

El trabajo colaborativo en la ciencia de datos desde un enfoque socioformativo

Collaborative Working in Data Science from a Socioformative Approach

Gabriela López Quesada galoqu@yahoo.com.mx Centro Universitario CIFE, México

Sergio Tobón Tobón stobon@cife.edu.mx Centro Universitario CIFE, México

Resumen: Se analizó el trabajo colaborativo desde un enfoque socioformativo en el proceso de la ciencia de datos mediante un análisis documental con enfoque cualitativo. Se realizó un análisis documental con enfoque cualitativo. Para la exploración se utilizó una serie de algoritmos de búsqueda, tanto en inglés como en español, enfocando la atención en documentos de años recientes, en su mayoría de 2019 a 2022. El resultado arroja que la falta de colaboración lleva a las áreas a trabajar de forma aislada empleando información de diferentes fuentes y enfoques limitando el logro del objetivo planteado.

Palabras clave: asignación de roles, socioformación, toma de decisiones, trabajo en equipo

Abstract: Collaborative work was analyzed from a socio-formative approach in the data science process through a documentary analysis with a qualitative approach. A documentary analysis was carried out with a qualitative approach. For the exploration, a series of search algorithms were used, both in English and Spanish, focusing attention on documents from recent years, mostly from 2019 to 2022. The result showed that the lack of collaboration leads the areas to work differently. in isolation using information from different sources and approaches, limiting the achievement of the stated objective.

Keywords: role assignment, data science, socio-formative approach, decision making

Introducción

La ciencia de datos (CD) obtiene información de diversas áreas con las que comparte sus principios y prácticas, como estadística, métodos científicos, inteligencia artificial y análisis de datos, entre otros, pueda reconocer, analizar, limpiar y dar valor a los datos obtenidos para su utilización y aplicación en la toma de decisiones a corto, mediano y largo plazo (Méndez & Porven, 2018; Lemus & Pérez, 2020; Oracle-México, 2020). En

este sentido, la CD es colaborativa ya que presenta momentos en los que las diferentes áreas del conocimiento deben trabajar en conjunto empleando diversos instrumentos para garantizar la calidad, pertinencia y validez de la información, sin dejar de lado temas de privacidad y seguridad.

Los principios de la CD involucran tomar los datos obtenidos y organizarlos; trabajar con información desconocida que también se deberá organizar; y reconocer el grado de imprecisión de los resultados para generar modelos y pilotos que consideren este factor a nivel local y global. La colaboración es la línea en la que convergen la CD y el enfoque socioformativo para contribuir con la comunicación de manera asertiva y la aplicación de los resultados desde la socialización a distintos niveles favoreciendo su comprensión (Nivela-Cornejo, Echeverría-Desiderio & Espinosa-Izquierdo, 2019).

El presente artículo analiza y documenta diferentes aspectos sobre la CD; el primero delimita y define el concepto mediante el uso de la cartografía conceptual. Tiene como objetivo ofrecer una breve visión sobre el trabajo colaborativo (TC) en la CD desde un enfoque socioformativo; el tercero se centra en algunos de los beneficios que la CD tiene; y el cuarto y último plantea el perfil profesional que el científico de datos debe tener, estos últimos tres artículos desde un análisis documental con enfoque cualitativo.

Método de investigación

Tipo de estudio

Para la realización de este artículo se llevó a cabo un análisis documental con enfoque cualitativo y de manera integral para realizar el estudio interno y externo de diversos documentos que apoyaron en la obtención de la información necesaria para la mejor comprensión (Dulzaidés & Molina, 2004; Pinto & Gálvez, 2011) sobre el TC y la contextualización de la información desde un enfoque socioformativo en la CD. El análisis documental es aquel que se centra en la búsqueda, localización, selección, organización, análisis, reflexión y comprensión de datos secundarios de diversas fuentes documentales, audiovisuales, impresas y electrónicas de fuentes primarias y secundarias mediante buscadores académicos, con referencias y apoyo de materiales electrónicos y en papel (Arias, 2016; Bermeo-Yaffar *et al.*, 2016) para responder al objetivo planteado en el estudio.

Como estrategia de investigación se desarrolló un análisis documental cualitativo e integral mediante un registro estructurado para describir e informar (Dulzaidés, & Molina, 2004; Pinto & Gálvez, 2011) sobre la relación que guardan el TC y el contexto desde la socioformación y la CD.

