



¿El científico de datos es un científico?: finalmente, las dudas se despejan

Cultura, 21/09/2020

Los datos parecerían estar recibiendo en la actualidad mucha atención por parte de una importante cantidad de individuos, empresas e instituciones a nivel mundial o, por lo menos, de los que estarían relacionados, de modo significativo, con tecnologías digitales. Así, por ejemplo, algunas empresas creerían haber encontrado en los datos una nueva mina de oro para ser explotada. De hecho, a los datos se los podría estar considerando hoy como el oro de la era digital, según sus sugerentes.

Se conoce, sin embargo, que los datos ya se habrían empleado en el antiguo Egipto al menos desde el lapso relativo a la 4.a dinastía faraónica (Antes de la era común) [1]. Tales datos, así, se habrían obtenido por medio de censos de población, y usados para mejorar las recaudaciones de impuestos, por ejemplo. O sea, la producción y utilización de datos, relativos al quehacer humano, no es algo que se podría decir que es moderno o contemporáneo.

Entonces ¿Por qué hoy los datos resultarían ser muy atractivos para algunos individuos y organizaciones? La respuesta central, así, podría estar vinculada a la generación sin precedentes de vastas cantidades de datos en el mundo en función sobre todo de las tecnologías digitales. Esto es, si bien es cierto que se podría decir que hace mucho tiempo ya existían los datos, pero no en la cantidad astronómica de la que hoy dispone la humanidad o, al menos, una parte importante de ella.

Ahora bien, el tratamiento de datos masivos parecería requerir de una nueva profesión (conocida por algunos, como Ciencia de datos y, por otros, como Analítica de datos) de naturaleza al menos interdisciplinaria. Y, en efecto, quien se dedicara a esa profesión sería conocido como científico de datos o analítico de datos [2]. Pero ¿el científico de datos es un científico?

Se pretende, en ese sentido, contribuir a la solución de la cuestión arriba formulada apelando a una comparación entre lo que se puede entender, de una parte, por científico de datos y, de otra, por científico. A su vez, con el fin de llevar a cabo tal cotejo se utilizarán algunos criterios y descripciones vinculados a aquellas profesiones.

Un científico de datos, así, puede ser caracterizado como un profesional quien para hacer su trabajo se vale de ciertas herramientas digitales y, además, de distintas competencias relativas a diferentes especialidades, como la informática, estadística y matemática [3]. Mientras que su desempeño tiene que ver con, entre otras cuestiones, analizar datos, generar hipótesis, información y conocimiento útiles a fin de contribuir a la resolución de problemas de la vida cotidiana, sobre todo en ciertos ámbitos, como en el de los negocios.

De modo que esa profesión podría ser fundamentalmente de naturaleza pragmática.

Un científico, en cambio, o investigador científico, puede ser distinguido como un profesional quien para realizar sus estudios se vale de algunos instrumentos, no necesariamente digitales y, asimismo, de diversas competencias relativas a experticias concretas de índole disciplinaria, interdisciplinaria o transdisciplinaria [4]. Su trabajo, asimismo, en términos generales, se trata de darle solución en la medida de lo posible a problemas de naturaleza científica que se identifican en trabajos científicos previamente realizados, con la posible excepción de problemas relativos a fenómenos que por primera vez son estudiados.

A todo esto, existen esencialmente tres tipos de investigaciones: básica, aplicada y experimental; según su propósito [5]. Así,

por ejemplo, los estudios básicos se relacionan con el conocer por conocer, es decir, no está orientada a la resolución de cuestiones prácticas. Mientras que los aplicados buscan apoyar, al menos en cierto grado, a la solución de dichas cuestiones.

Tampoco es inoportuno indicar que algunos científicos activos actualmente no se dedican o, al menos no solo se dedican, a hacer investigación, sino a realizar actividades de gestión y administrativas, por ejemplo.

Las dos definiciones anteriormente referidas, por tanto, contribuirían a la identificación de ciertos criterios, los que, a su vez, favorecerían a la determinación de algunas similitudes y semejanzas en relación a las dos profesiones que aquí interesan.

Así pues, la curiosidad, la argumentación y el pensamiento crítico serían parte de las particularidades que caracterizarían tanto al analítico de datos [6] como al científico [7]. Sobre este último, sin embargo, es importante mencionar que son competencias indispensables a fin de que él pueda realizar al menos sus investigaciones científicas.

