar nuevas fuentes de información tanto estructuradas como no estructuradas (p. ej., el contenido de las redes sociales), así como aplicar tecnologías para la visualización de la información. Para desempeñar estas funciones se requieren personas con conocimientos informáticos avanzados sobre el manejo de información, semántica de los datos y tecnologías de la información y, además, es necesario que tengan conocimientos sobre las áreas de operación de los clientes (Moneo, 2016).

Business Data Analyst o analista de datos es el que realiza el análisis de los datos (transforma los datos en información) y recoge las necesidades de los clientes para mostrárselas al *Data Scientist*. Son profesionales muy necesarios para las estrategias de marketing para las empresas (Juan, 2016). Según UNIR revista (2019), el analista de datos realiza diversas tareas en el ámbito de la ingeniería de datos. Entre estas funciones

están la identificación de los conjuntos de datos; la obtención con técnicas de recuperación de los datos para su análisis y procesamiento; la consulta de los datos obtenidos; y la limpieza de los datos incluyendo el análisis de la calidad de los valores presentes, la detección de valores inadecuados, la normalización de los datos y la detección de registros vacíos, entre otras cosas. Por otra parte, realiza el pre-procesamiento para transformar los datos y empaquetarlos en estructuras manejables.

Para ser un buen analista de datos se requiere tener visión de negocio, habilidades analíticas y creativas, capacidad de síntesis, capacidad crítica y facilidad para la comunicación. También tiene que tener conocimientos en informática, matemáticas y estadística, base tecnológica y de *Machine Learning*, conocimientos de sistemas de almacenamiento distribuidos, y experiencia en diseño y realización de cuadros de mando, sobre todo de inteligencia de negocio (UNIR revista, 2019). Son perfiles muy demandados y que según algunos estudios, en España, tienen un salario que pueden ir desde 24.000 euros para profesionales junior hasta los 50.000 para los seniors (UNIR revista, 2019), aunque esto puede variar mucho en función de la empresa y del país. Por eso, para otros autores, como Juan (2016), su salario puede oscilar los 50.000 y los 60.000 euros brutos anuales.

Data Evangelist es el que se encarga de promover la transición hacia un modelo organizacional basado en los datos y se aplica en varios campos como el del *big data* y el del *open data*. Su papel principal es lograr que haya una “corriente de pensamiento favorable a la adopción de tecnologías para la reutilización de datos”. Su función es cuasi-comercial, ya que se encargan de identificar “early-adopters” dentro de las organizaciones para adoptar nuevas tecnologías y colaboran en el desarrollo de herramientas para reutilizar los datos. También, tienen que identificar a los clientes potenciales, tanto internos como externos, y ayudarles a definir sus necesidades. Otra de sus tareas es documentar casos de éxito y participar en conferencias de sensibilización y eventos de capacitación para explicar los factores de éxito. Para realizar estas tareas se requiere un gran conocimiento del negocio y un conocimiento avanzado de la tecnología (Moneo, 2016).

Data Scientist o científico de datos es el encargado de traducir grandes volúmenes de datos y transformarlos en información útil para la empresa. Es la evolución del perfil de analista de datos, ya que combina las técnicas de análisis estadístico tradicional con nuevas técnicas para la recolección y análisis de grandes volúmenes de datos (*big data*). Es el responsable de decidir qué técnicas se aplicarán al análisis de los datos y de

determinar la fiabilidad de los resultados. Se sitúan entre la parte operativa y la corporativa, y buscan nuevas oportunidades negocio a través de la reutilización de la información. Se encargan de diseñar algoritmos para automatizar el procesamiento de distintas fuentes de datos, y también participan en el diseño de herramientas de *machine learning* para mejorar los resultados del análisis (Moneo, 2016).

Para realizar estas funciones se necesita tener conocimientos matemáticos, estadísticos y de programación -para un mayor detalle sobre este tipo de conocimientos, ver Casas (2017)- y, además, tener una clara visión del negocio y habilidades comunicativas, ya que son los que tienen que dar a conocer el resultado de su trabajo al resto de la organización (Juan, 2016). En las ofertas de trabajo se suelen pedir candidatos apasionados y orientados a la acción, idealistas, capaces de pensar en grande y sin aceptar las inercias de las organizaciones y curiosos y autónomos, ya que se espera que sean ellos los que guíen a la organización a establecer una nueva forma de trabajar (Moneo, 2016). Es otro de los perfiles más demandados con un sueldo variable, según su área de especialización, que puede oscilar entre los 55.000 a los 80.000 euros brutos anuales (Juan, 2016).

Data Engineer o ingeniero de datos es el encargado de distribuir datos de manera accesible al *Data Scientist* y de transformarlos y hacerlos llegar a los repositorios donde los usan el resto de los profesionales. Están más especializados en la gestión de bases de datos, en los sistemas de transmisión y en sistemas de procesamiento y de programación (Juan, 2016).

Data Steward o administrador de datos es el que se encarga de velar por la calidad, la seguridad, la integridad y la disponibilidad de los datos. Tiene que saber utilizar los datos dentro del proceso de negocio y presentarlos a toda la organización (Juan, 2016).

Citizen Data Scientist no tiene una formación específica en *Data Scientist* pero aporta valor por su experiencia y por su conocimiento del significado de los datos, realizando tareas analíticas y de gestión de datos a través de herramientas más sencillas (Juan, 2016).

