



TRABAJO DE FIN DE GRADO

# La Ciencia de Datos en la actualidad: Desarrollo, implicaciones y regulación

Data Science Today: Development, Implications and Regulation

Aitor Carrascoso Ballesteros

Tutora: Gema García Rojas

28 / 05 / 2023

Grado en Gestión Empresarial Basada en el Análisis de Datos

Curso 2022 -2023



# I. Resumen y palabras clave

## **Resumen:**

El fenómeno del Big Data y las prácticas de minería, limpieza, análisis y explotación de datos, comprendido todo ello en lo que se conoce como ciencia de datos, ha transformado en las últimas décadas todos los procesos, estructuras y comportamientos de los agentes económicos que conforman el sistema capitalista, nacido en la segunda mitad del siglo XVIII, consolidado a lo largo de los siglos XIX y XX, y llevado a su máxima expresión de la mano de una globalización resultada de los avances tecnológicos provenientes, sobre todo, de las telecomunicaciones y el transporte.

El objetivo principal del presente estudio es determinar las consecuencias materiales que el fenómeno, al que nos referiremos únicamente como Ciencia de Datos (CD en adelante) tiene sobre la sociedad desde los puntos de vista económico, sociológico y político, partiendo primeramente de un estudio genealógico y ontológico que permita determinar en primer lugar la constitución de la CD como disciplina propia hasta el papel que desempeña hoy como base estructural del sistema y, en segundo lugar, una definición exhaustiva de la disciplina, para poder así comprender correctamente el fenómeno al que nos referimos y el impacto real de este. Por último, llevaremos a cabo una aproximación crítica a la legislación encargada de regular el alcance de la CD en los tres aspectos que estudiamos previamente.

## **Palabras clave:**

Ciencia de datos, Ciencia, Tecnología, Ciencias de la información, Estadística, Informática, Economía, Política, Sociedad, Derecho, Educación.

## **Abstract:**

The phenomenon of Big Data and the practices of mining, cleaning, analysis, and data exploitation, all encompassed in what is known as data science, has transformed all the processes, structures, and behaviors of economic agents making up the capitalist system in recent decades. This system was born in the second half of the 18th century, consolidated throughout the 19th and 20th centuries, and brought to its maximum expression by globalization resulting from technological advancements, primarily from telecommunications and transportation.

The main objective of this study is to determine the material consequences that the phenomenon, to which we will refer only as Data Science (CD henceforth), has on society from the economic, sociological, and political points of view. This begins first with a genealogical and ontological study to initially determine the constitution of CD as its own discipline up to the role it plays today as the structural base of the system. Secondly, we aim to provide a comprehensive definition of the discipline to properly understand the phenomenon we are referring to and its real impact. Finally, we will carry out a critical approach to the legislation responsible for regulating the scope of DS in the three aspects we previously studied.

**Keywords:**

Data Science, Science, Technology, Information Science, Statistics, Computer Science, Economics, Politics, Society, Law, Education.

## II. Índice

I. Resumen y palabras clave .....	3
II. Índice.....	5
III. Introducción.....	7
IV. Objetivos .....	10
V. Desarrollo.....	11
1. Motivación.....	11
2. ¿Qué es la Ciencia de Datos? .....	13
1. Definición preliminar de Ciencia de Datos .....	13
2. Genealogía de la Ciencia de Datos .....	14
<b>Orígenes de la Ciencia de Datos</b> .....	14
<b>Evolución de la Ciencia de Datos</b> .....	21
3. Ontología de la Ciencia de Datos.....	29
<b>¿Qué es la ciencia?</b> .....	29
<b>¿Es la Ciencia de Datos una ciencia?</b> .....	40
<b>¿Qué es la Ciencia de Datos?</b> .....	44
3. Impacto de la Ciencia de Datos en la actualidad.....	47
1. Economía .....	48
<b>Teoría del valor</b> .....	50
<b>Economía del comportamiento</b> .....	53
<b>El dato como mercancía, el usuario como productor</b> .....	55
<b>Distribución de la riqueza. Monopolio de la información y las Big Tech</b> .....	57
<b>Conclusiones</b> .....	62
2. Política .....	64
<b>1984 y las sociedades de control</b> .....	66
<b>Vigilar y castigar</b> .....	68
<b>El caso de la República Popular China</b> .....	72
<b>El escándalo de Cambridge Analytica</b> .....	76
<b>Posverdad</b> .....	77
<b>Conclusiones</b> .....	80
3. Sociedad .....	82
<b>Las redes sociales como campo de batalla</b> .....	83
<b>Ideología y tribalismo</b> .....	86
<b>Arte y entretenimiento</b> .....	91

