



## Fomento de la alfabetización de datos en todos los ámbitos

Stephanie Arreguit O'Neill (\*)

(\*) Innobridge Services Sàrl

La autora de este artículo es empleada de Innobridge Services Sàrl, una consultoría con base en Suiza. Para más información póngase en contacto con [info@innobridge.com](mailto:info@innobridge.com).

Dirección postal: Innobridge Services Sàrl, EPFL Innovation Park, Bât. A, 1015 Lausana, Suiza

### RESUMEN

A medida que las tecnologías proliferan en nuestra sociedad, el mundo está cada vez más interconectado y se recopilan grandes cantidades de datos que contienen información y conocimientos valiosos. Esta creciente cantidad de datos ya ha sido reconocida como un recurso con un gran potencial para aumentar la eficiencia y la sostenibilidad de las organizaciones, las ciudades y las sociedades. Están apareciendo nuevos trabajos dedicados a los datos, y personas que no son expertas en datos que deben leer y gestionar datos en su trabajo diario. Por lo tanto, la alfabetización de datos es una habilidad esencial para avanzar en nuestra sociedad de la información. Es fundamental que las personas, los negocios, las instituciones educativas, las comunidades y las sociedades entiendan el valor y las oportunidades que residen en los datos que se utilizan, así como los problemas éticos, de seguridad y de privacidad que pueden surgir por su mal uso. Este artículo aborda la necesidad de la alfabetización de datos, estudia su definición y las habilidades relacionadas, considera la necesidad de dicha alfabetización desde distintas perspectivas, y analiza las consideraciones necesarias para fomentar la alfabetización de datos en distintos ámbitos.

Palabras Clave: datos, alfabetización de datos, sociedad de la información.

### Introducción

Los datos se han hecho omnipresentes. Se han convertido en una característica esencial de nuestra vida diaria, tanto cuando analizamos ciertos datos para apoyar las decisiones que tomamos en el trabajo, como cuando buscamos información a la hora de elegir el coche que queremos comprar o cuando escuchamos las últimas noticias. La cantidad de datos registrados aumenta cada minuto, y surgen nuevos modelos de negocio que aprovechan las oportunidades que plantean. Se ha previsto que el volumen total de la esfera de datos mundial crezca desde los 33 Zo en 2018 hasta los 175 Zo en 2025 ([Reinsel, Venkatraman, Gantz, & Rydning, 2019](#)). Hoy en día los datos son una herramienta utilizada en la mayoría de profesiones ([Ridsdale et al., 2015](#)). Hojas de cálculo, indicadores clave de rendimiento o cuadros de indicadores aparecen en múltiples sectores, impulsando las decisiones y apoyando la supervisión de los equipos, proyectos o procesos. Además, grandes cantidades de datos se ponen a disposición del público como resultado del movimiento de datos abiertos. Si bien los datos abiertos ofrecen el potencial de impulsar el desarrollo económico y sostenible, así como el cambio social y político, con una mayor implicación, transparencia o rendición de cuentas, no ha cumplido las expectativas iniciales por problemas como la falta de mantenimiento de los portales

de datos o el acceso desigual ([Frank, Walker, Attard, & Tygel, 2016](#)). De hecho, los inicios del movimiento se centraron en facilitar los datos, pero a menudo no tenían en cuenta quiénes eran sus usuarios potenciales y para qué los necesitarían ([Frank et al., 2016](#)). Cerrar esta brecha requiere una mejor gestión de las plataformas de datos en un extremo y el desarrollo de competencias de datos entre los usuarios en el otro. Este artículo explora esto último: a fin de obtener información valiosa a partir de los datos, ya sea como profesionales, ciudadanos, estudiantes o particulares, debemos adquirir nuevas habilidades y a aprender a "hablar datos". [Panetta \(2019\)](#) ilustra la necesidad de la alfabetización de datos con una organización imaginaria, "en la que el departamento de marketing habla en francés, los diseñadores de productos hablan en alemán, el equipo de análisis habla en español y nadie habla un segundo idioma" (párrafo 1), haciendo que la comunicación sea imposible, si bien la organización da una gran importancia a la tecnología digital, los datos y el análisis. Es fundamental que las personas adquieran capacidades relacionadas con los datos para impulsar el negocio y apoyar la colaboración entre disciplinas y departamentos. Como expresa [Crofts \(2018\)](#), "la alfabetización de datos será el idioma universal de la cuarta revolución industrial" (párrafo 1). Así pues, nuestra intención es definir cuáles son las competencias implicadas en la alfabetización de datos y cómo se pueden desarrollar.

La cuestión de la alfabetización de datos ha ganado importancia durante los últimos años. Las empresas quieren tomar decisiones “basadas en datos”, se exige a los dirigentes políticos que tomen decisiones y hagan leyes utilizando o considerando datos, y las personas aspiran al “yo cuantificado” (Rosenbaum, 2015) con la ayuda de smartphones y dispositivos wearable. Bowne-Anderson (2018) señala tres motivos por los que esta es una competencia importante para el siglo XXI: los cambios producidos en distintos sectores derivados de la ciencia de datos y la IA, la prevalencia de los datos y modelos como forma de comunicación en los medios de comunicación, y el impacto que nuestros datos personales tienen en nuestra interacción con el mundo. Destaca la falta de alfabetización de datos con el ejemplo de un modelo probabilístico publicado por una página web de periodismo de datos llamada FiveThirtyEight durante las elecciones presidenciales de 2016 en los Estados Unidos, que establecía las probabilidades de ganar de cada candidato: una probabilidad del 71,4 % para Clinton frente a un 28,6 % para Trump. Cuando ganó este último, Downey (2016) señaló que este resultado no nos debería sorprender puesto que un 28,6 % de hecho es más probable que echar una moneda al aire dos veces y sacar cara dos veces; es decir, no es un escenario imposible (Bowne-Anderson, 2018; Downey, 2016). Los modelos predictivos de este tipo exigen un análisis crítico y un entendimiento fundamental de los datos y las estadísticas para realizar interpretaciones inteligentes. Ciertamente, la forma en la que leemos e interpretamos los datos, las estadísticas o la información pueden tener efectos significativos: será más o menos probable que los expertos jurídicos opten por un veredicto de culpabilidad en función de cómo se les presentan las estadísticas sobre la precisión de las pruebas de ADN, mientras que el diagnóstico de un experto médico

puede cambiar dependiendo de cómo se enmarca la fiabilidad del cribaje del cáncer de mama (Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2000). Estas decisiones pueden tener consecuencias profundas para nuestras vidas y las de los demás. No obstante, “la mayoría de personas siguen confiando en un pequeño número de intermediarios tales como aplicaciones especializadas, periodistas de datos, grupos de presión y partidos políticos para que les seleccionen e interpreten los datos” (Frank et al., 2016, pág. 5). A fin de tomar decisiones informadas mejores, reconocer los sesgos y extraer interpretaciones bien fundadas, aumentar nuestro nivel de alfabetización de datos es crucial.

El primer apartado de este artículo revisa la definición de alfabetización de datos y sus competencias relacionadas en la bibliografía. A continuación exploraremos el valor de la alfabetización de datos desde la perspectiva de las empresas, las sociedades y las comunidades, la educación y los particulares. El último apartado analizará las consideraciones necesarias para desarrollar habilidades de alfabetización de datos entre la población.

### Definición de la alfabetización de datos

El advenimiento de la tecnología ha traído nuevas “alfabetizaciones”, como la alfabetización numérica o informática, referida como tal en comparación con la capacidad de leer, comprender y usar textos (Frank et al., 2016). El creciente interés por la alfabetización de datos es, por lo tanto, una consecuencia natural del incremento de los datos en toda la sociedad. Sin embargo, antes de analizar la importancia de la alfabetización de datos, es necesario establecer los límites de este término. La Tabla 1 presenta una selección de sus definiciones.

Tabla 1  
Definiciones de alfabetización de datos

Fuente	Definición de alfabetización de datos
(Jacob Carlson, Fosmire, Miller, & Sapp Nelson, 2011, pág. 633)	“Por lo general, la alfabetización de datos implica comprender qué significan los datos, lo cual incluye cómo leer gráficos y cuadros de manera adecuada, sacar conclusiones correctas de los datos y reconocer cuándo los datos se están utilizando de manera engañosa o inapropiada”.
(Hogenboom, Holler Phillips, & Hensley, 2011, pág. 410)	“La alfabetización de datos es la capacidad de leer e interpretar datos, utilizar el pensamiento crítico ante las estadísticas y usar las estadísticas como evidencia”.
(Calzada Prado & Marzal, 2013, pág. 124)	“Conjunto de competencias relacionadas con la adquisición, evaluación, manejo, análisis e interpretación de datos que se encuentran fuera del alcance de las competencias estadísticas”.
(McAuley, Rahemtulla, Goulding, & Souch, 2010, pág. 89)	“La capacidad de identificar, recuperar, evaluar y usar información para formular y responder preguntas significativas”.
(Ridsdale et al., 2015, pág. 2)	“La alfabetización de datos es la capacidad de recopilar, gestionar, evaluar y aplicar los datos de manera crítica”.
(Crusoe, 2016, pág. 38)	“La alfabetización de datos es el conocimiento de qué son los datos, cómo se recopilan, analizan, visualizan y comparten, y es el entendimiento de cómo se aplican los datos en beneficio o en perjuicio, dentro del contexto cultural de seguridad y privacidad”.
(Frank et al., 2016, pág. 5)	“La alfabetización de datos se refiere a la capacidad de comprender y usar los datos, particularmente en el contexto de internet”.
(Wolff, Gooch, Cavero Montaner, Rashid, & Kortuem, 2016, pág. 23)	“La alfabetización de datos es la capacidad de formular y responder preguntas del mundo real a partir de conjuntos de datos grandes y pequeños a través de un proceso de consulta, considerando el uso ético de los datos. Se basa en habilidades básicas prácticas y creativas, con la capacidad de ampliar el conocimiento de las habilidades de gestión de datos especializados de acuerdo con los objetivos. Esto incluye las habilidades de seleccionar, limpiar, analizar, visualizar, criticar e interpretar los datos, así como de comunicar historias a partir de los datos y usar los datos como parte de un proceso de diseño”.

(Panetta, 2019, párrafo 5)	"Gartner define la alfabetización de datos como la capacidad de leer, escribir y comunicar datos en contexto, incluyendo el entendimiento de las fuentes y conjuntos de datos, los métodos analíticos y las técnicas aplicadas, y la capacidad de describir el caso de uso, la aplicación y el valor resultante".
(Pedersen & Caviglia, 2019, pág. 1)	"La alfabetización de datos se puede definir como una competencia compuesta que consta de un cierto nivel de competencia en estadística, visualización de datos y competencias más genéricas de resolución de problemas utilizando diferentes datos".