Data Protection Officer se encarga de la protección de datos y de asesorar sobre la recogida, uso, acceso o tratamiento de datos personales dentro de una organización. Entre sus funciones están la de crear, difundir y verificar el cumplimiento de las políticas corporativas de tratamiento de datos personales, y vigilar la seguridad y cumplimiento de los sistemas, de forma que los datos personales se traten conforme a la normativa aplicable. Se relacionan directamente con la dirección de las empresas y se

coordinan con las unidades tecnológicas para detectar escenarios de riesgo y adaptar las políticas corporativas, de forma que se pueda asegurar la actualización constante de las medidas de seguridad y reducir las posibles amenazas al tratamiento inadecuado de datos de carácter personal. En Europa, el reglamento de protección de datos ha hecho que sea un perfil muy relevante para las empresas. De hecho, para muchas de ellas es un perfil obligatorio. Estos profesionales tienen que tener conocimientos jurídicos sobre protección de datos y medidas técnicas y organizativas y, también, tienen que tener conocimientos específicos del sector en la que desempeñan sus funciones. Además, se valora que tenga conocimientos tecnológicos en temas relacionados con ciberseguridad y manejo de riesgos (Moneo, 2016).

Data Artist es el encargado de seleccionar y crear los gráficos, las infografías o las herramientas visuales que ayuden a la organización a entender los volúmenes de datos complejos (Juan, 2016). Es el responsable de seleccionar las cartografías y la forma de presentar los datos geo-posicionados respecto del territorio.

Data Journalist o periodista de datos es el que produce información a partir de datos, por lo que proviene del mundo del periodismo. Es el encargado de generar historias basadas en datos y de difundirlas en los medios de comunicación y redes sociales. Entre sus tareas principales también está extraer de los resultados de los análisis los mensajes institucionales más adecuados. Además, se encarga de realizar alianzas internas y externas para asegurar que llega, a las audiencias clave, la información precisa. Son profesionales que tienen que tener gran capacidad de redacción, análisis y conocimientos avanzados de marketing y diseño web, así como conocimientos del tratamiento de datos. También es necesario que conozcan bien la misión de la organización (Moneo, 2016).

Fuentes de información utilizadas

- Abella, A.; Ortiz-de-Urbina-Criado, M.; De-Pablos-Heredero, C. (2019). Meloda 5: A metric to assess open data reusability. *El profesional de la información*, 28(6), e280620. <https://doi.org/10.3145/epi.2019.nov.20>
- Casas, J. (2017). El perfil de un *Data Scientist*. Artículo publicado en el blog de la UOC, el 06/07/2017. Disponible en <http://informatica.blogs.uoc.edu/2017/07/06/el-perfil-de-un-data-scientist/> [15/03/2020].
- COTEC (2017)._Generación de talento *big data* en España. Informe elaborado por el grupo de trabajo de COTEC sobre *big data*, impulsado y coordinado por

Telefónica. Disponible en <http://informecotec.es/media/BIG-DATA-FINAL-web.pdf> [15/03/2020].

- Data Center Market (2019). Eda de oro del *big data*: buen salario y alta demanda. Artículo publicado en *computing*, el 23/09/2019. Disponible en <https://www.computing.es/analytics/noticias/1114223046201/edad-de-oro-del-big-data-buen-salario-y-alta-demanda.1.html> [15/03/2020].
- Juan, C. (2016). Los perfiles profesionales más buscados de *big data*. Artículo publicado en IEBS, el 25/10/2016. Disponible en <https://www.iebschool.com/blog/profesionales-mas-buscados-big-data/> [15/03/2020].
- Moneo, A. (2016). 5 nuevos perfiles para la industria de los datos. Artículo publicado en <https://blogs.iadb.org/>, el 23/06/2016. Disponible en <https://blogs.iadb.org/conocimiento-abierto/es/5-nuevos-perfiles-para-la-industria-de-los-datos/> [15/03/2020].
- UNIR revista (2019). Analista de datos: ¿cuál es el perfil de estos profesionales? Artículo publicado en *UNIR revista*, el 27/11/2019. Disponible en <https://www.unir.net/ingenieria/revista/noticias/analista-de-datos-cual-es-el-perfil-de-estos-profesionales/549204692005/> [15/03/2020].

Fecha del caso: marzo, 2020.

Palabras clave: análisis de puesto de trabajo, perfil profesional, *big data*, *open data*.

Preguntas sobre el caso

1. Imagínese que trabaja en un departamento de recursos humanos como “Analista de Puestos” y que le han encomendado la tarea de realizar un estudio sobre los puestos de trabajo asociados a *big data*. Haga una propuesta de proceso de realización del análisis del puesto de trabajo, identificando las etapas y lo que hay que hacer en cada una de ellas.
2. Identifique las fuentes de información que se pueden utilizar para realizar el análisis de puestos de trabajo asociados a *big data*.
3. Elija dos de los perfiles profesionales que aparecen en este caso, cree una idea de puesto de trabajo y proponga una descripción y especificación para cada uno de ellos.

4. ¿Qué desafío supone para la organización que los salarios de algunos de estos profesionales sean superiores a los de sus responsables operativos?