<b>Conclusiones</b> .....	95
4. Análisis crítico de la legislación en materia de datos .....	97
1. Protección, privacidad y propiedad de los datos .....	100
<b>Aciertos</b> .....	100
<b>Errores</b> .....	101
<b>Carencias</b> .....	102
2. Monopolios de información .....	105
<b>Aciertos</b> .....	105
<b>Errores</b> .....	106
<b>Carencias</b> .....	107
3. Responsabilidad y transparencia de las empresas de redes sociales.....	109
<b>Aciertos</b> .....	109
<b>Errores</b> .....	110
<b>Carencias</b> .....	111
4. Regulación de publicidad y contenido sensible .....	113
<b>Aciertos</b> .....	113
<b>Errores</b> .....	114
<b>Carencias</b> .....	114
5. Control de la desinformación .....	117
<b>Aciertos</b> .....	117
<b>Errores</b> .....	118
<b>Carencias</b> .....	119
6. Control de contenido sesgado.....	121
<b>Aciertos</b> .....	121
<b>Errores</b> .....	122
<b>Carencias</b> .....	123
7. Alfabetización digital y mediática.....	125
<b>Aciertos</b> .....	125
<b>Errores</b> .....	126
<b>Carencias</b> .....	127
VI. <b>Conclusiones</b> .....	130
VII. <b>Bibliografía</b> .....	133
VIII. <b>Anexos</b> .....	141

### III. Introducción

Vivimos en un mundo en constante cambio donde, en ocasiones, parecemos movidos por las grandes innovaciones de un lugar para otro, fijando la mirada un instante en un punto para fijarla a continuación en otro distinto. Un mundo en transformación por unas tecnologías cada vez más poderosas, resultado de los grandes avances científicos de los últimos años, cuyos despliegues se extienden por el largo y ancho del entorno, ocupando hasta aquellos lugares que creíamos inmunes a toda innovación tecnológica. El papel del ser humano, incapaz de comprender por completo la variedad y diversidad de movimientos procedentes de todos los campos del saber, se resigna a una posición de observación, desde la cual, lejos de operar y tomar partido en la constitución de estos poderes, ejerce la fuerza necesaria para impulsarlos. Necesaria para un pueblo que debe adaptarse a un entorno que no espera rezagados y necesaria para las innovaciones tecnológicas que requieren del apoyo masivo para alcanzar sus niveles más altos de desarrollo.

La ciencia de datos es uno de estos grandes avances, constituida no a partir del desarrollo de una sola ciencia sino de la combinación de múltiples, que transforma por completo las posibilidades y alcance de un entorno que monitoriza por completo. Sin embargo, la ciencia de datos consta de unas características especiales que han permitido que su presencia, a pesar de ser imprescindible, pase completamente inadvertida por la sociedad. Su propósito como medio y no como fin en sí mismo, como son otros muchos avances cuya razón de ser es su constitución y mantenimiento, hace a la ciencia de datos un caso especial, una especie de herramienta que se adapta e incorpora los conocimientos propios de otros campos sobre los que explota las posibilidades de la información generada en estos. Su avance y extensión ha llegado al punto de extenderse a todos los campos del saber y del hacer, siendo esta segunda parte la que más preocupa en este estudio, pues la ciencia de datos se ha convertido en el *modus operandi* de aquellos que tienen el poder de desplegar estas herramientas sobre el resto de agentes con los que se relacionan, estando estos, en ocasiones, en igualdad de condiciones, y en ocasiones en una desventaja absoluta derivada del desconocimiento de estas tecnologías.

En este contexto, donde los individuos podemos vernos desbordados, a veces incluso sepultados, por la magnitud de estas transformaciones, buscamos, desde este trabajo, la calma y la observación sobre este fenómeno de la ciencia de datos, como el único medio posible para comprenderlo por completo y ser capaces de adaptarnos a él como sociedad. La continua indiferencia y el movimiento por inercia al que nos limitamos, como una suerte de estoicismo

tecnológico, no es viable cuando las innovaciones traen consigo posibilidades tan peligrosas como las de la ciencia de datos. La única forma de controlar sus consecuencias es abordando estas innovaciones desde una perspectiva crítica que nos permita conocer su razón de ser, sus implicaciones y las posibilidades de regularlo desde una de las herramientas más poderosas que el ser humano ha desarrollado en su historia: el derecho.

Con todo esto, los objetivos que nos proponemos en la presente investigación son, en primer lugar, desarrollar un conocimiento férreo de la ciencia de datos, examinando sus orígenes más remotos, su constitución desde otras ciencias y el conjunto de herramientas y conocimientos que le dan forma. Para ello, desarrollamos un estudio genealógico, a través del cual determinaremos los antecedentes de esta disciplina recién creada y cómo se combinan para darle forma, y un estudio ontológico, donde la atención se centrará en determinar qué es el objeto sobre el que investigamos. Con esta primera parte queremos determinar y delimitar el objeto de estudio para, a continuación, poder observar hasta dónde se extiende su uso y cuales son las problemáticas que acarrea.