Criterios para el estudio

Para el estudio se seleccionaron y analizaron una serie de artículos, libros y capítulos de libros que siguieron los parámetros descritos a continuación:

1. Se realizaron diversas búsquedas de artículos empíricos, artículos de revisión, libros y capítulos de libros en bases de datos bibliográficas como Scopus, Web of Science, Scielo, Redalyc, Dialnet, World Wide Science, Refseek, Elsevier y Google Scholar, básicamente; la exploración se llevó a cabo empleando una serie de algoritmos de búsqueda, tanto en inglés como en español, enfocando la atención en documentos de años recientes, siendo en su mayoría de 2019 a 2022, de acceso libre como limitado y consultados, en la medida de lo posible, en el sitio en donde originalmente fueron publicados.

2. Las palabras clave propuestas fueron "ciencia(s) de datos", "contexto", "problema contextualizado", "situaciones del contexto", "contextualizar", "socioformación", "enfoque socioformativo", "trabajo colaborativo" y "trabajo en equipo", "colaboración" y "cooperación" combinadas con una o varias de las siguientes palabras "colaboración", "cooperación", "equipo", "grupo", "rol/es", "contexto", "problema", "situación", "enfoque socioformativo", "socioformación", "información", "datos", "ciencia", "tecnología", "comunicación", "necesidades", "global", "local" y "área(s) del conocimiento" todas ellas tanto en inglés como en español.

3. Los documentos seleccionados fueron artículos de revistas indexadas, capítulos de libros y libros de editoriales reconocidas, así como materiales ofrecidos por centros de investigación, empresas y universidades, todos ellos de acceso libre y restringido. Todos los recursos empleados cuentan con la información mínima requerida para ser referenciados bajo las normas APA y se comparten las ligas de acceso en general y siempre que fue posible obtenerlas. En su mayoría, los artículos fueron consultados en la revista o sitio de origen y en el idioma en que se escribieron –inglés y español-. Cabe mencionar que se revisaron diversas páginas Web de instituciones inmersas en las

ciencias de datos, el big data y la minería de datos, así como conferencias, eventos y documentos desde la socioformación como complemento y para cotejar la información obtenida con puntos de vista de expertos actuales y vigentes.

4. Todos los documentos empleados sirvieron como base para la elaboración de este estudio, mismos que se enlistan en las referencias de este artículo.

Resultados

Socialmente hablando, el ser humano está en constante transformación, empleando y desechando, teniendo diversas experiencias y enfrentando situaciones que pueden ser mejoradas o modificadas. Es así como las personas generan información que los expertos toman y traducen para que sea usada en beneficio de quienes la requieren. En este sentido, el concepto de CD regularmente refiere a la información digital, sea cual sea su volumen, que se genera y procesa desde los fenómenos sociales que la incrementan y modifican a gran velocidad en contextos domésticos y generales, todos ellos interrelacionados como una red que requiere de un minucioso análisis para ser aplicados en la mejora de situaciones reales (Del-Castillo & Arza, 2021) o resolución de problemas complejos contextualizados (Klenzi, Malberti & Beguerí, 2019; Lemus & Pérez, 2020; Rodríguez, 2020; Vázquez, 2020) acordes a las necesidades del entorno, partiendo siempre de un problema contextualizado que se comprende y cuestiona con base en un objetivo, para desde el enfoque socioformativo, cumplir el desafío de mejorar una situación mediante el desarrollo de un producto que se pueda poner en acción (González, 2018). Este proceso de elección de alternativas y generación de soluciones se da en la vida diaria para enfrentar conflictos que posibilitan aprender y tomar decisiones futuras.

Así como los individuos emplean su conocimiento, experiencia de vida y lo que les rodea para optar por la mejor opción, la CD se enfoca en diversas áreas del conocimiento y de tanta información como le es posible para ofrecer alternativas y soluciones. Esto implica procesos que involucran a las ingenierías, matemáticas, estadística, análisis de datos, procesamiento de datos, actividades comerciales, comunicación y a los expertos en el área de la que se trate (Gibert *et al.*, 2018; Méndez & Porven, 2018; Vázquez, 2020) de manera colaborativa para que puedan traducir diversos volúmenes de información en formatos comprensibles para la toma de decisiones y predicción de comportamientos en contextos reales (UTEQ, 2021), así como, para generar nuevo conocimiento al recabar,

manejar, procesar, limpiar, explorar, modelar e interpretar datos (Yu & Kumbier, 2020). La CD requiere de diversas disciplinas para el logro de la meta fijada que sin la colaboración de sus equipos los resultados no estarían completos ni serían comunicados en tiempo y forma.