Las definiciones de la Tabla 1 evocan la capacidad de leer, comprender e interpretar datos, o de tener habilidades básicas de gestión de datos, como su recopilación, análisis, gestión o evaluación, como componentes clave de la alfabetización de datos. Sin embargo, las definiciones difieren en los detalles que rodean estas competencias básicas. Algunas definiciones dan importancia a los objetivos finales del uso de los datos, la creación de valor (Panetta, 2019) o la resolución de un problema (Pedersen & Caviglia, 2019), y la capacidad de responder a preguntas "significativas" o "del mundo real" (McAuley et al., 2010; Wolff et al., 2016). Otras evocan un contexto, ya sea internet (Frank et al., 2016), el contexto cultural en relación con la privacidad y la seguridad (Crusoe, 2016), o un caso de uso que enmarca las aplicaciones de datos (Panetta, 2019). Consideraciones éticas (Wolff et al., 2016) y el pensamiento crítico o el reconocimiento de la verdad (Jacob Carlson et al., 2011; Hogenboom et al., 2011; Ridsdale et al., 2015) también se presentan como aspectos de la alfabetización de datos. Ciertamente, lo que nos puede llevar de la información al conocimiento es el "sentido crítico para saber filtrar datos de manera discriminatoria" (Álvarez-Arregui, Rodríguez-Martín, Agudo Prado, & Arreguit, 2017, pág. 98). Dichos aspectos extienden la alfabetización de datos desde las habilidades básicas de comprensión y gestión de los datos hasta la adopción de una "cultura de datos" (Crusoe, 2016).

Hogenboom, Holler Phillips y Hensley (2011), así como Pedersen y Caviglia (2019), incluyen la competencia estadística en su definición, contrastando con la definición de Calzada Prado y Marzal (2013), que establece las competencias estadísticas como un elemento separado. Puede resultar difícil definir las líneas que separan los dos conceptos, ya que la alfabetización de datos y la alfabetización estadística están relacionadas e incluso entrelazadas. A partir de los datos podemos crear modelos, calcular probabilidades o extraer estadísticas que requieren competencias estadísticas para su comprensión e interpretación. Shield (2005) propone un punto de vista sobre la relación entre la alfabetización de datos, la alfabetización estadística y la alfabetización informacional que se muestra en la Figura 1.

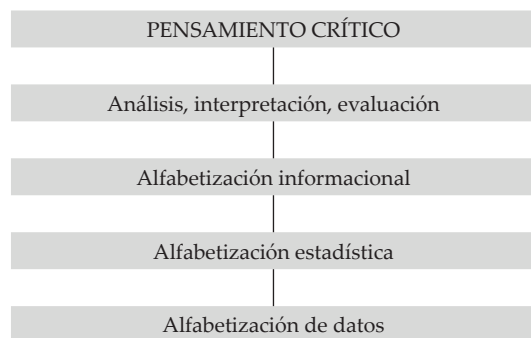


Figura 1. Relación entre la alfabetización de datos, la alfabetización estadística y la alfabetización informacional desde una perspectiva de pensamiento crítico (Shield, 2005, pág. 8)

Las tres formas de alfabetización, junto con las habilidades de análisis, interpretación y evaluación, son aspectos de la capacidad de pensamiento crítico. La alfabetización de datos es crucial para la alfabetización estadística e informacional. Las tres están íntimamente relacionadas y "son más generales que específicas, cada una implica un estudio interdisciplinario y se ocupan de los fundamentos", lo que las hace "útiles para los estudiantes de cualquier especialidad" (Shield, 2005, pág. 9).

La alfabetización de datos también debe diferenciarse de la ciencia de datos con respecto al nivel de competencia: si bien esta última pertenece a un campo de estudio al que los especialistas dedican años de formación, la alfabetización de datos es la capacidad de los no especialistas de aprender sobre la ciencia de datos y aplicar el conocimiento adquirido para resolver problemas (Pedersen & Caviglia, 2019).

Definir la alfabetización de datos y distinguirla de otros términos, como la alfabetización informacional, es útil para comunicar claramente el tema y su importancia (Hunt, 2005). Sin embargo, la búsqueda de una definición no debe ir en detrimento de la flexibilidad, ya que establecer límites a un concepto relativamente nuevo "es tan prescriptivo como descriptivo" (Frank et al., 2016, pág. 6). Las definiciones de alfabetización de datos varían según el contexto en el que se formularon (Wolff et al., 2016). Es natural que la profundidad de la definición y las competencias consideradas se ajustaran a las necesidades específicas de un individuo o grupo, ya que los diferentes campos y problemas requieren distintos grados de alfabetización de datos. Sin embargo, lo que es seguro es que todos los campos de nuestra sociedad de la información requerirán que las personas adquieran cierta capacidad para comprender y utilizar los datos, ya sea en un nivel avanzado o básico.

### La necesidad de la alfabetización de datos

A medida que el mundo está cada vez más interconectado y nuestra sociedad de la información recopila cantidades crecientes de datos, la alfabetización de datos deviene tan útil y esencial como la alfabetización textual, brindando mayores oportunidades a quienes la adquieran y respaldando el progreso de las sociedades (Frank et al., 2016). A menudo se dice que "los datos son el nuevo petróleo" (una frase generalmente atribuida a Clive Humby, quien la pronunció en 2006 (Haupt, 2018)), un recurso que se puede extraer, refinar y utilizar como combustible para la innovación. Sin embargo, esta metáfora ha sido criticada por fallar en varios aspectos, como por ejemplo la comparación entre el recurso finito del petróleo y el flujo interminable y reutilizable de los datos (Marr, 2018) (véase García Martínez, 2019; Haupt, 2018; James, 2019; van Zeeland, 2019 para otros ejemplos de deconstrucción de este mantra frecuentemente utilizado). En el año 2010, durante una charla sobre visualización de datos, David McCandless sugirió una metáfora diferente, imaginando que los datos eran el nuevo suelo, un "medio fértil y creativo" (McCandless, 2010). A través de esta lente, los datos son un recurso valioso en el que se puede trabajar para desarrollar nuevas ideas, perspectivas, proyectos y soluciones. ¿Pero qué beneficios con-

cretos pueden aportar los datos, y por lo tanto la alfabetización de datos, a las personas y a la sociedad? En este apartado consideraremos la necesidad de la alfabetización de datos desde la perspectiva de las empresas, las sociedades y las comunidades, la educación y las personas.

### Empresas

Los datos tienen el potencial de dar a las empresas una ventaja competitiva ([Andrejevic, 2014](#); [Morrow, 2018](#)). Las nuevas tecnologías y aplicaciones como el Internet de las cosas, la robótica o el aprendizaje automático, que han echado raíces en empresas de todos los sectores, utilizan y se benefician de los datos que pueden ayudar a mejorarlas, de la misma forma que pueden mejorar la eficiencia de sus operaciones y procesos ([Reinsel et al., 2019](#)). Los datos pueden respaldar la toma de decisiones y favorecer la rendición de cuentas, y pueden usarse para identificar cambios y sus causas ([Wu, Chen, Wu, & Lytras, 2018](#)). Desde los datos hasta la información y el conocimiento, los conocimientos extraídos de los datos permiten a las empresas tomar decisiones informadas e implementar soluciones basadas en hechos y no solo en el instinto, permitiéndoles abordar las ineficiencias y aumentar la sostenibilidad. Según un estudio realizado por [Harvard Business Review Analytic Services \(2012\)](#), “más del 70 por ciento de las organizaciones que habían implementado análisis en todas sus organizaciones informaron sobre un mejor rendimiento financiero, una mayor productividad, menores riesgos y una toma de decisiones más rápida” (pág. 2).

Las empresas, así como las organizaciones sin ánimo de lucro u organizaciones de otro tipo, pueden beneficiarse del uso de los datos. Para sacar provecho de este potencial, es necesario fomentar la alfabetización de datos en todas las organizaciones. Debe convertirse en un segundo idioma en el que los empleados de todos los departamentos puedan comunicarse entre sí a medida que los datos crecen y se convierten en una parte fundamental de su trabajo ([Panetta, 2019](#)). Además, mientras van apareciendo nuevos puestos dedicados al análisis de datos y son ocupados por especialistas, es necesario que estos expertos en datos adquieran experiencia específica de su dominio para hacer el mejor uso de los datos, formular las preguntas correctas y proporcionar información relevante para el contexto ([Andrejevic, 2014](#)). Para conseguirlo, el equipo de datos de una empresa debe compartir puntos en común con el resto de la organización, comunicándose con los miembros de otros departamentos para identificar sus necesidades. Si el personal sabe leer y escribir datos, se facilita la identificación de necesidades y preguntas y la comunicación de resultados y perspectivas. Además, los expertos de la organización que no dominan los datos deben tener la habilidad de evaluar críticamente los hallazgos proporcionados por los expertos en datos ([Wolff et al., 2016](#)). Esto es esencial para garantizar que los conocimientos se entiendan adecuadamente, se pongan en contexto y se conviertan en acciones apropiadas en respuesta a las necesidades de la organización. La alfabetización de datos también capacita a las personas para evaluar la confiabilidad de la información y reconocer cuándo están siendo engañadas deliberadamente por informes y cifras que pueden parecer válidos a primera vista ([Pedersen & Caviglia, 2019](#); [Wolff et al., 2016](#)), analizándolos de manera crítica y formulando las preguntas correctas.

Otro aspecto esencial de la alfabetización de datos para las empresas es la cuestión de la responsabilidad social. A medida que los datos aumentan en importancia en todas las organizaciones, la alfabetización de datos debe incluir cuestiones de ética ([Chakrabarti et al., 2009](#); [Frank et al., 2016](#); [Ridsdale et al., 2015](#);

[Wolff et al., 2016](#)) o seguridad y privacidad ([Andrejevic, 2014](#); [Crusoe, 2016](#); [Ridsdale et al., 2015](#)). Las empresas deben comprender y adaptarse a la nueva legislación en materia de datos, manejar los datos de manera ética, respetar la privacidad de sus clientes y otras partes interesadas y garantizar la seguridad de los sistemas para evitar fugas de datos. Comprender estos aspectos es crucial para todos los empleados que entran en contacto con datos en su trabajo, ya sea en el departamento jurídico, de marketing, de ventas o en cualquier otro departamento.

A medida que los mercados se transforman a un ritmo acelerado, empujando a las empresas a innovar, mejorar y adaptarse, los datos tienen el potencial de ser una herramienta inestimable. Fomentar la alfabetización de datos y formar a los empleados para que puedan comprender, gestionar y comunicarse con los datos es esencial para que las empresas crezcan de manera eficiente y sostenible en la sociedad de la información.