En segundo lugar, comprender el alcance de la ciencia de datos en, como hemos dicho antes, los campos del hacer, organizados analíticamente como las relaciones entre agentes económicos, las relaciones entre los ciudadanos y las esferas políticas, y las relaciones entre individuos como miembros de una sociedad. En este punto, es importante aclarar que el foco se pondrá directamente sobre las problemáticas y abusos derivados de la ciencia de datos sobre estos tres tipos de relaciones, porque solo así podremos llegar preparados a una última parte donde nos ocuparemos de poner a prueba la potencia de la legislación en materia de datos para hacer frente a las problemáticas a las que nos hemos referido.

Así, el último objetivo del estudio es realizar una aproximación crítica sobre las distintas regulaciones que, desde los aparatos legislativos europeos principalmente, se han propuesto establecer un marco sobre el que tanto individuos como instituciones deben subordinarse para hacer de la ciencia de datos un instrumento que favorezca el progreso y la libertad y no lo contrario.

Como se verá, la aproximación, al igual que el objeto de estudio, es multidisciplinar, primeramente, desde una perspectiva histórica, hacia los campos del saber donde encontramos las primeras piedras de la ciencia de datos, y una perspectiva técnica hacia los distintos materiales que conforman el campo de la ciencia de datos, examinando las distintas herramientas, lenguajes de programación y sistemas que le dan forma. A continuación, desde una perspectiva económica, política y sociológica se examinan las implicaciones derivadas y, por

último, desde una perspectiva legal la incorporación al marco regulatorio comunitario. La razón de este acercamiento desde distintas perspectivas es asegurar un desarrollo orgánico de la investigación, no limitándonos a una perspectiva legal, por ejemplo, que carecería de sentido sin la determinación del objeto a legislar y las consecuencias de este.

Por todo esto, este trabajo se embarcará en una exploración profunda y crítica de la ciencia de datos, sus implicaciones y posibles soluciones a los problemas que plantea. Con esto, buscamos proporcionar un marco sólido para futuras investigaciones, establecer un diálogo constructivo y fomentar un mayor entendimiento de este campo tan relevante y transformador. De esta forma, con una comprensión más completa y una perspectiva crítica, podremos navegar con mayor confianza en este mundo en constante evolución, y, lo que es más importante, podremos hacerlo de manera que sea en beneficio de todos.

## IV. Objetivos

El objetivo del estudio es analizar los comportamientos de las sociedades contemporáneas perpetrados por la Ciencia de Datos.

Los objetivos específicos que permitirán alcanzar la meta del estudio son los siguientes:

1. Examinar la constitución de la CD a partir de otras ciencias hasta la constitución de una disciplina multidisciplinar y transdisciplinar propia.
2. Determinar las características esenciales del objeto del estudio, la CD.
3. Analizar el impacto de la CD desde un punto de vista económico: economía de comportamiento, datos como mercancía, monopolios de información.
4. Analizar el impacto de la CD desde un punto de vista político: sociedades de control, el caso de la República Popular China, el escándalo de Cambridge Analytica, Posverdad.
5. Analizar el impacto de la CD desde un punto de vista sociológico: redes sociales, tribus digitales, arte y entretenimiento en el entorno digital.
6. Estudio crítico de la legislación en materia de datos.

## V. Desarrollo

# 1. Motivación

Es paradójico que la principal razón que soy capaz de localizar como motor de esta investigación sea la progresiva pérdida de motivación hacia el grado universitario que doy cierre con este trabajo. Si bien la ilusión e interés era alto en un principio, a medida que avanzaba superando las distintas asignaturas, otras disciplinas conseguían captar fácilmente mi atención, desviándola de un desempeño académico que se ha visto perjudicado por ello.

De la mano de Richard Dawkins, Roger Penrose, Douglas Hofstadter y Brian Greene me adentré en los libros de divulgación científica, que, sin alejarme aun de los contenidos de la carrera, ampliaron mis horizontes y mis intereses hasta los textos de filosofía en los que me sumergí completamente, ahora sí, cuestionándome una elección de carrera que cada día que pasaba se me hacía más precipitada. De esta manera, proseguí una carrera universitaria asediado por la duda de elegir entre un futuro en la ciencia de datos, aparentemente cómodo en el ámbito laboral, y un reinicio académico en alguna facultad de filosofía, destinado a la docencia en el mejor de los casos. Fue la primera de las opciones la que más fuerza tuvo y la que, a día de hoy, creo era la más acertada.