La colaboración entre diversas áreas hace de la CD una multidisciplinaria que involucra a todos los que manejan cada uno de los procesos, como aquellos que mejor conocen el sector en el que se está trabajando; asimismo, es primordial conocer a fondo la situación para la cual se tomarán las decisiones con base en datos verídicos y confiables (Yu & Kumbier, 2020) teniendo especial cuidado en que cada uno de los equipos a colaborar emplee la misma información, la cual, generalmente, es almacenada en espacios a los que todos los colaboradores tienen acceso. Es entonces que la CD converge con otras áreas y enfoques como el de la socioformación para lograr su proyecto.

Al TC con un enfoque socioformativo lo caracterizan 6 ejes que fácilmente se alinean al proceso que la CD realiza.

1. El TC requiere de la definición del objetivo o meta a lograr para la resolución del problema contextualizado; en tanto que el objetivo o meta esperado lo establece la empresa en cuestión con la guía y colaboración del experto en CD;
2. El TC se basa en el establecimiento de un plan a seguir que deberá incluir la descripción de las acciones, los tiempos y momentos, la responsabilidad y los recursos con los que se cuenta; para la CD dicho plan también considera los resultados que se puedan lograr;
3. Para el TC como para la CD, la distribución de roles se determina y reformula en cada uno de los procesos y pasos que se siguen;
4. Para la CD, posterior a la planificación y determinación de roles se cubren los procesos de generación, aplicación, evaluación, reformulación y re-aplicación de los pilotos y modelos. Para el TC estos procesos conforman el proceso de mejora continua;
5. La interacción y comunicación de manera asertiva se contempla desde el inicio del TC y de la CD; cabe mencionar que la comunicación clara y oportuna son características que la CD tiene para alcanzar el éxito esperado;

6. Para el TC como para la CD el reconocimiento de la responsabilidad individual y colectiva son factores que alejan al proyecto de riesgos y fracasos.

El TC y la CD cierran su ciclo de trabajo con la determinación de logros y áreas de mejora tanto del proceso como del producto o modelos generados (Juárez et al. 2019, Oracle-México, 2020, ORACLE, 2022).

La CD, desde el enfoque socioformativo, se basa en el TC al compartir el conocimiento al asignar roles y responsabilidades bien definidos beneficiando el desarrollo del proyecto y la resolución de problemas contextualizados mediante la identificación, interpretación y argumentación desde el área de conocimiento de cada uno de los involucrados (Nivela-Cornejo, Echeverría-Desiderio & Espinosa-Izquierdo, 2019, Ruiz-Velazco & Bárcenas-López, 2019).

Quienes llevan a cabo los procesos necesitan tener conocimientos en diversas áreas y colaborar con los expertos en cuestión, como pueden ser ciencias, matemáticas, estadística y ciencias computacionales, entre otras (Gibert *et al.*, 2018; Del-Castillo & Arza, 2021) ya que las situaciones que se busca mejorar son manejadas mediante proyectos multidisciplinarios considerando su contexto con base en la metodología que la socioformación propone (Juárez *et al.*, 2019; Martínez, Tobón & Soto, 2021) debido a que sus procesos imponen el TC directo e indirecto con especialidades como las TIC, estadísticas, minería de datos y análisis de datos ya que cada una de ellas pone especial atención en una faceta del proceso ofreciendo panoramas claros sobre la situación que se busca resolver (Rollins, 2015; Hernández-Leal, Duque-Méndez & Moreno-Cadavid, 2017; Gibert *et al.*, 2018; Favaretto *et al.*, 2020; Zhang, Muller & Wang, 2020; Linares-Morales, 2021). En caso contrario, cada una de las áreas involucradas estaría trabajando con aquellos datos que considera prudente sin que necesariamente fuesen los mismos y con metas contrarias que podrían arrojar soluciones desfavorables.