### Sociedades y comunidades

Los beneficios de la alfabetización de datos para los gobiernos y las organizaciones públicas son un reflejo de los de las empresas: mayor eficiencia y sostenibilidad, toma de decisiones basadas en hechos, un lenguaje universal para mejorar la comunicación entre los departamentos y con los ciudadanos, así como una mayor conciencia y comprensión de consideraciones éticas, de seguridad y de privacidad. Además, las aplicaciones exitosas de datos podrían respaldar la seguridad nacional y la ventaja militar, mejorar la salud, la eficiencia, la capacidad de respuesta y la conveniencia en todas las sociedades, así como una mayor transparencia y rendición de cuentas para las entidades públicas ([Andrejevic, 2014](#)). La tendencia de los proyectos de ciudades inteligentes ha visto aparecer tecnologías de recogida de datos y aplicaciones relacionadas en todos los espacios públicos, con ejemplos como la ciudad de Santander en España, con 18.000 sensores para 180.000 residentes en 2015 ([Treboux, Jara, Dufour, & Genoud, 2015](#)), el proyecto Virtual Singapore dedicado a la creación de un mapa en 3D de la ciudad utilizando sus datos de construcción, terreno y medio ambiente ([Chia, 2016](#)), o el centro de operaciones de Río de Janeiro, construido en asociación con IBM y que obtiene datos de 30 agencias ([Kitchin, 2014](#)). La colaboración entre sectores es necesaria para garantizar que las aplicaciones de los datos tengan éxito en el contexto urbano, implementar sistemas interoperativos y proporcionar acceso a conjuntos de datos ([Bertot, Ursula, Jaeger, Sarin, & Choi, 2014](#)). Por lo tanto, de nuevo la alfabetización de datos puede servir como idioma universal entre sectores. Los datos pueden apoyar la planificación urbanística y el desarrollo de políticas ([Lim, Kim, & Maglio, 2018](#)), ya que las entidades públicas deben extraer valor de los datos de manera eficiente, equitativa y sostenible para hacer frente a los desafíos sociales o al cambio climático ([Batty et al., 2012](#)). Los funcionarios públicos alfabetizados en datos pueden comprender mejor el pasado, gestionar el presente y prepararse para el futuro extrayendo información procesable de los datos.

Cuando deben tomar decisiones e implementar leyes y políticas relacionadas con la gestión de datos, los datos abiertos, la seguridad y la privacidad, o tomar decisiones utilizando los datos de los ciudadanos, los funcionarios públicos deben comprender el fondo de los problemas. Es fundamental que nuestros líderes y administradores políticos se familiaricen con los datos, comprendan los datos en sí mismos y la ética que los rodea. Los resultados obtenidos a través del análisis de datos podrían mostrar patrones que conducirían a decisiones poco éticas o incluso ilegales, discriminando a grupos específicos ([Chakrabarti et al., 2009](#)). Las predicciones proporcionadas por la minería de datos

son probabilísticas ([Andrejevic, 2014](#)) y deben entenderse como tales, lo que significa que los responsables de las decisiones deben comprender las limitaciones y posibles dificultades de los datos y los análisis y no aplicar a ciegas las acciones que se alinean con el resultado más probable. Como se ha señalado anteriormente, independientemente de lo improbable que sea que se vaya a producir un acontecimiento según lo indicado por los datos, esto no hace que dicho acontecimiento sea imposible. La interpretación crítica y la contextualización de los resultados, así como la transparencia, son esenciales para comprender las implicaciones y consecuencias de las decisiones. Las entidades públicas también deben abogar por tener unos ciudadanos alfabetizados en datos. Según [Andrejevic \(2014\)](#), existe una “gran brecha de datos”, un desequilibrio de poder “entre los que recopilan, almacenan y extraen grandes cantidades de datos, y aquellos a quienes se dirige la recogida de los datos” (pág. 1). Los datos pueden fomentar un público informado y participativo, ([Al Nuaimi, Al Neyadi, Mohamed, & Al-Jaroodi, 2015](#)), y pueden usarse como una herramienta de comunicación entre los gobiernos y sus ciudadanos, lo que significa que la alfabetización de datos es una “habilidad cívica importante” ([McAuley et al., 2010, pág. 89](#)), así como “una plataforma para una democracia efectiva” ([Frank et al., 2016, pág. 5](#)).

### Educación

[Pedersen y Caviglia \(2019\)](#) opinan que “la alfabetización de datos se sugiere como un elemento central en la educación que prepara a todos los jóvenes para convertirse en ciudadanos de una sociedad de la información” (pág. 1), pero según un estudio solo el 21% de las personas entre 16 y 24 años sabían leer y escribir datos ([Olik, 2018](#)). Como se ha señalado anteriormente, están apareciendo nuevas oportunidades y carreras en relación con los datos, si bien las del futuro todavía no se han inventado ([Morrow, 2019](#)), y hay escasez de científicos de datos ([Davenport & Patil, 2012](#)). Es crucial preparar a los estudiantes, la fuerza laboral del mañana, para que sepan recopilar, gestionar, interpretar y aplicar los datos de forma crítica, tareas que formarán parte de su trabajo en la mayoría de sectores ([Bowne-Anderson, 2018; Ridsdale et al., 2015](#)). La educación superior tiene un papel importante que desempeñar en el uso innovador de la creciente cantidad de datos de la sociedad ([McAuley et al., 2010](#)). La alfabetización de datos respalda el proceso de aprendizaje del conocimiento específico del dominio, es una habilidad valiosa tanto a nivel personal como profesional, y se alinea con el “objetivo general de obtener graduados bien equipados, de pensamiento crítico y competitivos a nivel mundial” ([Ridsdale et al., 2015, pág. 2](#)). Las instituciones educativas deben aceptar y adaptarse a los cambios de la sociedad, desarrollando planes de estudio que cumplan los nuevos requisitos ([Álvarez-Arregui et al., 2017](#)). Además, para garantizar que los estudiantes inicien sus estudios superiores preparados para la integración de nuevas habilidades relacionadas con los datos además de los conocimientos de su dominio específico, es necesario comenzar temprano enseñando los fundamentos de la alfabetización de datos en las escuelas ([Wolff et al., 2016](#)), ampliando las habilidades básicas de alfabetización estadística que puedan encontrarse ya en los planes de estudios ([Frank et al., 2016](#)). Como explican [Pedersen y Caviglia \(2019\)](#), “la reforma educativa y la innovación tienen un papel clave a la hora de llevar la alfabetización de datos a la población con estudios de todas las disciplinas y planes de estudio” (pág. 3).

Desde un punto de vista interno, la alfabetización de datos también es inestimable para impulsar la investigación y el ámbito académico. Los datos pueden ser una fuente de investigación

o su resultado, y se puede compartir y reutilizar entre distintos proyectos, lo que significa que es necesario contar con habilidades de trabajo con los datos y estándares en este sentido ([Shorish, 2015](#)). Si bien se espera que los estudiantes de posgrado tengan la capacidad de gestionar datos al entrar a trabajar en un laboratorio, una tarea que deberán asumir con distintos grados de responsabilidad, a menudo no están lo suficientemente preparados ([Jacob Carlson et al., 2011; Shorish, 2015](#)). Por lo tanto, deben aprender a medida que avanzan, pero cada estudiante puede aprender de manera diferente en ausencia de unas prácticas estandarizadas, lo cual dificulta la interoperabilidad y la preservación de los datos en el futuro ([Jacob Carlson et al., 2011](#)). Los estudiantes de posgrado a menudo son cruciales en el proceso de investigación, llevan a cabo muchas tareas relacionadas con los datos, pero puede suceder que no comprendan las prácticas adecuadas de documentación y almacenamiento, cuestiones relativas a la propiedad o el valor potencial de los datos a largo plazo. ([Jake Carlson, Nelson, Johnson, & Koshoffer, 2015](#)). De hecho, a medida que la investigación evoluciona, “comprender que los datos tienen el potencial de influir no solo en la propia investigación sino también en el trabajo de otros, en campos que pueden parecer no relacionados, puede ayudar a crear una conciencia del sistema plural de becas” ([Shorish, 2015, pág. 11](#)). Esto significa que capacitar a los estudiantes, así como al profesorado, de manera holística e implementar estándares claros de gestión de datos es crucial para eliminar obstáculos para el avance de la investigación.

### Personas

Bajo el prisma de las empresas, la sociedad y la educación, hemos explorado la importancia de la alfabetización de datos para las personas, ya sea como empleados, ciudadanos, estudiantes o académicos. Brinda acceso a nuevas oportunidades de negocio, permite a las personas desempeñarse mejor en su trabajo, respalda sus decisiones con hechos e impulsa la innovación de su organización, les permite comunicarse y comprender mejor a la administración pública y los gobiernos, participar como ciudadanos activos, extraer opiniones e ideas informadas, y acelerar la investigación científica compartiendo datos e ideas en todos los campos y mejorando y adaptando continuamente la educación a las necesidades de la sociedad.

La alfabetización de datos también tiene una gran importancia a nivel personal. Las personas entran en contacto con los datos cada vez más a menudo, usándolos como base para tomar decisiones como qué comprar, a quién votar o dónde vivir, y para elegir cómo se pueden usar sus propios datos ([Wolff et al., 2016](#)). [Frank et al. \(2016\)](#) argumentan que la capacidad de usar datos y obtener información a partir de ellos “se está convirtiendo rápidamente en un requisito para participar en la vida moderna, tan fundamental como la capacidad de usar un teléfono o dinero” (pág. 5), dejando en desventaja a quienes no poseen estas habilidades. Como se ha señalado anteriormente, hay una brecha entre quienes recopilan y utilizan los datos y quienes son los titulares de estos datos, dado que quienes tienen grandes cantidades de datos de muchas personas pueden extraer mucha más información que una persona con acceso solamente a su propio punto de datos único ([Andrejevic, 2014](#)). Los datos personales se pueden usar para determinar decisiones tales como el destinatario de un préstamo o una beca, lo que puede dar lugar a una discriminación poco ética ([Chakrabarti et al., 2009](#)). Las personas deben comprender cómo se utilizan sus datos en su favor o en su contra, y cómo esto afecta su vida y a las opciones que tienen a su disposición. También deben comprender las implicaciones para

la privacidad y seguridad que plantea el hecho de compartir sus datos personales con el fin de poder tomar decisiones informadas sobre su uso (Crusoe, 2016; Ridsdale et al., 2015). Esto exige transparencia por parte de los recopiladores y administradores de los datos, nuevas leyes y políticas que regulen el uso de datos y habilidades de alfabetización de datos en toda la población.