Así, continué superando los distintos desafíos que el trabajo académico exigía, ampliados por la compaginación con el empleo y una educación en filosofía cuya exigencia autoimpuesta era mayor incluso que la universitaria. Sin embargo, esta forma de actuar dejaba ver ciertos signos de contradicción, pues cuanto más me adentraba en asuntos filosóficos, más atención me requerían y menos interés me suscitaban la ingeniería de datos, la dirección estratégica o las herramientas de CRM.

Y así era hasta hace alrededor de un año, cuando di con la obra del filósofo Gustavo Bueno, clave en este estudio, donde abordaba el saber filosófico desde una perspectiva que no había apreciado hasta entonces. Para el filósofo riojano, la filosofía, al contrario de otras muchas opiniones que la consideran como la ciencia primera o el saber por antonomasia (así lo creía yo también), es un saber de segundo grado, cuyo campo de estudio no es otro que los saberes de primer grado, es decir, los saberes técnicos, políticos, matemáticos... La filosofía no tiene un campo propio sobre el que investigar, sino que es un saber del presente, de un presente definido por los avances técnicos, los momentos políticos y las relaciones sociales. El objetivo, entonces,

de la filosofía es determinar, a través de la crítica y la dialéctica, las Ideas inherentes a los conceptos que moldean nuestro presente, para poder comprender su funcionamiento, su interpretación y las consecuencias que derivan. Es por ello que Gustavo Bueno piensa incluso en la desaparición de las facultades de filosofía y propone una escuela de filosofía posterior a la universidad, donde estudiantes procedentes de distintas disciplinas acudan para reflexionar sobre los conceptos pertenecientes a los campos categoriales que trabajan.

Es mediante este marco establecido por Gustavo Bueno a través del cual me reconcilio con el grado universitario y con la ciencia de datos, sobre la que crece un interés ya no práctico, como se suele abordar esta disciplina, sino teórico y filosófico, que con la elaboración de la presente investigación creo que doy por satisfecho. He aquí, a continuación, el resultado de mis investigaciones.

## 2. ¿Qué es la Ciencia de Datos?

### 1. Definición preliminar de Ciencia de Datos

La CD es un campo interdisciplinario que combina técnicas y conocimientos de estadística, matemáticas, informática y otras áreas del conocimiento para extraer información y conocimientos valiosos a partir de datos estructurados y no estructurados (CSIC). Esta disciplina se centra en el análisis, la interpretación, la visualización y la comunicación de datos, con el objetivo de apoyar la toma de decisiones y resolver problemas en diversos campos, como la industria, la investigación científica, la medicina, las finanzas, el marketing y el gobierno, entre otros.

Una de las características principales de la CD es su enfoque práctico y aplicado, lo que significa que busca resolver problemas del mundo real mediante el uso de datos y algoritmos. Los científicos de datos son profesionales que combinan habilidades técnicas, como la programación y el conocimiento de algoritmos, con habilidades analíticas y de comunicación para explorar, interpretar y presentar resultados a partir de datos.

A continuación, trataremos de exponer una genealogía de la CD. El porqué de una genealogía, en contraposición a una historia, es que la CD es un campo que, aunque combina múltiples disciplinas de largas trayectorias, es bastante moderna en su concepción, pues no podemos hablar propiamente de la CD hasta bien entrada la segunda mitad del siglo XX. Hacer una historia de la CD nos obligaría a comenzar desde el momento en que se empieza a considerar una disciplina propia, construyendo de esta manera una idea equivocada de un campo cuyos antecedentes son esenciales.

## 2. Genealogía de la Ciencia de Datos

### Orígenes de la Ciencia de Datos

Para comprender correctamente a lo que nos referimos cuando hablamos de CD, debemos retrotraernos a las disciplinas que la componen y sobre las que se apoya al llevar a cabo su actividad propia. La estadística y la computación, como hemos expuesto en nuestra definición, suelen nombrarse como los principales precedentes de la CD; la ciencia de la información también queremos considerarla como tal, en especial por la Teoría de la Información de Claude Shannon, sobre la que se apoya gran parte de la teoría aplicada a la CD.

Debemos tener en cuenta además que la CD no solo es interdisciplinar, sino también transdisciplinar, pues a través de las técnicas, herramientas y teorías propias de ciencias como la matemática o la computación, aborda problemas propios de todo tipo de disciplinas sobre las que el científico de datos debe apoyarse. Esta, como veremos, es una de las mayores potencias de la CD, pues que sea aplicable a cualquier campo de conocimiento en el que se generen datos y estos puedan ser tratados, convierte a la CD en una disciplina de vital importancia para la sociedad en la que vivimos, donde todo se mide con el objetivo de mejorarlo.