Otra cuestión en la que, igualmente, esos dos profesionales coincidirían tiene que ver, a más de contar con las competencias requeridas, con la identificación de un espacio en el que tanto el analítico de datos [8] como el científico [9] se sientan cómodos para hacer sus aportaciones. Pero, esta situación por lo general no sucedería así, sobre todo, en el caso de los científicos, es decir, estos comúnmente no tomarían decisiones sin que algún otro actor y/o elemento intervenga en absoluto respecto a las acciones a realizarse conectadas al ejercicio de su profesión.

Puesto que las decisiones, por ejemplo, respecto a dónde laborar por parte de un científico y, más que todo, las que tienen que ver con el proceso de producción del conocimiento científico pueden también estar influenciadas por el aspecto social [10]. Por ejemplo, una institución que contrata a un científico para hacer estudios puede terminar disponiéndolo al hombre de ciencia qué temas investigar, y no ser él quien, como experto, los determine.

También la oportunidad de trabajar con datos, conjeturas, hipótesis, tesis e inferencias sería compartida en cierta medida por quienes se desempeñan, respectivamente, en las dos ocupaciones antes citadas.

Pero ¿Por qué ocurriría esa semejanza a medias? Porque el analítico de datos trabajaría solo con hipótesis empíricas [11]. Mientras que el científico también trabaja con hipótesis teóricas, al menos los investigadores que son partidarios de algunas de las vertientes de la epistemología Racioempirista [12], por ejemplo.

Algunos autores, a propósito, consideran que los datos son construcciones y, por tanto, ellos pueden ser falsos o verdaderos, o estar correctos o no; pero los hechos no y, por ende, ellos son reales o imaginarios [13]. Mientras que otros sostienen que ambos son construcciones [14], es decir, tanto los datos como los hechos pueden ser falsos o verdaderos, o estar correctos o no.

Ahora bien, algo, por ejemplo, en lo que sí se puede observar una clara diferencia entre el científico de datos y el científico se relaciona con que los conceptos, preguntas, sus respuestas, metodologías y saberes con los que trabaja un científico se encuentran articulados a un cuerpo de conocimientos científicos acumulados, de manera no lineal, resultado de estudios que se han llevado a cabo en el transcurso del tiempo y, asimismo, con que ellos dependen de sus respectivas disciplinas o transdisciplinas científicas [15]. Mientras que a los datos con los que trabaja un científico de datos no les ocurre lo mismo [16].

Esa diferencia, así, quizás se puede esclarecer, de alguna forma, por medio de la siguiente analogía: un analítico de datos es como un fontanero (o gasfitero). De modo que como un gasfitero quien para solucionar una fuga de agua en el baño de una casa no necesariamente requiere revisar todo el sistema de tubería, aquel analítico no necesita precisamente examinar todo el sistema organizacional a fin de resolver una traba de una empresa.

Ahora bien, es preciso mencionar que en ciertos casos el diseño de una solución para resolver un problema específico de una empresa puede, a su vez, crear otro inconveniente en otros procesos organizacionales, por ejemplo, que, a la vez, se lo tenga

que arreglar antes de implementar la solución en cuestión, pero en todo caso esa situación no va a llegar a ser de naturaleza estructural.

En contraste, un científico, a diferencia de un fontanero, sí tiene que pasar revista al menos a casi todo el sistema de tubería. Esto es, un hombre de ciencia a fin de identificar un problema de índole científico, proponer una solución posible y determinar cómo llegar a dicho desenlace debe realizar una lectura de, al menos, casi todo el conocimiento que existe, relativo al fenómeno a investigar, u objeto de estudio, de su interés, hasta la fecha en la que se propone iniciar su investigación, con la eventual excepción de fenómenos que no hayan sido previamente indagados en absoluto.

En conclusión, y en esta primera aproximación a la solución de la pregunta que dio pie al desarrollo del presente escrito, el examen realizado, y expuesto en los párrafos anteriores, sugiere la formulación de, al menos, las siguientes tres declaraciones: la primera, un científico de datos no es un científico. La segunda, la transición de científico de datos a científico podría ser menos fácil que de este a aquel. Y la tercera, un profesional dedicado a la Ciencia de datos podría ser más bien identificado, de un modo apropiado, como anatómico de datos (sobre esto ver nota #2) o, por último, como analítico de datos o analista de datos masivos. Y, efectivamente, a la Ciencia de datos se la podría más bien llamar, de manera adecuada, Anatómica de datos o, si se prefiere, Analítica de datos.