En una era tecnológica con acceso constante a las historias de los medios de comunicación, una afluencia interminable de información y la posibilidad de que surjan cámaras de eco o filtros burbuja, (Flaxman, Goel, & Rao, 2013), las personas deben adquirir las habilidades para evaluar críticamente la información que se les presenta, así como la capacidad de buscar, seleccionar, analizar e interpretar datos relevantes para sus preguntas. Además, deben ser capaces de evaluar la validez de los datos, cuestionar los métodos de muestreo e identificar posibles sesgos para llegar a conclusiones sólidas. Los datos o las interpretaciones sesgadas, que pueden generar conocimientos incorrectos y malas decisiones, pueden surgir inadvertidamente a partir de una escasa alfabetización de datos o ser introducidas voluntariamente con la intención de manipular a los lectores (Wolff et al., 2016). Por lo tanto, estas habilidades son esenciales para tomar decisiones basadas en datos en todos los aspectos de la vida. Además, en un mundo en constante cambio, las personas deben aprender a buscar e integrar nuevos datos, adaptando sus opiniones y acciones de acuerdo con los cambios situacionales y ambientales. Esto se alinea con la necesidad de aprendizaje continuo y adaptación en todos los campos (Álvarez-Arregui et al., 2017), ya que los sectores se ven alterados por las innovaciones y el mundo cambia a un ritmo acelerado.

### Fomento de la alfabetización de datos

Por lo tanto, está claro que la alfabetización de datos es esencial para cualquier persona o grupo que entre en contacto con los datos, que ahora son un aspecto omnipresente de la sociedad. ¿Cómo podemos fomentar la alfabetización de datos en toda la población y aprovechar las oportunidades que ofrece? Hay varias consideraciones que son cruciales y que vamos a analizar en este apartado.

#### *Fijación de objetivos en relación con el contexto*

Alcanzar la alfabetización de datos no implica convertirse en un científico de datos (Pedersen & Caviglia, 2019). Implica comprender qué son los datos y tener la capacidad de recopilarlos, analizarlos, visualizarlos y compartirlos, así como desarrollar una conciencia sobre las cuestiones éticas, incluyendo los aspectos de seguridad y privacidad, relacionadas con el uso de datos (Crusoe, 2016; Wolff et al., 2016). Se requieren diferentes niveles de competencia para diferentes situaciones (Frank et al., 2016) y las personas pueden alcanzar la alfabetización pertinente para un rol específico, como comunicador, lector, creador o científico de datos (Wolff et al., 2016). Es necesario reconocer las habilidades relevantes para un contexto dado y para las necesidades de una persona u organización. Los fundamentos de la alfabetización de datos pueden ser comunes en todos los campos, pero diferentes dominios requerirán diferentes conjuntos de habilidades. Por ejemplo, la creación de visualizaciones y la capacidad de contar historias de datos son útiles para aquellos que trabajan en marketing, mientras que el almacenamiento, la documentación y la preservación de los datos considerando su posible uso en el futuro son conceptos que deben dominar las personas del ámbito académico. Algunos necesitarán una sólida capacidad de gestionar los datos y analizarlos, algunos pueden centrarse en

usarlos para comunicarse, mientras que otros verán cubiertas sus necesidades con una conciencia y comprensión básica de los datos, incluida la capacidad de evaluar críticamente las interpretaciones proporcionadas por otros, convirtiéndolas en decisiones y acciones apropiadas. Esto demuestra la necesidad de planes de estudios o programas de formación escalables y flexibles, adaptando los objetivos a los niveles requeridos de alfabetización de datos (Schneider, 2013). Para motivar a los estudiantes a hacer frente al desafío de desarrollar su alfabetización de datos, se deben describir claramente los beneficios potenciales en su contexto y la relevancia para sus necesidades (Ridsdale et al., 2015). Además, dado que los alumnos deben poner en práctica la teoría para integrar completamente nuevas habilidades de datos y la disponibilidad de datos abiertos continúa aumentando, es posible desarrollar habilidades con la práctica utilizando datos del mundo real. Esto permite que estudiantes de todos los orígenes puedan relacionarse con la comunidad más amplia, comprender mejor el impacto potencial de sus contribuciones y generar soluciones innovadoras para problemas del mundo real (Ridsdale et al., 2015).

#### *Aprendizaje continuo*

Para desarrollar verdaderamente una cultura de alfabetización de datos, su importancia como habilidad cívica debe ser entendida y reconocida en toda la sociedad, tanto por las instituciones educativas como por los gobiernos (Ridsdale et al., 2015). La enseñanza de sus fundamentos debe comenzar en la escuela (Ridsdale et al., 2015; Wolff et al., 2016), debe enseñarse en todos los niveles (Calzada Prado & Marzal, 2013) y como una asignatura transversal (Calzada Prado & Marzal, 2013; Ridsdale et al., 2015; Vahey et al., 2012; Wolff et al., 2016). Anteriormente hemos destacado el papel clave de la educación superior en el desarrollo de una fuerza laboral alfabetizada en datos y el valor que el ámbito académico puede obtener con una mayor alfabetización de datos. A medida que la sociedad, los procesos y las tecnologías cambian a un ritmo rápido, también será esencial que nunca se deje de aprender. Las personas necesitarán adaptarse a nuevas formas de datos e integrar nuevas posibilidades en su experiencia. Esto significa que la adquisición de nuevas habilidades de datos no debe detenerse al final de la formación de las personas; las organizaciones que deseen obtener valor a partir de los datos deberán capacitar a sus miembros para mantenerse al día de las innovaciones y las nuevas oportunidades. Los objetivos y las habilidades que constituyen la alfabetización de datos deberán revisarse para reflejar los cambios sociales y ambientales.

#### *Una competencia de la comunidad*

Como resultado de esta necesidad de adaptación y aprendizaje constantes, el desarrollo de la alfabetización de datos en toda la población requerirá la colaboración y la combinación de recursos entre las partes interesadas. Las organizaciones, las instituciones educativas, los gobiernos y los individuos tienen sus intereses y todos tienen un papel que desempeñar para inculcar una cultura de datos. La mejor forma de conseguirlo de manera coherente y holística es a través de la comunicación y la colaboración, por ejemplo entre universidades y empresas, para garantizar que los graduados estén preparados para entrar en el mundo laboral y puedan satisfacer los requisitos de las empresas, o entre educadores para eliminar lagunas o superposiciones en los planes de estudio (Ridsdale et al., 2015). Los profesores de habilidades relacionadas con los datos pueden no ser especialistas en datos, pero pueden enseñarlos en el contexto del dominio

específico de su experiencia, por lo que será necesario el apoyo, la colaboración y el aprendizaje entre pares para proporcionar una gama completa de habilidades de datos a los estudiantes ([Pedersen & Caviglia, 2019](#)).

Como se ha comentado anteriormente, la alfabetización de datos es un concepto flexible que depende del contexto y las necesidades específicas de una persona u organización, y que evoluciona de acuerdo con los cambios sociales y tecnológicos. Por lo tanto, las personas alcanzarán distintos grados de alfabetización de datos y se especializarán en ciertos aspectos según sus necesidades. Cada persona tendrá sus propias habilidades de datos para contribuir, “de forma que la presencia de algunas personas que puedan encontrar datos, algunas personas que puedan manipularlos y algunas que puedan presentar el resultado pueden constituir la alfabetización de datos para esa comunidad” ([Frank et al., 2016, pág. 6](#)). De esta manera, la alfabetización de datos es una competencia colectiva en la que cada miembro de la comunidad aporta algo diferente y que debe enseñarse de manera colaborativa e interdisciplinaria, uniendo competencias para fomentarla en comunidades heterogéneas ([Pedersen & Caviglia, 2019](#)). Reunir a personas de muchas disciplinas también tiene el enorme potencial de extraer más valor de los datos. Dado que los conjuntos de datos se pueden usar, analizar e interpretar de muchas maneras diferentes, lo que lleva a una variedad de conclusiones diferentes, es importante crear un debate en torno a los datos e incluir las ideas, la experiencia y las perspectivas de personas de diversos orígenes y campos. [Silberzahn y Uhlmann \(2015\)](#) pidieron a 29 equipos que estudiaran el mismo conjunto de datos, buscando la respuesta a la misma pregunta, y descubrieron que sus conclusiones variaban según los métodos analíticos elegidos. Argumentan que una investigación crowdsourcing de este tipo puede atenuar el efecto de los “resultados llamativos” y “da espacio a opiniones disidentes” (pág. 191). Además, crear debates sobre los datos y dar voz a personas de distintas disciplinas nos permite acceder a la “base de datos mental” ([see Forrester, 2009, pág. 12](#)), completando nuestros conjuntos de datos digitales con esos datos que no pueden ser capturados y que existen en la mente de las personas: ideas, experiencias, comprensión de las sutilezas culturales y contextuales, y así sucesivamente. Es crucial incluir estos datos para llegar a conclusiones sólidas y tomar decisiones informadas. El uso de los datos como un punto común de debate, intercambiando perspectivas, ideas y opiniones a través del lenguaje compartido y universal de la alfabetización de datos, ofrece muchas posibilidades de progreso y la oportunidad de encontrar e imaginar mejores soluciones a los desafíos de hoy y de mañana.

## Conclusión

Nuestra revisión de varias definiciones de alfabetización de datos en la bibliografía ha mostrado diferentes interpretaciones del concepto: hace referencia a distintas habilidades, como

la comprensión, el uso y la comunicación de datos, junto con la capacidad de buscar soluciones y resolver problemas con el apoyo de los datos, con relevancia para un contexto dado y con la conciencia de consideraciones éticas como la privacidad o la seguridad. Las variadas perspectivas que rodean el concepto convergen en un punto: la alfabetización de datos es una habilidad crucial para hacer frente a los desafíos y aprovechar las oportunidades de una sociedad de la información con tecnologías ubicuas de recopilación de datos. Las habilidades relacionadas con los datos pueden ayudar a las empresas a obtener una ventaja competitiva, así como aumentar la eficiencia y la sostenibilidad de sus operaciones; son una habilidad cívica importante para que los ciudadanos participen en la configuración de sus ciudades, sus países y el mundo; son esenciales para los administradores, los responsables de la toma de decisiones o legisladores que están llamados a desarrollar políticas relacionadas con la tecnología, la seguridad y la privacidad, o que usan datos para tomar decisiones sobre otros temas; pueden impulsar la investigación académica y son competencias cruciales que la educación debe impartir a la futura fuerza laboral; y son inestimables para que todas las personas, en todos los campos, puedan tomar decisiones informadas y basadas en hechos a lo largo de la vida.

Fomentar la alfabetización de datos en todos los campos requerirá un reconocimiento de su valor en todos los niveles de la sociedad. Los objetivos para cada individuo u organización deben ser relevantes para sus necesidades y su contexto, y no deben ser inamovibles: a medida que cambien los entornos y la sociedad, tendremos que adaptarnos y aprender nuevas habilidades que reflejen el mundo en el que vivimos. Esto significa que la alfabetización de datos debe inscribirse en un proceso de aprendizaje continuo. Además, es necesario tratar la alfabetización de datos como la competencia interdisciplinaria que es: su valor se extiende a todos los ámbitos y puede servir como un lenguaje universal para conversar entre departamentos, sectores, contextos y disciplinas. Las diferentes perspectivas y métodos extraen diferentes conclusiones de los datos, por lo que es esencial colaborar y debatir sobre sus interpretaciones para llegar a conclusiones y decisiones acertadas. También es una competencia de la comunidad, lo que significa que una comunidad en la que cada miembro contribuye con sus propias habilidades individuales puede lograr un alto grado de alfabetización de datos. Por lo tanto, la colaboración entre disciplinas nos ayudará a hacer el mejor uso de las grandes cantidades de datos que tenemos a nuestra disposición.