Ahora enumeraremos los distintos conceptos que, pertenecientes a campos de otras disciplinas, forman también parte del campo de la CD, ya como parte constituyente de él, ya como parte accesoria. Enumeraremos, pero no detallaremos, pues el objeto de esta primera parte del estudio no es desarrollar minuciosamente conceptos matemáticos o computacionales, sino mostrar cómo distintos avances en distintas ciencias son reestructurados e incorporados al campo de la CD.

#### 1. Estadística y matemáticas.

El primer paso en nuestro recorrido por los orígenes de la CD debe ser necesariamente el trazado a lo largo de la historia por los distintos avances en materia de matemáticas y estadística. Ambas disciplinas, como veremos, son de vital importancia en la CD tal y como la conocemos, la estadística por el conjunto de técnicas y teorías que se ocupan de recolectar, analizar, interpretar, presentar y organizar datos, y las matemáticas por ser el lenguaje en que se han podido desarrollar desde las primeras piedras del edificio estadístico hasta los complejos algoritmos que emplea hoy en día la CD.

La relación entre la estadística y la CD se remonta a los primeros días de la estadística como disciplina, cuando se enfocaba en la recolección y el análisis de datos, con el objetivo de extraer información y tomar decisiones. El concepto de probabilidad, introducido por Blaise Pascal y Pierre de Fermat en el siglo XVII, fue uno de los primeros avances significativos en estadística y matemáticas que condujo al desarrollo de la CD (Hacking, 1990).

Más adelante, en el siglo XVIII, el matemático suizo Jakob Bernoulli desarrolló el teorema central del límite y el británico Thomas Bayes formuló el teorema que lleva su nombre, el teorema de Bayes. El primero resultó fundamental en la teoría de la probabilidad y el segundo en el desarrollo de técnicas estadísticas y de aprendizaje automático, como la inferencia bayesiana y la clasificación bayesiana ingenua.

A finales del siglo XIX y principios del XX, el matemático británico Karl Pearson y el estadístico británico Ronald A. Fisher hicieron contribuciones significativas en estadística que han influido en el campo de la CD. Pearson desarrolló el coeficiente de correlación de Pearson y la regresión lineal y Fisher, por otro lado, es considerado uno de los fundadores de la estadística moderna, y sus contribuciones incluyen el diseño de experimentos, la prueba de hipótesis y la distribución de Fisher-Snedecor.

Andrey Kolmogorov, un matemático ruso, y Norbert Wiener, un matemático estadounidense, también hicieron contribuciones fundamentales en la teoría de la probabilidad y el análisis matemático en la primera mitad del siglo XX. Kolmogorov desarrolló una teoría axiomática de la probabilidad que ha sido fundamental en el estudio de los procesos estocásticos y las matemáticas aplicadas en la CD. Wiener, por su parte, es considerado el padre del ciberespacio y la teoría de la información, y su trabajo en el análisis armónico y la teoría de la predicción ha influido en el desarrollo de algoritmos y técnicas en la CD.

La aparición de la computación y la programación en el siglo XX permitió el desarrollo de técnicas estadísticas computacionales y la implementación de algoritmos. John W. Tukey, un estadístico estadounidense, fue pionero en la estadística exploratoria y en el análisis de datos, y sus ideas han influido en gran medida en el enfoque práctico y computacional de la CD.

En el campo de las matemáticas puras, podemos destacar la teoría de grafos, introducida por Leonhard Euler en el siglo XVIII, que ha sido fundamental en el estudio de las relaciones entre datos y en la modelización de redes complejas. El álgebra lineal, por su parte, desarrollada por Carl Friedrich Gauss y otros matemáticos, ha proporcionado las bases para resolver sistemas de ecuaciones lineales y ha sido esencial en el análisis de datos multidimensionales y en el aprendizaje automático (Strang, 1988).

La optimización matemática es otra área que ha contribuido significativamente al desarrollo de la CD. Los algoritmos de optimización, como el método del gradiente descendente, son fundamentales para el entrenamiento y la actualización de modelos en el aprendizaje automático y las redes neuronales (Ruder, 2016).

También el cálculo, desarrollado por Isaac Newton y Gottfried Wilhelm Leibniz, ha sido crucial en la comprensión y modelización de funciones continuas y en la derivación e integración de funciones matemáticas. Esto ha permitido el desarrollo de modelos de regresión y clasificación en la CD, así como en el análisis de series temporales y datos funcionales (Ramsay & Silverman, 2005).

En resumen, la estadística y las matemáticas han sido fundamentales en el desarrollo de la CD. Los avances en estas disciplinas han permitido la creación de técnicas y algoritmos que forman la base de la CD actual. La genealogía de la CD se puede rastrear a través de las contribuciones de matemáticos y estadísticos a lo largo de la historia, y su legado continúa siendo evidente en las aplicaciones y enfoques modernos de la CD.