Desde el enfoque socioformativo, la colaboración se da para desarrollar proyectos que serán aplicados en diferentes ámbitos de la vida mediante la distribución de responsabilidades y compartiendo el conocimiento sin priorizar o minimizar los resultados parciales, sino que se trabaja de manera dinámica y horizontal para diversificar el proyecto en cuestión y la obtención de resultados recomendables, resaltando los logros y compartiendo retroalimentación y gestión de los resultados (Zhang, Muller & Wang, 2020); para ello, la CD se nutre de la colaboración de diversas

áreas en las que cada individuo y equipo adquiere un rol primordial para lograr la resolución, parcial o total, de una situación a corto, mediano y/o largo plazo (Martínez, Tobón & Soto, 2021) El trabajo multidisciplinario y el colaborativo benefician la generación de nuevos contenidos que se compartirán para la toma de decisiones de una forma más innovadora y creativa.

El TC es, entonces, un ejercicio que privilegia y potencia la investigación para dar mayor conocimiento sobre las problemáticas que se buscan resolver presentando alternativas que cumplan con opciones vigentes y posibles para definir cursos de acción. Desde un enfoque socioformativo, dichas opciones serán socializadas para propiciar un mayor beneficio y su posible réplica en otros ambientes con situaciones similares (Ruiz-Velazco & Bárcenas-López, 2019), acción que la CD lleva a cabo al comunicar los resultados del proceso y proponer el camino a seguir.

Tanto la socioformación, como la CD, no dejan de lado el trabajo local, global ni a la tecnología en diversos ámbitos, como el laboral, lo cotidiano y lo académico, entre otros y proponen identificar, interpretar, argumentar y formular para resolver problemas contextualizados tomando ventaja de los talentos, conocimiento y habilidades de cada involucrado (Tobón *et al.*, 2015). Debido a que el TC refiere a trabajar con otros para lograr un objetivo que se tiene en conjunto, se pueden proponer una meta común o varias submetas para ir sorteando problemas menores combinando el conocimiento y recursos con que cada uno cuenta y en total comunicación partiendo de una propuesta o plan de trabajo que exige la definición de los roles clara y oportuna.

Desde la CD, la colaboración permite obtener diferentes perspectivas de una misma situación, documentando recomendaciones y descubrimientos conservando el rol que a cada quien le corresponde (Zhang, Muller & Wang, 2020). La interacción para la colaboración se da en cada una de las fases del proyecto en los diferentes ambientes y en la puesta en marcha de los resultados (Zhang, Muller & Wang, 2020).

El flujo de la colaboración inicia con la empresa o grupo de personas que presentan un problema sobre el que deben tomar decisiones y del que cuentan con información limitada; para continuar con el equipo de trabajo de la CD que intervendrá para complementar la información con datos vitales fijando los roles que asumirán, dependiendo de las habilidades y conocimiento que cada área tenga, colaboración que no solo se da entre los equipos de científicos de datos e ingenieros para alinear sus

metas, sino con analistas de negocios, gerentes y expertos que dominan el tema para que se garantice el cumplimiento de los objetivos empresariales para resolver problemas desde la innovación (Zhang, Muller & Wang, 2020). La colaboración no termina hasta la última fase del proyecto, una en la que se comunican los resultados y se ejecutan acciones con base en la toma de decisiones (Nivela-Cornejo, Echeverría-Desiderio, & Espinosa-Izquierdo, 2019; Ruiz-Velazco & Bárcenas-López, 2019). Por tanto, la colaboración es trabajar con un plan y de manera organizada, estructurada y manteniendo el control de cada paso que se da en el proceso.

Conclusiones

Cada vez más organizaciones requieren de la CD para impulsar los resultados que obtendrán tras llevar a cabo la toma de decisión, para ello, las empresas proporcionan información que ha sido extraída y almacenada y que regularmente no es suficiente por lo que requieren acceder a los datos correctos que le permitan generar un modelo y que son extraídos de diversas fuentes y comparaciones que se establecen con empresas de áreas similares o iguales; es decisión de quienes emplean la CD el realizar procesos de forma colaborativa o independiente valiéndose solamente de los versados en datos generando una brecha enorme entre los resultados a obtenerse y la meta que se pretende alcanzar en la toma de decisiones. Cuando la CD adopta un entorno colaborativo se enfrenta a los desafíos que mantienen a la organización actualizada y le dan un mayor valor comercial; la colaboración también ofrece la comprensión y entendimiento compartidos y más profundos de la información y de los modelos que se proponen como más productivos y confiables permitiendo que la información sea accesible para quienes no manejan datos codificados.