Los datos no tomarán las decisiones por nosotros (y no deben hacerlo); somos nosotros quienes debemos extraer las informaciones, ideas, interpretarlas críticamente y convertirlas en decisiones y acciones apropiadas. La tecnología y los datos son herramientas inestimables con el potencial de hacer avanzar a nuestra sociedad. Debemos aprender a entenderlos, así como a utilizarlos de forma efectiva y responsable.







## Fostering Data Literacy Across Fields

Stephanie Arreguit O'Neill (\*)

(\*) Innobridge Services Sàrl, a consulting company based in Switzerland. For more information, please contact [info@innobridge.com](mailto:info@innobridge.com).

Mailing address: Innobridge Services Sàrl, EPFL Innovation Park, Bât. A, 1015 Lausanne, Switzerland

### ABSTRACT

As technologies proliferate throughout society, the world becomes more interconnected and large amounts of data are collected, holding valuable information and knowledge. This increasing amount of data has been recognized as a resource with great potential to increase the efficiency and sustainability of organizations, cities and societies. New jobs dedicated to data are appearing and non-data experts are called to read and handle data in their everyday work. Data literacy is thus an essential skill to move forward in our information society. It is crucial for individuals, businesses, educational institutions, communities and societies to understand the value and opportunities held in the use of data, as well as the ethical, security and privacy issues that may arise from its misuse. This paper addresses the need for data literacy, studies its definition and related skills, considers the need for it from various perspectives, and discusses necessary considerations to foster data literacy across fields.

Keywords: data, data literacy, information society.

### Introduction

Data has become ubiquitous. It has become an essential feature of our everyday lives, whether we analyze data to support decisions in the workplace, search for information when choosing which car to buy, or hear about the latest data breach on the news. The amount of recorded data increases by the minute, as new business models emerge from its opportunities. The total volume of the world's datasphere is predicted to grow from 33 Zo in 2018 to reach 175 Zo by 2025 (Reinsel, Venkatraman, Gantz, & Rydning, 2019). Today, data is a tool used in most professions (Ridsdale et al., 2015). Spreadsheets, key performance indicators or dashboards appear across sectors, driving decisions and supporting the monitoring of teams, projects or processes. Moreover, large amounts of data are publicly available as a result of the open data movement. While open data offers the potential to drive economic and sustainable development, as well as social and political change, with increased engagement, transparency or accountability, it has not met early expectations as issues such as lack of maintenance of data portals or unequal access have arisen (Frank, Walker, Attard, & Tygel, 2016). Indeed, the beginnings of the movement focused on making data available, but often failed to account for who its potential users were and for what they would need to make use of it (Frank et al., 2016). Bridging this gap requires better management of data platforms on one end and the development of data competencies among users on the other. This paper explores the latter: in order to draw valuable insights from data, whether as professionals, citizens, students or private individuals, people must acquire new skills and learn to "speak data". Panetta (2019) illustrates the need for data literacy with an imaginary organization, "where the marketing department speaks French, the product designers speak German, the analytics team speaks Spanish and no one speaks a second language" (para. 1), making communication impossible, even if the organization places importance on digital

technology, data and analytics. It is crucial for people to acquire data-related skills to drive business forward and to support collaboration across disciplines and departments. As expressed by Crofts (2018), "data literacy will be the universal language of the fourth industrial revolution" (para. 1). Thus, we seek to define which competencies are involved in data literacy and how they can be developed.

The question of data literacy has gained importance in recent years. Businesses seek to make "data-driven" decisions, political leaders are called to make decisions and set laws utilizing or regarding data, and individuals reach towards the "quantified self" (see Rosenbaum, 2015) with the help of smartphones and wearable devices. Bowne-Anderson (2018) notes three reasons why it is an important competence for the 21<sup>st</sup> century: the changes across industries resulting from data science and AI, the prevalence of data and models as a means of communication in news media, and the impact our personal data has on our interaction with the world. He points out the lack of data literacy with the example of a probabilistic model published by a data journalism website named Five Thirty-Eight during the 2016 presidential election in the United States, which set the probabilities of either candidate winning: a 71.4% chance for Clinton, against a 28.6% chance for Trump. When the latter occurred, Downey (2016) noted that we shouldn't be surprised by the result, as 28.6% is in fact more likely than flipping a coin twice and getting two heads; in other words, not an impossible scenario (Bowne-Anderson, 2018; Downey, 2016). Predictive models such as this one require critical analysis and a fundamental understanding of data and statistics to make sensible interpretations. Indeed, how we read and interpret data, statistics or information can have significant effects: legal experts are more or less likely to decide on a guilty verdict depending on how statistics about the accuracy of DNA tests is presented to them, while a medical expert's diagnosis can change depending on how the reliability of breast cancer

screening is framed (Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2000). Such decisions can have profound consequences on our lives and those of others. However, “most people still rely on a small number of intermediaries such as specialist applications, data journalists, pressure groups and political parties to select and interpret data on their behalf” (Frank et al., 2016, p. 5). In order to make better, informed decisions, to recognize biases and to extract sound interpretations, increasing our level of data literacy is critical.

The first section of this paper will review the definition of data literacy and its related competencies in the literature. We will then explore the value of data literacy from the perspectives of businesses, of societies and communities, of education, and of private individuals. The final section will discuss the necessary considerations to develop data literacy skills throughout the population.

### Defining Data Literacy

The advent of technology has brought on new “literacies”, for example numerical or IT literacy, referred to as such in a comparison to the ability to read, understand and use text (Frank et al., 2016). The growing interest in data literacy is thus a natural consequence of the rise of data throughout society. Before discussing the importance of data literacy, however, it is necessary to set the boundaries of the term. Table 1 presents a selection of its definitions.

The definitions in Table 1 evoke the ability to read, understand and interpret data, or having basic data handling skills, such as collecting, analyzing, managing or evaluating it,

as key components of data literacy. However, the definitions differ on the details surrounding these core competencies. Some definitions place an importance on the end-goals of using data, achieving the creation of value (Panetta, 2019) or the resolution of a problem (Pedersen & Caviglia, 2019), and the capacity to answer “meaningful” or “real-world” questions (McAuley et al., 2010; Wolff et al., 2016). Others evoke context, whether it is the Internet (Frank et al., 2016), the cultural context in relation to privacy and security (Crusoe, 2016), or a use case framing the data applications (Panetta, 2019). Ethical considerations (Wolff et al., 2016) and critical thinking or truth recognition (Jacob Carlson et al., 2011; Hogenboom et al., 2011; Ridsdale et al., 2015) are also brought forward as aspects of data literacy. Indeed, what can take us from information to knowledge is the “critical sense to filter data discriminately” (Álvarez-Arregui, Rodríguez-Martín, Agudo Prado, & Arreguit, 2017, p. 98). Such aspects extend data literacy from the basic skills of understanding and handling data to the adoption of a “dataculture” (Crusoe, 2016).

Hogenboom, Holler Phillips and Hensley (2011), as well as Pedersen and Caviglia (2019), include statistical competence into their definition, contrasting with the definition provided by Calzada Prado and Marzal (2013), which sets statistical competencies as a separate item. The lines separating the two concepts can be difficult to define, as data literacy and statistical literacy are related and even interlinked. From data we can create models, draw probabilities, extract statistics, which require statistical competencies to understand and interpret. Shield (2005) proposes a viewpoint of the relationship between data literacy, statistical literacy and information literacy shown in Figure 1.

Table 1  
Definitions of Data Literacy

Source	Definition of Data Literacy
(Jacob Carlson, Fosmire, Miller, & Sapp Nelson, 2011, p. 633)	“Typically, data literacy involves understanding what data mean, including how to read graphs and charts appropriately, draw correct conclusions from data, and recognize when data are being used in misleading or inappropriate ways.”
(Hogenboom, Holler Phillips, & Hensley, 2011, p. 410)	“Data literacy is the ability to read and interpret data, to think critically about statistics, and to use statistics as evidence.”
(Calzada Prado & Marzal, 2013, p. 124)	“A suite of data acquisition-, evaluation-, handling-, analysis- and interpretation-related competencies that lie outside the scope of statistical competencies.”
(McAuley, Rahemtulla, Goulding, & Souch, 2010, p. 89)	“The ability to identify, retrieve, evaluate and use information to both ask and answer meaningful questions.”
(Ridsdale et al., 2015, p. 2)	“Data literacy is the ability to collect, manage, evaluate, and apply data, in a critical manner.”
(Crusoe, 2016, p. 38)	“Data literacy is the knowledge of what data are, how they are collected, analyzed, visualized and shared, and is the understanding of how data are applied for benefit or detriment, within the cultural context of security and privacy.”
(Frank et al., 2016, p. 5)	“Data literacy refers to the ability to understand and use data, particularly in the context of the Internet.”
(Wolff, Gooch, Cavero Montaner, Rashid, & Kortuem, 2016, p. 23)	“Data literacy is the ability to ask and answer real-world questions from large and small data sets through an inquiry process, with consideration of ethical use of data. It is based on core practical and creative skills, with the ability to extend knowledge of specialist data handling skills according to goals. These include the abilities to select, clean, analyse, visualize, critique and interpret data, as well as to communicate stories from data and to use data as part of a design process.”
(Panetta, 2019, para. 5)	“Gartner defines data literacy as the ability to read, write and communicate data in context, including an understanding of data sources and constructs, analytical methods and techniques applied – and the ability to describe the use case, application and resulting value.”
(Pedersen & Caviglia, 2019, p. 1)	“Data literacy can be defined as a compound competence consisting of some level of competence in statistics, data visualization and more generic competencies in problem-solving using different data.”

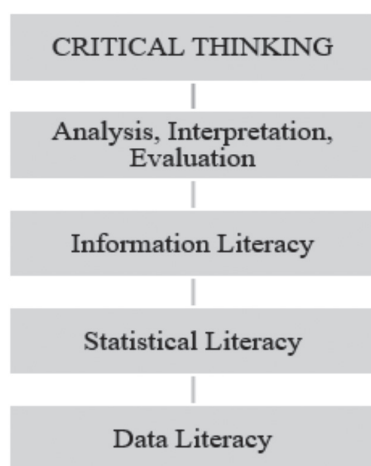


Figure 1. The relation between data literacy, statistical literacy and information literacy from a critical thinking perspective (Schield, 2005, p. 8)

All three forms of literacy, along with the skills of analysis, interpretation and evaluation, are aspects of the capacity for critical thinking. Data literacy is crucial to both statistical and information literacy. The three are closely related and they “are more general than specific, they each involve interdisciplinary study and they deal with fundamentals”, which makes them “useful to students in any major” (Schield, 2005, p. 9).