## 2. Ciencia de la información.

Otra disciplina fundamental en el desarrollo de la CD es la ciencia de la información, que se ocupa de la adquisición, almacenamiento, organización, recuperación, difusión y protección de la información en sus diversas formas (Borko, 1968). Aunque la ciencia de la información se originó en la gestión de documentos y archivos, la explosión de información en el siglo XX, especialmente en forma digital, ha llevado a la convergencia de la ciencia de la información con la informática, la estadística y otros campos relacionados con la CD.

El origen de la ciencia de la información se remonta a la antigüedad, con la creación de bibliotecas y archivos para conservar y organizar documentos, como la Biblioteca de Alejandría en el siglo III a.C. (Bazerman, 1988). Sin embargo, fue en el siglo XX cuando la ciencia de la información experimentó un rápido desarrollo, impulsado por el crecimiento exponencial de la información y la aparición de tecnologías de la información y la comunicación.

En la década de 1940, la ciencia de la información comenzó a adoptar enfoques matemáticos y computacionales para la organización y recuperación de información. Claude Shannon, en su trabajo seminal sobre la teoría de la información (1948), estableció las bases matemáticas para la comprensión de la información y su transmisión en sistemas de comunicación. La teoría de la información ha sido fundamental en el desarrollo de algoritmos y técnicas de compresión de

datos, codificación de información y corrección de errores, que son esenciales en la CD. La teoría de la información también ha influido en áreas como el aprendizaje automático y la minería de datos, en particular en la selección y evaluación de características, donde la entropía y la información mutua se utilizan para medir la importancia y la relación entre las características y las variables objetivo.

La recuperación de información es otro aspecto clave de la ciencia de la información que ha influido en la CD. La recuperación de información se ocupa de la búsqueda y extracción de información relevante a partir de grandes conjuntos de datos, como colecciones de documentos, bases de datos o la World Wide Web. A partir de la década de 1950, la recuperación de información comenzó a adoptar enfoques automáticos y computacionales, como el desarrollo de algoritmos de clasificación, agrupamiento y ranking (Luhn, 1957; Salton, 1968).

La aparición de la World Wide Web en la década de 1990 transformó la ciencia de la información y llevó a la creación de nuevos enfoques y técnicas para organizar, buscar y extraer información en línea. Los motores de búsqueda, como Google, han sido fundamentales en la evolución de la ciencia de la información y la CD, al aplicar algoritmos de recuperación de información, minería de datos y aprendizaje automático para indexar y buscar información en la Web (Page et al., 1999).

La ciencia de la información también ha influido en la CD en términos de ética, privacidad y seguridad de la información. La gestión y análisis de datos a gran escala plantean desafíos éticos y legales relacionados con la privacidad, la propiedad intelectual y el uso responsable de la información. La ciencia de la información ha proporcionado marcos teóricos y prácticos para abordar estos problemas, como las políticas de privacidad, la anonimización de datos y la protección de datos personales (Sweeney, 2002; Cavoukian, 2011). La creciente conciencia de la importancia de la privacidad y la seguridad de la información ha llevado a la adopción de regulaciones y leyes, como el Reglamento General de Protección de Datos (RGPD) en la Unión Europea (EU, 2016).

La gestión del conocimiento es otro aspecto de la ciencia de la información que ha influido en la CD. La gestión del conocimiento se ocupa de la creación, organización, almacenamiento, recuperación, aplicación y transferencia de conocimiento en las organizaciones (Davenport & Prusak, 1998). La CD ha adoptado enfoques y técnicas de gestión del conocimiento, como la representación del conocimiento, la ontología y la minería de texto, para extraer información valiosa de los datos y aplicarla en la toma de decisiones y la innovación (Hearst, 1999; Gruber, 1993).

### 3. Informática y programación.

La informática y la programación también han sido fundamentales en el desarrollo de la CD, proporcionando herramientas y plataformas que permiten el procesamiento, almacenamiento, análisis y visualización de grandes volúmenes de datos. La informática es la ciencia que estudia los fundamentos teóricos de la información y la computación, así como su implementación y aplicación en sistemas computacionales. La programación, por otro lado, es el proceso de diseño, codificación, depuración y mantenimiento de código fuente de programas de computadora.

El nacimiento de la informática como disciplina se puede rastrear hasta la invención de las máquinas analíticas por Charles Babbage en el siglo XIX (Babbage, 1837). Sin embargo, fue en el siglo XX cuando la informática y la programación se convirtieron en campos fundamentales para el desarrollo de la CD. Los avances en la teoría de la computación, como la máquina de Turing (Turing, 1936) y la teoría de la información de Claude Shannon (Shannon, 1948), sentaron las bases para la creación de algoritmos y técnicas computacionales que más tarde se aplicaron en la CD.