La colaboración debe iniciar entre la empresa y el equipo de CD. Posteriormente se deberá establecer entre los científicos de datos, analistas, ingenieros y expertos que se asociarán y colaborarán en la estandarización y traducción de los datos, pero cada uno trabajando desde su estructura o lugar teniendo especial cuidado en emplear la misma información, por lo que la colaboración se basará en el uso y sustento de los mismos documentos, fuentes o lugares de almacenamiento de la información. Al entrelazar la CD y el TC es necesario considerar la multidisciplinariedad de ambas, entendida esta como el colaborar de las ramas y especialidades que pueden ser involucradas en el proceso desde su inicio, es decir, desde que se busca el conocimiento de la empresa y

hasta la toma de decisiones, incluso, en la propia aplicación de los modelos ya generados.

La CD ofrece a las empresas la aplicación rápida y oportuna de los modelos que ya han sido probados y ajustados para las decisiones tomadas para la obtención de resultados tangibles y positivos pretendiendo resolver los problemas planteados o mejorar las situaciones que la institución está viviendo. Dichos procedimientos llegan a puntos en los que las diferentes áreas se concentran en solo su labor, es el TC quien aleja a los grupos y equipos de trabajo de realizar su tarea de manera independiente y asilada empleando datos de fuentes y perspectivas diferentes, así, la colaboración lleva a todos los equipos a tomar como punto de partida una única fuente de datos compilados en los documentos que todas las áreas emplearán para realizar su labor, es común alojar estos documentos y fuentes de datos en nubes que propician que se acceda a ellos de forma remota.

El reto más grande que puede enfrentar la CD al trabajar de manera colaborativa es la misma transmisión de la información que se está trabajando considerando que las locaciones en las que las áreas se pueden encontrar son verdaderamente lejanas, en tiempo y espacio, así como, tomando en cuenta que el envío de información puede sufrir filtraciones y amenazas, de las que la misma empresa involucrada no se libra, un vacío que requiere de una profunda investigación desde la perspectiva de la ciberseguridad para brindar mayor seguridad y confianza a las empresas que requieren de la labor de la CD.

Referencias bibliográficas

- Arias, F. (2016). El proyecto de investigación: introducción a la metodología científica (Episteme (ed.); 6th ed.). <https://evidencia.com/wp-content/uploads/2014/12/EL-PROYECTO-DE-INVESTIGACIÓN-6ta-Ed.-FIDIAS-G.-ARIAS.pdf>.
- Bermeo, F., Hernández, J. S. & Tobón, S. (2016). Análisis documental de la v heurística mediante la cartografía conceptual. *Ra Ximhai*, 103–122. <https://doi.org/10.35197/rx.12.01.e3.2016.05.fb>.
- Castillo-Medina, C. A. (2021). Ciberseguridad: Por dónde Empezar *Revista de Tecnología*, 18(1), 1–12. <https://revistas.unbosque.edu.co/index.php/RevTec/article/view/3803>.

López Quesada, G., Tobón Tobón, S.

- Del-Castillo, M. & Arza, V. (2021). La ciencia digital en América Latina: alcance y beneficios. *Arbor*, 197(799), a595. <https://doi.org/10.3989/arbor.2021.799008>.
- Dulzaides, M. E. & Molina, A. M. (2004). Análisis documental y de información: dos componentes de un mismo proceso. *Acimed*, 12(2), 1-5. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84899874589&partnerID=40&md5=2c01cbbe5636c36240c08941c50e2d34>.
- Favaretto, M., de Clercq, E., Schneble, C. O. & Elger, B. S. (2020). What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. *PLoS ONE*, 15(2), 1-20. <https://doi.org/10.1371/journal.pone.0228987>
- Gibert, K., Horsburgh, J. S., Athanasiadis, I. N. & Holmes, G. (2018). Environmental Data Science. *Environmental Modelling and Software*, 106, 4-12. <https://doi.org/10.1016/j.envsoft.2018.04.005>.
- González, M. d. L. (2018). UVE socioformativa: estrategia didáctica para evaluar la pertinencia de la solución a problemas de contexto. *IE Revista de Investigación Educativa de La REDIECH*, 9(16), 133-153. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2448-85502018000100133.
- Hernández-Leal, E. J., Duque-Méndez, N. D. & Moreno-Cadavid, J. (2017). Big Data: an exploration of research, technologies and application cases. *Tecnológicas*, 20(39). <https://www.redalyc.org/pdf/3442/344251476001.pdf>
- Juárez, L. G., Tobón, S., Salas, G., Jerónimo, A. E. & Martínez, M. G. (2019). Desarrollo Sostenible: Educación y Sociedad. *Revista Electrónica de Medio Ambiente*, 20(1), 54-72. <https://www.ucm.es/iuca/volumen-20-numero-1-revista.elecde-medioambiente>.
- Klenzi, R., Malberti, A. & Beguerí, G. (2019). Propuesta didáctica inherente al área de ciencia de datos. XXI Workshop de investigadores en Ciencias de la Computación (WICC 2019, Universidad Nacional de San Juan). <http://sedici.unlp.edu.ar/handle/10915/77174>.