Data literacy must also be differentiated from data science with regard to the level of competence: while the latter is a field of study to which specialists devote years of training, data literacy is the ability of non-specialists to learn about data science and apply the acquired knowledge to solve problems (Pedersen & Caviglia, 2019).

Defining data literacy and distinguishing it from other terms such as information literacy is useful to communicate clearly on the subject and its importance (Hunt, 2005). However, the search for a definition should not inhibit flexibility, as setting boundaries to a relatively new concept “is as much prescriptive as descriptive” (Frank et al., 2016, p. 6). Definitions of data literacy vary according to the context in which they were set (Wolff et al., 2016). It is natural that the depth of the definition and the competences considered should fit to the specific needs of an individual or group, as different fields and problems will require varying degrees of data literacy. What is certain, however, is that all fields in our information society will require people to acquire some ability to understand and use data, whether at an advanced or basic level.

### The Need for Data Literacy

As the world becomes more interconnected and our information society collects increasing amounts of data, data literacy becomes as useful and essential as textual literacy, providing greater opportunities those who acquire it and supporting the progress of societies (Frank et al., 2016). It is often said that “data is the new oil” (a phrase generally attributed to Clive Humby, who said it in 2006 (Haupt, 2018)), a resource that can be mined, refined and used as fuel for innovation. This metaphor has however been criticized as flawed on several aspects, for example the comparison between the finite resource of oil to the unending and reusable stream of data (Marr, 2018) (see García Martínez, 2019; Haupt, 2018; James, 2019; Van Zeeland, 2019 for other examples of deconstructions of the oft-

used mantra). In a 2010 talk about data visualization, David McCandless suggested a different metaphor, imagining data to be the new soil, a “fertile, creative medium” (McCandless, 2010). Through this lens, data is a valuable resource that can be worked to grow new ideas, insights, projects and solutions. But what concrete benefits can data, subsequently data literacy, bring to people and society? In this section, we will consider the need for data literacy from the perspectives of businesses, societies and communities, education and individuals.

### Businesses

Data holds the potential to give businesses a competitive edge (Andrejevic, 2014; Morrow, 2018). New technologies and applications such as the Internet of Things, robotics or machine learning, which have taken root in businesses across industries, utilize and benefit from the data insights that can help improve them, just as operations and processes can be made more efficient (Reinsel et al., 2019). Data can support decision-making and favor accountability, and it can be used to identify changes and their causes (Wu, Chen, Wu, & Lytras, 2018). From data to information to knowledge, the insights drawn from data enable businesses to make informed decisions and implement solutions based on facts rather than only on instinct, enabling businesses to tackle inefficiencies and increase sustainability. According to a study conducted by Harvard Business Review Analytic Services (2012), “more than 70 percent of the organizations that had deployed analytics throughout their organizations reported improved financial performance, increased productivity, reduced risks, and faster decision-making” (p. 2).

Businesses, as well as non-profit or other organizations, can thus benefit from making use of data. To capitalize on this potential, it is necessary to foster data literacy throughout organizations. It should become a second language for employees in all departments to communicate with each other, as data grows to become a fundamental part of their work (Panetta, 2019). Furthermore, as new positions devoted to data analytics appear and are filled by specialists, it is necessary for these data experts to acquire domain-specific expertise in order to make the best use of the data, ask the right questions and provide insights relevant to the context (Andrejevic, 2014). To achieve this, a company’s data team must share common ground with the rest of the organization, communicating with members of other departments to identify their needs. If the workforce is data literate, the identification of needs and questions and the communication of results and insights are facilitated. Furthermore, non-data experts in the organization should have the skill to critically evaluate the findings provided by data experts (Wolff et al., 2016). This is essential to ensure that the insights are properly understood, placed in context and turned into appropriate action in response to organizational needs. Data literacy also empowers people to assess the reliability of information and recognize when they are being willfully misled by reports and numbers that may seem valid at first glance (Pedersen & Caviglia, 2019; Wolff et al., 2016), by analyzing them in a critical manner and asking the right questions.

Another essential aspect of data literacy for businesses is the question of social responsibility. As data rises in importance across organizations, data literacy must include the questions of ethics (Chakrabarti et al., 2009; Frank et al., 2016; Ridsdale et al., 2015; Wolff et al., 2016) or security and privacy (Andrejevic, 2014; Crusoe, 2016; Ridsdale et al., 2015). Companies must understand and adapt to new data laws, handle data ethically, respect the privacy of their clients and other stakeholders, and ensure the

security of systems to prevent data leaks. Understanding these questions is crucial for all employees who come into contact with data in their work, whether in the legal, marketing, sales, or any other department.

As markets change at an accelerating rate, pushing businesses to innovate, improve and adapt, data has the potential to be an invaluable tool. Fostering data literacy and training employees to understand, manage and communicate with data is essential for companies to grow efficiently and sustainably in the information society.

### *Societies and Communities*

The benefits of data literacy for governments and public organizations mirror those for businesses: increased efficiency and sustainability, fact-based decision-making, a universal language to enhance communication between departments and with citizens, as well as increased awareness and understanding of ethical, security and privacy considerations. Successful applications of data could moreover support national security and military advantage, improved health, efficiency, responsiveness and convenience throughout societies, as well as increased transparency and accountability for public entities (Andrejevic, 2014). The trend of smart city projects has seen data-collecting technologies and related applications appear throughout the public spaces, with examples such as the municipality of Santander in Spain with its 18,000 sensors for 180,000 residents in 2015 (Trebourg, Jara, Dufour, & Genoud, 2015), the Virtual Singapore project which is devoted to the creation of a 3D map of the city using its building, land and environmental data (Chia, 2016), or Rio de Janeiro's operations center, built in partnership with IBM and sourcing data from 30 agencies (Kitchin, 2014). Collaboration across sectors is necessary to ensure successful data applications in the urban context, implementing interoperable systems and providing access to datasets (Bertot, Ursula, Jaeger, Sarin, & Choi, 2014). Thus, data literacy can once again serve as the universal language between sectors. Data can support urban planning and policy development (Lim, Kim, & Maglio, 2018), as public entities must draw value from data in efficient, equitable and sustainable ways to take on social challenges or climate change (Batty et al., 2012). Data literate public officials can better understand the past, manage the present and prepare for the future by drawing actionable insights from data.

As they are called to make decisions and implement laws and policies regarding data management, open data, security and privacy, or to make decisions using citizens' data, public officials must understand the issues at heart. It is critical for our political leaders and administrators to become data literate, understanding data itself and the ethics surrounding it. The results acquired through data analytics could indeed show patterns that would lead to unethical or even illegal decisions that discriminate against specific groups (Chakrabarti et al., 2009). Predictions provided by data mining are probabilistic (Andrejevic, 2014) and should be understood as such, meaning that deciders must understand the limitations and possible pitfalls of data and analyses and not blindly apply the actions that align with the most likely outcome. As was pointed out previously, regardless of how unlikely the data shows a potential future event to be, it does not render this event impossible. Critical interpretation and contextualization of results, as well as transparency, are essential to understanding the implications and consequences of decisions. Public entities should also advocate for data literate citizens. According to Andrejevic (2014), there is a "big data divide", a power imbalance "between those who collect, store,

and mine large quantities of data, and those whom data collection targets" (p. 1). Data can foster an informed, participative public (Al Nuaimi, Al Neyadi, Mohamed, & Al-Jaroodi, 2015), and it can be used as a communication tool between governments and their citizens, meaning that data literacy is an "important civic skill" (McAuley et al., 2010, p. 89), as well as "a plank for effective democracy" (Frank et al., 2016, p. 5).

### *Education*

Pedersen and Caviglia (2019) find that "data literacy is suggested as a central element in education preparing all young people to become citizens in an information society" (p. 1), yet one study found that a mere 21% of people between the ages of 16 and 24 were data literate (Olik, 2018). As noted previously, new opportunities and careers are appearing with relation to data, while future ones have not yet been invented (Morrow, 2019), and there is a shortage of data scientists (Davenport & Patil, 2012). It is crucial to prepare students, tomorrow's workforce, to critically collect, manage, interpret and apply data, tasks that will be a part of their work in most industries (Bowne-Anderson, 2018; Ridsdale et al., 2015). Higher education has an important role to play in using society's increasing amounts of data in innovative ways (McAuley et al., 2010). Data literacy supports the learning process of domain-specific knowledge, it is a valuable skill on both a personal and professional level, and it aligns with the "overarching goal of producing globally competitive, critically thinking, well-equipped graduates" (Ridsdale et al., 2015, p. 2). Educational institutions must accept change and adapt to changes in society, evolving curricula to meet new requirements (Álvarez-Arregui et al., 2017). Furthermore, to ensure students enter higher education prepared to integrate new data skills and their specific domain's knowledge, it is necessary to start early by teaching the foundations of data literacy in schools (Wolff et al., 2016), extending from the basic skills of statistical literacy that may already be included in curricula (Frank et al., 2016). As explained by Pedersen and Caviglia (2019), "educational reform and innovation has a key role to play in bringing data literacy to the broad educated population across disciplines and curricula" (p. 3).

From an internal standpoint, data literacy is also invaluable to driving research and academia forward. Data can be a source of research or its output, which can be shared and reused across projects, meaning that data skills and standards are necessary (Shorish, 2015). While graduate students are expected to have the ability to manage data when joining a lab, a task they must take on with varying degrees of responsibility, they are often insufficiently prepared (Jacob Carlson et al., 2011; Shorish, 2015). They must then learn as they go, but each student may learn differently without standard practices, making future interoperability and preservation of data difficult (Jacob Carlson et al., 2011). Graduate students are often central to the research process, carrying out many data-related tasks, yet they may lack the understanding of proper documentation and storage practices, of ownership concerns, or of the data's potential long-term value (Jake Carlson, Nelson, Johnston, & Koshoffer, 2015). Indeed, as research is continuously evolving, "understanding that data have the potential to impact not only one's own research, but also the work of others – in fields that may appear unrelated – can help build an awareness of the diverse scholarship system" (Shorish, 2015, p. 11). This means that training students, as well as faculty, in a holistic manner and implementing clear data management standards is crucial to remove barriers to the advancement of research.

## Individuals

Through the lenses of business, society and education, we have explored the importance of data literacy for individuals as employees, citizens, students and scholars. It provides access to new business opportunities, enables people to perform better in their work, to support their decisions with facts and to drive the innovation of their organization, allows them to better communicate with and understand public administration and governments, to participate as active citizens, to draw informed opinions and ideas, and to accelerate scientific research, sharing data and insights across fields and continuously improving and adapting education to societal needs.