[1]Para más información, remítase a Moreno, Juan (2008). Estates (Old Kingdom). In Elizabeth Froid and Willeke Wendrich (eds.), UCLA Encyclopedia of Egyptology, Los Angeles. Disponible en:

https://www.researchgate.net/publication/235412579_Estates_Old_Kingdom

[2]El nombre del supuesto nuevo profesional, como analítico de datos, a criterio del autor del presente escrito no lo podría identificar de una manera adecuada a él. Puesto que este no solamente tiene que ver con realizar análisis, sino también con hacer síntesis, por ejemplo. Y, aunque esto sea una cuestión trivial, un nombre que lo podría identificar mejor sería anatómico de datos. A su vez anatómico es una palabra compuesta por ciertas sílabas que pertenecen a los vocablos analista y sintético.

[3]Para más información, remítase a Peng, Roger y Matsui, Elizabeth (2015). The Art of Data Science. A Guide for Anyone Who Works with Data. Canada: Leanpub book, y a Patil, D.J. (2011). Building data science teams. The skills, tools and perspectives behind great data science groups. USA: O'Reilly.

[4]Para más información, remítase a Hernández, Roberto, Hernández, Carlos y Baptista, Pilar (2014). Metodología de la investigación. México: McGraw Hill Education.

[5]Organización para la Cooperación y el Desarrollo Económico (2013). Manual de Frascati. Propuesta de prácticas estándar para encuestas sobre proyectos de investigación y desarrollo experimental. Chile: F-Iniciativas Ltda.

[6]Para más información, remítase a Patil, D. J. et al. (s. f.). The Data science Handbook. Advice and insights from 25 amazing data scientists.

[7]Para más información, remítase a **Sánchez, Verónica (2016)**. Curiosidad, el motor de los científicos: Fernando Magaña. *Organización de Estados Iberoamericanos*. Disponible en:

<https://www.oei.es/historico/divulgacioncientifica/?Curiosidad-el-motor-de-los-cientificos-Fernando-Magana>

[8]Para más información, remítase a Álvarez, José Antonio y Coll-Serrano, Vicente (2018). "Científico de datos", la profesión del presente. *Métodos de Información*, 9, 16, 71-87. Disponible en:

<https://www.metodosdeinformacion.es/mei/index.php/mei/article/view/IIMEI9-N16-113129/967>

[9]Para más información, remítase a [Wasserman](#), Moisés (2019). Curiosidad, conocimiento y ciencia. Es inocua la discusión

sobre si es más importante la ciencia básica o la aplicada. El Tiempo. Disponible en:

<https://www.eltiempo.com/opinion/columnistas/moises-wasserman/curiosidad-conocimiento-y-ciencia-columna-de-moises-wasserman-419>

[10]Para más información, remítase a Kreimer, Pablo (1999). De probetas, computadores y ratones. La construcción de una mirada sociológica sobre la ciencia. Bernal: Universidad Nacional de Quilmes. Disponible en:

https://www.researchgate.net/publication/286928245_De_probetas_computadoras_y_ratones

[11]Para más información, remítase a Usama Fayyad and Hamutcu, H. (2020). Toward Foundations for Data Science and Analytics: A Knowledge Framework for Professional Standards. *Harvard Data Science Review*. Disponible en:

<https://assets.pubpub.org/ktdlin1u/6b19368a-b697-4991-9d27-9c9643ebe2db.pdf>

[12]Bunge, Mario (1999). *Buscar la Filosofía en las Ciencias Sociales*. México: Siglo Veintiuno Editores, S. A. de C. V.

[13]Ibídem, Bunge, 1999.

[14]Para más información, remítase a Knorr Cetina, Karin (2005). *La fabricación del conocimiento. Un ensayo sobre el carácter constructivista y contextual de la ciencia*. Bernal: Universidad Nacional de Quilmes Editorial.

[15]Para más información, remítase a Batthyány, Karina y Cabrera, Mariana (2011). *Metodología de la investigación en las Ciencias Sociales. Apuntes para un curso inicial*. Montevideo: Udelar CSE. Disponible en:

https://www.colibri.udelar.edu.uy/jspui/bitstream/20.500.12008/9491/1/01_FCS_Batthianny_2011-07-27-lowres.pdf; Ibídem, Hernández, Hernández y Baptista, 2014; y a Kuhn, Thomas (2004). *La estructura de las revoluciones científicas*. México: Fondo de Cultura Económica.

[16]Para más información, remítase a Aguilar, Luís (2013). *Big data. Análisis de grandes volúmenes de datos en organizaciones*. México: Alfaomega Grupo Editor, S.A. de C.V.