El desarrollo de lenguajes de programación, como FORTRAN, COBOL y, más tarde, C y C++, permitió a los científicos de datos implementar algoritmos y técnicas estadísticas en programas de computadora. Estos lenguajes de programación facilitaron la creación de software para procesar y analizar datos, como los paquetes estadísticos y los sistemas de bases de datos.

La aparición de la programación orientada a objetos en la década de 1980, con lenguajes como Smalltalk y, posteriormente, Java y Python, permitió una mayor abstracción y modularidad en el diseño de software, facilitando el desarrollo de bibliotecas y herramientas especializadas en el procesamiento y análisis de datos. En la actualidad, Python es uno de los lenguajes de programación más utilizados en la CD, gracias a su facilidad de uso, versatilidad y amplio ecosistema de bibliotecas y herramientas, como NumPy, pandas, scikit-learn y TensorFlow.

Otro desarrollo clave en informática y programación que ha impulsado la CD es la creación de sistemas de gestión de bases de datos (DBMS) en la década de 1970. Los DBMS permiten el almacenamiento, la consulta y la manipulación eficiente de grandes volúmenes de datos. Los primeros DBMS se basaban en el modelo jerárquico o en el modelo de red (Tsichritzis & Lochovsky, 1982), pero fue el modelo relacional, propuesto por Edgar F. Codd (1970), el que sentó las bases para la mayoría de los sistemas de bases de datos modernos.

El modelo relacional fue revolucionario porque introdujo una forma estructurada y matemáticamente sólida de representar y manipular datos en tablas. El lenguaje de consulta estructurado (SQL) se convirtió en el estándar para interactuar con bases de datos relacionales, permitiendo a los científicos de datos extraer, transformar y cargar datos de manera eficiente (Chamberlin & Boyce, 1974). Con el tiempo, los sistemas de bases de datos han evolucionado para adaptarse a las necesidades de la CD, incluyendo el desarrollo de sistemas de bases de datos NoSQL, como MongoDB, que ofrecen mayor flexibilidad en la estructura y el almacenamiento de datos.

La computación distribuida y en paralelo también ha sido fundamental en el avance de la CD, especialmente con el crecimiento exponencial de los datos disponibles. Técnicas de computación distribuida, como MapReduce, han permitido procesar y analizar grandes conjuntos de datos en clústeres de computadoras, proporcionando escalabilidad y eficiencia en el análisis de big data. Frameworks como Hadoop y Spark han sido desarrollados para facilitar la implementación de algoritmos de procesamiento y análisis de datos en entornos distribuidos.

La visualización de datos es otra área en la que la informática y la programación han influido significativamente en la CD. La visualización es esencial para comprender y comunicar los resultados del análisis de datos, y se han desarrollado numerosas herramientas y bibliotecas para facilitar la creación de gráficos y visualizaciones interactivas. Entre las más destacadas se encuentran D3.js, ggplot2 y Tableau.

#### 4. Minería de datos y aprendizaje automático.

La minería de datos y el aprendizaje automático son dos áreas estrechamente relacionadas que han impulsado el desarrollo de la CD y se han convertido en componentes clave en la extracción de conocimientos a partir de grandes conjuntos de datos.

La minería de datos es un proceso que involucra la exploración, análisis y descubrimiento de patrones útiles en grandes conjuntos de datos. La minería de datos utiliza enfoques de estadística, inteligencia artificial, aprendizaje automático y visualización de datos para identificar patrones y relaciones subyacentes en los datos que pueden ser útiles para la toma de decisiones y la generación de conocimiento (Han et al., 2011).

El aprendizaje automático, por otro lado, es un subcampo de la inteligencia artificial que se centra en el desarrollo de algoritmos y modelos que permiten a las computadoras aprender de los datos y mejorar su desempeño en tareas específicas sin ser programadas explícitamente

(Samuel, 1969; Mitchell, 1997). El aprendizaje automático utiliza técnicas de minería de datos, estadísticas y optimización para construir y entrenar modelos que pueden predecir, clasificar, agrupar y generar recomendaciones basadas en datos de entrada.

El auge de la minería de datos y el aprendizaje automático en las últimas décadas ha sido impulsado por factores como:

- El crecimiento exponencial de datos generados por dispositivos digitales, sensores, redes sociales y otras fuentes, lo que ha llevado a una mayor demanda de técnicas y herramientas capaces de analizar y extraer información valiosa de estos datos (Manyika et al., 2011).
- El avance en las tecnologías de computación y almacenamiento, como la computación en la nube, la computación distribuida y el almacenamiento de big data, que han hecho posible procesar y analizar grandes conjuntos de datos de manera eficiente y rentable (Dean & Ghemawat, 2008).
- La aparición de nuevas técnicas y algoritmos de aprendizaje automático, como las redes neuronales profundas (Hinton et al., 2012), los modelos de ensamble (Breiman, 2001) y los métodos no supervisados (von Luxburg, 2007), que han mejorado significativamente el rendimiento y la aplicabilidad de los modelos de aprendizaje automático en una amplia variedad de dominios y tareas.