Data literacy also holds great importance on a personal level. Individuals come into contact with data more and more often, using it as a basis for decisions such as what to buy, who to vote for or where to live, and make choices regarding how their own data can be used (Wolff et al., 2016). Frank et al. (2016) argue that the ability to use data and derive information "is rapidly becoming a requirement to participate in modern life – as fundamental as the ability to use a telephone or money" (p. 5) and that it leaves those without these skills at a disadvantage. As noted previously, there is a divide between those who collect and use data and those who are the subjects of this data, as those with large amounts from many individuals can extract far more information than an individual with access only to their own, single data point (Andrejevic, 2014). Personal data can be used to determine decisions such as the recipient of a loan or scholarship, which can result in unethical discrimination (Chakrabarti et al., 2009). Individuals must understand how their data is being used for or against them, and how this affects their life and the options available to them. They should also understand the privacy and security implications of sharing their personal data, in order to make informed decisions about its use (Crusoe, 2016; Ridsdale et al., 2015). This requires transparency from data collectors and handlers, new laws and policies regulating the use of data, and data literacy skills across the population.

In a technological era with constant access to media stories, a never-ending influx of information, and the possibility of echo chamber or filter bubbles (see Flaxman, Goel, & Rao, 2013), people must acquire the skills to critically assess the information being presented to them, as well as the capacity to search for, select, analyze and interpret data relevant to their questions. They should further be capable of evaluating the validity of data, question sampling methods and identify possible biases in order to reach sound conclusions. Biased data or interpretations, which may bring forth incorrect knowledge and bad decisions, can arise inadvertently from poor data literacy or be introduced willfully with the intent to manipulate readers (Wolff et al., 2016). Thus, these skills are essential to making data-driven decisions in all aspects of life. Furthermore, in an ever-changing world, people must learn to seek out and integrate new data, adapting their opinions and actions according to situational and environmental changes. This aligns with the need for continuous learning and adaptation across fields (Álvarez-Arregui et al., 2017), as industries are disrupted by innovations and the world changes at an accelerating pace.

## Fostering Data Literacy

It is thus clear that data literacy is essential to any individual or group that come into contact with data, which is now an omnipresent aspect of society. How can we foster data literacy

throughout the population and grasp the opportunities it offers? Several considerations are crucial. We will discuss them in this section.

## Setting Objectives with Relation to Context

Attaining data literacy does not require one to become a data scientist (Pedersen & Caviglia, 2019). It involves understanding what data is and having the ability to collect, analyze, visualize and share it, as well as developing awareness of the ethical questions, including security and privacy issues, related to the use of data (Crusoe, 2016; Wolff et al., 2016). Different levels of competence are required for different situations (Frank et al., 2016) and people can attain literacy pertinent to a specific role, such as data communicator, reader, maker or scientist (Wolff et al., 2016). It is necessary to recognize the skills relevant to a given context and to the needs of an individual or an organization. The foundations of data literacy may be common across fields, but different domains will require different sets of skills. For example, the creation of visualizations and the capacity for data storytelling are useful for those who work in marketing, while the storage, documentation and preservation of data with consideration for the possibility of future use should be understood by those in academia. Some will need in-depth abilities to handle data and analyze it, some may focus on using it to communicate, while others will meet their needs with a basic awareness and understanding of data, including the ability to critically evaluate the interpretations provided by others, turning them into appropriate decisions and actions. This demonstrates the necessity of scalable and flexible curricula or training programs, adapting objectives to the required levels data literacy (Schneider, 2013). In order to motivate learners to take on the challenge of developing their data literacy, its potential benefits in their given context and its relevance to their needs should be clearly outlined (Ridsdale et al., 2015). Moreover, as learners must put theory into practice to fully integrate new data skills and as the availability of open data continues to rise, it is possible to develop skills with practice using real-world data. Learners from all backgrounds can then relate to the wider community, better understand the potential impact of their contributions, and generate innovative solutions to real-world problems (Ridsdale et al., 2015).

## Continuous Learning

To truly develop a culture of data literacy, its importance as a civic skill should be understood and acknowledged throughout society, whether by educational institutions or governments (Ridsdale et al., 2015). The teaching of its fundamentals should begin in school (Ridsdale et al., 2015; Wolff et al., 2016), it should be taught at all levels (Calzada Prado & Marzal, 2013) and as a cross-curricular subject (Calzada Prado & Marzal, 2013; Ridsdale et al., 2015; Vahey et al., 2012; Wolff et al., 2016). We have previously highlighted the key role of higher education in developing a data-literate workforce and the value academia can gain from increased data literacy. As society, processes and technologies change at a quick pace, it will furthermore be essential to never stop learning. People will need to adapt to new forms of data and integrate new possibilities into their experience. This means that acquiring new data skills should not stop at the end of an individual's education; organizations that wish to draw value from data will need to train their members to keep up-to-date with innovations and new opportunities. The objectives and the skills that constitute data literacy will need to be revised to reflect societal and environmental changes.

## A Community Competence

As a result of this need for constant adaptation and learning, developing data literacy across the population will require collaboration and the pooling of resources between stakeholders. Organizations, educational institutions, governments and individuals all have their interests and their parts to play in instilling a data culture. This can be better achieved in a coherent and holistic manner through communication and collaboration, for example between universities and businesses to ensure graduates are prepared to enter the workforce and meet business requirements, or amongst educators to eliminate gaps or overlaps in curricula (Ridsdale et al., 2015). Teachers of data skills may themselves not be data specialists, rather teaching them in the context of a specific domain of expertise, so support, collaboration and peer learning between teachers will be necessary to provide a full range of data skills to students (Pedersen & Caviglia, 2019).

As previously discussed, data literacy is a flexible concept which depends on an individual's or organization's specific context and needs, and which evolves according to societal and technological changes. Thus, people will attain varying degrees of data literacy and specialize in certain aspects depending on their needs. Each individual will then have their own data skills to contribute, "so that the presence of some people who can find data, some people who can manipulate it, and some who can present the result might constitute data literacy for that community" (Frank et al., 2016, p. 6). In this way, data literacy is a collective competence to which each community member brings something different and that should be taught in collaborative, interdisciplinary ways, joining competences to foster it in heterogeneous communities (Pedersen & Caviglia, 2019). Bringing people from many disciplines together also holds enormous potential to draw more value from data. Because datasets can be used, analyzed and interpreted in many different ways, leading to a variety of different conclusions, it is important to create discussion around data and include the ideas, experience and perspectives of people from various backgrounds and fields. Silberzahn and Uhlmann (2015) asked 29 teams to study the same dataset, seeking the answer to the same question, and found that their conclusions varied depending on their chosen analytical methods. They argue that crowdsourcing research in this manner can temper the attraction of "flashy results" and "gives space to dissenting opinions" (p. 191). Moreover, creating discussions around data and giving voices to people from various disciplines allows us to tap into the "mental database" (see Forrester, 2009, p. 12), completing our digital datasets with that data which cannot be captured and exists in people's heads: ideas, experiences, understanding of cultural and contextual subtleties, and so forth. This data is crucial to include in order to reach sound conclusions and make informed decisions. Using data as a common point of discussion, exchanging perspective, ideas and opinions through the shared, universal language of data literacy, holds many possibilities for progress, with the opportunity to find and imagine better solutions to the challenges of today and tomorrow.

## Conclusion

Our review of various definitions of data literacy in the literature showed different interpretations of the concept: it refers to various abilities such as understanding, using and communicating data, along with the capability to search for solutions and solve problems with the support of data,

with relevance to a given context and with the awareness of ethical considerations such as privacy or security. The various perspectives surrounding the concept converge on one point: data literacy is a crucial skill to face the challenges and to draw on the opportunities of an information society with ubiquitous data-collecting technologies. Data skills can help businesses get a competitive edge, as well as increase the efficiency and sustainability of their operations; they are an important civic skill for citizens to engage in shaping their cities, countries and the world; they are essential for administrators, decision-makers or legislators who are called to develop policy concerning technology, security and privacy, or using data to make decisions on other subjects; they can drive forward academic research and are crucial competencies which education must impart on the future workforce; and they are invaluable to all individuals, in all fields, to make informed, fact-based decisions throughout life.

Fostering data literacy across fields will require a recognition of its value at all levels of society. The objectives for each individual or organization should be relevant to their needs and their context, and should not be immovable: as environments and society shift, we will need to adapt and learn new skills that reflect the world we live in. This means that data literacy should be inscribed in a process of continuous learning. It is furthermore necessary to treat data literacy as the interdisciplinary competence that it is: its value spans across fields and it can serve as a universal language to converse between departments, sectors, backgrounds and disciplines. Different perspectives and methods extract different conclusions from data, so it is essential to collaborate and discuss interpretations in order to reach sound conclusions and decisions. It is also a community competence, meaning that a high degree of data literacy can be attained by a community to which each member contributes their own individual skills. Thus, collaboration across disciplines will help us make the best use of the large amounts of data available to us.

Data will not (and should not) make our decisions for us; it is up to us to draw insights, to interpret them critically and to turn them into appropriate decisions and actions. Technology and data are invaluable tools with the potential to move our society forward. We must learn to understand them, as well as to use them effectively and responsibly.

## Referencias /References

- Al Nuaimi, E., Al Neyadi, H., Mohamed, N., & Al-Jaroodi, J. (2015). Applications of big data to smart cities. *Journal of Internet Services and Applications*, 6(25). <https://doi.org/10.1186/s13174-015-0041-5>
- Álvarez-Arregui, E., Rodríguez-Martín, A., Agudo Prado, S., & Arreguit, X. (2017). Inclusive ecosystem model for the management of knowledge, training and innovation: Progress, difficulties and challenges. *Aula Abierta*, 46(2), 97–104. <https://doi.org/10.17811/rifie.46.2017.97-104>
- Álvarez-Arregui, E.; Rodríguez-Martín, A.; Madrigal Maldonado, R.; Grossi-Sampedro, B. G.; y Arreguit, X. (2017). Ecosystems of media training and competence. International assessment of its implementation in Higher Education. *Comunicar*, 51, XXV, 105-114. <https://doi.org/10.3916/C51-2017-10>
- Andrejevic, M. B. (2014). The big data divide. *International Journal of Communication*, 8(1), 1673–1689.
- Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., ... Portugali, Y. (2012). Smart cities of the future. *The European Physical Journal Special Topics*, 214(1), 481–518. <https://doi.org/10.1140/epjst/e2012-01703-3>