López Quesada, G., Tobón Tobón, S.

Lemus, D. & Pérez, R. (2020). Ciencia de datos y estudios globales: aportaciones y desafíos metodológicos. *Colombia Internacional*, 102, 41–62. <https://doi.org/10.7440/colombiaint102.2020.03>.

Linares-Morales, J. (2021). Una mirada desde el sur al tema de la ciberseguridad. *IPSA Scientia. Revista Científica Multidisciplinaria*, 6(1), 123–124. <https://latinjournal.org/index.php/ipsa%0ARevista>.

Martínez, J. E., Tobón, S. & Soto, J. A. (2021). Ejes claves del modelo educativo socioformativo para la formación universitaria en el marco de la transformación hacia el desarrollo social sostenible. *Formación Universitaria*, 14(1), 53–66. <https://doi.org/10.4067/S0718-50062021000100053>.

Méndez, N. & Porven, J. (2018). Ciencia de datos: una revisión del estado del arte. *UCE Ciencia*, 6(3). <https://doi.org/10.1016/j.aci.2014.10.001>.

Nivela-Cornejo, M. A., Echeverría-Desiderio, S. V. & Espinosa-Izquierdo, J. G. (2019). Herramientas digitales en el trabajo colaborativo. *Espiraes Revista Multidisciplinaria de Investigación*, 103–111. <https://doi.org/https://doi.org/10.31876/re.v3i25.444>.

Oracle-México. (2020). *Base de Datos*. <https://www.oracle.com/mx/database/what-is-database/>.

Pineda-de-Alcazar, M. Y. (2018). La Internet de las Cosas, el Big Data y los nuevos problemas de la comunicación en el Siglo XXI. *Mediaciones Sociales*, 17, 11–24. <https://doi.org/10.5209/meso.60190>.

Pinto, M. & Gálvez, C. (2011). Análisis Documental de Contenido. Procesamiento de la Información. In *E. Síntesis* (Ed.) Manual de ciencias de la información y documentación.

Rodríguez, J. M. (2020). Cambio tecnológico y adaptación de la oferta educativa a la nueva demanda de habilidades en el Uruguay. *Documentos de Proyectos*, 82. <https://repositorio.cepal.org/handle/11362/45902>.

Rollins, J. B. (2015). Metodología Fundamental para la Ciencia de Datos. 6. <https://www.ibm.com/downloads/cas/WKK9DX51>.

López Quesada, G., Tobón Tobón, S.

Ruiz-Velazco, E. & Bárcenas-López, J. (2019). *Trabajo colaborativo en entornos virtuales* (A. C. Sociedad Mexicana de Computación en la Educación (ed.)). <http://www.telematica.ccadet.unam.mx/recursos/eBook/libros2019/trabajoc.pdf>.

Tobón, S., González, L., Nambo, J. S. & Vazquez-Antonio, J. M. (2015). La Socioformación: Un Estudio Conceptual. *Paradigma*, 36(1), 7–29. <https://doi.org/10.13140/RG.2.1.2957.5289>.

UTEC. (2021). ¿Qué es la Ciencia de Datos y para qué se utiliza? Ciencia de Datos. <https://www.utec.edu.pe/blog-de-carreras/ciencias-de-datos/que-es-la-ciencia-de-datos-y-para-que-se-utiliza>.

Vázquez, A. (2020). Ciencia de Datos para Gente Sociable. Bookdown. https://bitsandbricks.github.io/ciencia_de_datos_gente_sociable/index.html.

Yu, B. & Kumbier, K. (2020). Veridical data science. *Proceedings of the National Academy of Sciences of the United States of America*, 117(8), 3920–3929. <https://doi.org/10.1073/pnas.1901326117>.

Zhang, A. X., Muller, M. & Wang, D. (2020). How do Data Science Workers Collaborate? Roles, Workflows, and Tools. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1), 1–23. <https://doi.org/10.1145/3392826>.