- Bertot, J. C., Ursula, G., Jaeger, P. T., Sarin, L. C., & Choi, H. (2014). Big data, open government and e-government: Issues, policies and recommendations. *Information Polity*, 19(1,2), 5–16. <https://doi.org/10.3233/IP-140328>
- Bowne-Anderson, H. (2018, October 23). Your Data Literacy Depends on Understanding the Types of Data and How They're Captured. *Harvard Business Review*. Retrieved from <https://hbr.org/2018/10/your-data-literacy-depends-on-understanding-the-types-of-data-and-how-theyre-captured>
- Calzada Prado, J., & Marzal, M. Á. (2013). Incorporating Data Literacy into Information Literacy Programs: Core Competencies and Contents. *Libri*, 63(2), 123–134. <https://doi.org/10.1515/libri-2013-0010>
- Carlson, Jacob, Fosmire, M., Miller, C. C., & Sapp Nelson, M. (2011). Determining Data Information Literacy Needs: A Study of Students and Research Faculty. *Portal: Libraries and the Academy*, 11(2), 629–657. <https://doi.org/10.1353/pla.2011.0022>
- Carlson, Jake, Nelson, M. S., Johnston, L. R., & Koshoffer, A. (2015). Developing Data Literacy Programs: Working with Faculty, Graduate Students and Undergraduates. *Bulletin of the Association for Information Science and Technology*, 41(6), 14–17. <https://doi.org/10.1002/bult.2015.1720410608>
- Chakrabarti, S., Cox, E., Frank, E., Güting, R. H., Han, J., Jiang, X., ... Witten, I. H. (2009). *Data Mining: Know It All* (S. Chakrabarti, Ed.). Burlington, MA: Elsevier/Morgan Kaufmann Publishers.
- Chia, E. S. (2016). Singapore's smart nation program—Enablers and challenges. *2016 11th System of Systems Engineering Conference (SoSE)*, 1–5. <https://doi.org/10.1109/SYS-OSE.2016.7542892>
- Crofts, J. (2018, April 29). Data Literacy: The Universal Language Of The Fourth Industrial Revolution. Retrieved August 28, 2019, from LinkedIn website: <https://www.linkedin.com/pulse/data-literacy-universal-language-fourth-industrial-jane-crofts>
- Crusoe, D. (2016). Data Literacy defined pro populo: To read this article, please provide a little information. *The Journal of Community Informatics*, 12(3), 27–46.
- Davenport, T. H., & Patil, D. J. (2012). Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review*, 90(5), 70–76.
- Downey, A. (2016, November 14). Why are we so surprised? [Blog]. Retrieved September 20, 2019, from Probably Overthinking It website: <http://alldowney.blogspot.com/2016/11/why-are-we-so-surprised.html>
- Flaxman, S., Goel, S., & Rao, J. M. (2013). Ideological Segregation and the Effects of Social Media on News Consumption. *SSRN Electronic Journal*, 80. <https://doi.org/10.2139/ssrn.2363701>
- Forrester, J. W. (2009, January 29). *Some Basic Concepts in System Dynamics*. Retrieved from <http://static.clexchange.org/ftp/documents/system-dynamics/SD2009-02SomeBasicConcepts.pdf>
- Frank, M., Walker, J., Attard, J., & Tygel, A. (2016). Data Literacy: What is it and how can we make it happen? Editorial. *The Journal of Community Informatics*, 12(3), 4–8.
- García Martínez, A. (2019, February 26). No, Data Is Not the New Oil. Retrieved September 23, 2019, from Wired website: <https://www.wired.com/story/no-data-is-not-the-new-oil/>
- Harvard Business Review Analytic Services. (2012). *The Evolution of Decision Making: How Leading Organizations Are Adopting a Data-Driven Culture*. Retrieved from [https://hbr.org/resources/pdfs/tools/17568\\_HBR\\_SAS%20Report\\_webview.pdf](https://hbr.org/resources/pdfs/tools/17568_HBR_SAS%20Report_webview.pdf)
- Haupt, M. (2018, April 23). "Data is the New Oil" — A Ludicrous Proposition. Retrieved September 23, 2019, from Medium website: <https://medium.com/project-2030/data-is-the-new-oil-a-ludicrous-proposition-1d91bba4f294>
- Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating Statistical Information. *Science*, 290(5500), 2261–2262. <https://doi.org/10.1126/science.290.5500.2261>
- Hogenboom, K., Holler Phillips, C. M., & Hensley, M. (2011). Show Me the Data! Partnering with Instructors to Teach Data Literacy. *Declaration of Interdependence: The Proceedings of the ACRL 2011 Conference, March 30-April 2, Edited by Dawn M. Mueller*, 410–417. Philadelphia, PA.
- Hunt, K. (2005). The Challenges of Integrating Data Literacy into the Curriculum in an Undergraduate Institution. *IASSIST Quarterly*, 28(2–3), 12–16. <https://doi.org/10.29173/iq791>
- James, J. (2019, July 11). Data as the new oil: The danger behind the mantra. Retrieved September 23, 2019, from The Enterprisers Project website: <https://enterpriseproject.com/article/2019/7/data-science-data-can-be-toxic>
- Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, 79(1), 1–14. <https://doi.org/10.1007/s10708-013-9516-8>
- Lim, C., Kim, K.-J., & Maglio, P. P. (2018). Smart cities with big data: Reference models, challenges, and considerations. *Cities*, 82, 86–99. <https://doi.org/10.1016/j.cities.2018.04.011>
- Marr, B. (2018, March 5). Here's Why Data Is Not The New Oil. Retrieved September 23, 2019, from Forbes website: <https://www.forbes.com/sites/bernardmarr/2018/03/05/heres-why-data-is-not-the-new-oil/>
- McAuley, D., Rahemtulla, H., Goulding, J., & Souch, C. (2010). How open data, data literacy and linked data will revolutionise higher education. In L. Coiffait & J. Hill, *Blue Skies: New Thinking about the Future of Higher Education. A collection of short articles by leading commentators* (pp. 88–93). Retrieved from [https://www.academia.edu/1945211/The\\_Open\\_Data\\_Revolution\\_and\\_Data\\_Literacy\\_in\\_Higher\\_Education](https://www.academia.edu/1945211/The_Open_Data_Revolution_and_Data_Literacy_in_Higher_Education)
- McCandless, D. (2010). *David McCandless: The beauty of data visualization* [Video file]. Retrieved from [https://www.ted.com/talks/david\\_mccandless\\_the\\_beauty\\_of\\_data\\_visualization](https://www.ted.com/talks/david_mccandless_the_beauty_of_data_visualization)
- Morrow, J. (2018). *Developing a data literate workforce: A strategy and framework for the enterprise*. Retrieved from Qlik website: <https://www.qlik.com/us/-/media/files/training/global-us/qlik-education-data-literacy-program-strategy-and-framework.pdf>
- Morrow, J. (2019, March 9). To Prepare Millennials for the Workplace Add Data Literacy to Mainstream Curriculum. Retrieved September 3, 2019, from Getting Smart website: <https://www.gettingsmart.com/2019/03/to-prepare-millennials-for-the-workplace-add-data-literacy-to-mainstream-curriculum/>
- Panetta, K. (2019, February 6). A Data and Analytics Leader's Guide to Data Literacy. Retrieved September 3, 2019, from Gartner website: <http://www.gartner.com/smarterwithgartner/a-data-and-analytics-leaders-guide-to-data-literacy/>
- Pedersen, A. Y., & Caviglia, F. (2019). Data Literacy as a Compound Competence. In T. Antipova & A. Rocha (Eds.), *Digital Science. DSIC18 2018. Advances in Intelligent Systems and Computing* (Vol. 850, pp. 166–173). [https://doi.org/10.1007/978-3-030-02351-5\\_21](https://doi.org/10.1007/978-3-030-02351-5_21)
- Qlik. (2018). *Lead with Data: How to Drive Data Literacy in the Enterprise*. Retrieved from QlikTech International AB website: [https://www.qlik.com/us/bi/-/media/08F37D711A58406E83BA8418EB1D58C9.ashx?ga-link=datlitreport\\_resource-library](https://www.qlik.com/us/bi/-/media/08F37D711A58406E83BA8418EB1D58C9.ashx?ga-link=datlitreport_resource-library)
- Reinsel, D., Venkatraman, A., Gantz, J. F., & Rydning, J. (2019). *La datasphère EMEA: expansion rapide et migration vers la*

- périphérie* (p. 12). Retrieved from IDC website: <https://www.seagate.com/files/www-content/our-story/trends/files/data-age-emea-fr-regional-idc.pdf>
- Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., ... Wuetherick, B. (2015). *Strategies and Best Practices for Data Literacy Education: Knowledge Synthesis Report*. Retrieved from Dalhousie University website: <https://dalspace.library.dal.ca/handle/10222/64578>
- Rosenbaum, S. (2015, May 17). The Quantified Self—Measuring To Curate Your Life. Retrieved September 20, 2019, from Forbes website: <https://www.forbes.com/sites/stevenrosenbaum/2015/05/17/the-quantified-self-measuring-to-curate-your-life/>
- Schild, M. (2005). Information Literacy, Statistical Literacy, Data Literacy. *IASSIST Quarterly*, 28(2-3), 6-11. <https://doi.org/10.29173/iq790>
- Schneider, R. (2013). Research Data Literacy. In S. Kurbanoglu, E. Grassian, D. Mizrachi, R. Catts, & S. Špiranec (Eds.), *Worldwide Commonalities and Challenges in Information Literacy Research and Practice* (Vol. 397, pp. 134-140). [https://doi.org/10.1007/978-3-319-03919-0\\_16](https://doi.org/10.1007/978-3-319-03919-0_16)
- Shorish, Y. (2015). Data Information Literacy and Undergraduates: A Critical Competency. *College & Undergraduate Libraries*, 22(1), 97-106. <https://doi.org/10.1080/10691316.2015.1001246>
- Silberzahn, R., & Uhlmann, E. L. (2015). Crowdsourced research: Many hands make tight work. *Nature*, 526(7572), 189-191. <https://doi.org/10.1038/526189a>
- Treboux, J., Jara, A. J., Dufour, L., & Genoud, D. (2015). A predictive data-driven model for traffic-jams forecasting in Smart Santander city-scale testbed. *2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 64-68. <https://doi.org/10.1109/WCNCW.2015.7122530>
- Vahey, P., Rafanan, K., Patton, C., Swan, K., van 't Hooft, M., Kratcoski, A., & Stanford, T. (2012). A cross-disciplinary approach to teaching data literacy and proportionality. *Educational Studies in Mathematics*, 81(2), 179-205. <https://doi.org/10.1007/s10649-012-9392-z>
- Van Zeeland, J. (2019, February 11). Data is Not the New Oil. Retrieved September 23, 2019, from Medium website: <https://towardsdatascience.com/data-is-not-the-new-oil-721f5109851b>
- Wolff, A., Gooch, D., Cavero Montaner, J. J., Rashid, U., & Kortuem, G. (2016). Creating an Understanding of Data Literacy for a Data-driven Society. *The Journal of Community Informatics*, 12(3), 9-26.
- Wu, S., Chen, T., Wu, Y., & Lytras, M. (2018). Smart Cities in Taiwan: A Perspective on Big Data Applications. *Sustainability*, 10(1), 106. <https://doi.org/10.3390/su10010106>