La minería de datos y el aprendizaje automático han llevado a importantes avances y aplicaciones en la CD, como la personalización de contenidos y publicidad, la detección de fraudes, el diagnóstico médico, la predicción del comportamiento del consumidor, la identificación de tendencias en redes sociales, la toma de decisiones basada en datos, entre otros. Algunos de los principales métodos y técnicas de minería de datos y aprendizaje automático utilizados en la CD incluyen:

- **Clasificación:** La clasificación es una tarea de aprendizaje supervisado que consiste en asignar etiquetas o categorías a objetos o instancias de datos basándose en un conjunto de entrenamiento etiquetado.
- **Regresión:** La regresión es otra tarea de aprendizaje supervisado que implica predecir una variable continua en función de una o más variables de entrada. Los métodos de regresión más comunes incluyen la regresión lineal, la regresión logística y la regresión de soporte vectorial.
- **Agrupamiento (clustering):** El agrupamiento es una tarea de aprendizaje no supervisado que consiste en dividir un conjunto de datos en grupos o clústeres de objetos similares.

- Reducción de la dimensionalidad: La reducción de la dimensionalidad es un enfoque utilizado para simplificar datos de alta dimensión, manteniendo la mayor cantidad de información posible.
- Aprendizaje por refuerzo: El aprendizaje por refuerzo es un enfoque de aprendizaje automático en el que un agente aprende a tomar decisiones en un entorno mediante la interacción y la recepción de retroalimentación en forma de recompensas o castigos (Sutton & Barto, 1998).
- Redes neuronales profundas: Las redes neuronales profundas son un tipo de modelo de aprendizaje automático que utiliza múltiples capas de neuronas artificiales para aprender representaciones jerárquicas de datos y realizar tareas de clasificación, regresión y generación (LeCun et al., 2015). Las redes neuronales profundas han revolucionado la CD en áreas como el reconocimiento de imágenes, el procesamiento del lenguaje natural y la traducción automática (Krizhevsky et al., 2012; Sutskever et al., 2014).

La minería de datos y el aprendizaje automático han sido fundamentales en el desarrollo de la CD, proporcionando herramientas y técnicas poderosas para analizar y extraer información valiosa de grandes conjuntos de datos. La creciente disponibilidad de datos y la mejora de las tecnologías de computación han permitido la aplicación de la minería de datos y el aprendizaje automático en una amplia variedad de dominios y tareas, lo que ha llevado a avances significativos y a la transformación de industrias, como la atención médica, las finanzas, el marketing, la energía, el transporte y la educación, entre otras.

## Evolución de la Ciencia de Datos

### 1. Concepto de Ciencia de Datos.

A finales del siglo XX, la cantidad de datos almacenados por las empresas y las organizaciones públicas ya era demasiado grande como para seguir siendo tratada por la ciencia de la información de la que hemos hablado. La unidad de medida ya no era la información en sentido amplio, sino el dato, la unidad mínima de información. La materia tratada era cada vez mayor y más compleja, y las herramientas para procesarla también se hicieron más sofisticadas.

Ya en 1962, John W. Tukey anticipaba los cambios que estaban empezando a producirse en el campo de la información en su artículo *The Future of Data Analysis*, en el que expone la necesidad de separar la estadística en general y el análisis de datos en particular de las

matemáticas que los englobaba, pues desde el surgimiento de las computadoras modernas estos dos ámbitos de estudio habían evolucionado hasta poder ser considerados una ciencia empírica autónoma (Tukey, 1962). Más adelante, en 1977, Tukey insiste en este punto argumentando que las herramientas de análisis de datos pueden y deben emplearse para la formulación de hipótesis (Tukey, 1977). En el mismo año se funda la Asociación Internacional de Computación Estadística como una sección del Instituto Internacional de Estadística, con el objetivo de vincular la metodología estadística tradicional, la tecnología informática y el conocimiento de expertos en cualesquiera dominios en los que se quiera convertir los datos en conocimiento.

En los años noventa el fenómeno comienza a ser imparable y los distintos medios ya hablan de empresas que recogen información sobre los clientes para adaptar el marketing a los gustos del público (BusinessWeek, 1994). Las instituciones académicas también hablan de la CD, la minería de datos y el KDD (Descubrimiento de conocimiento en bases de datos) a partir de algoritmos capaces de extraer patrones en los datos. Esta fiebre de los datos llega hasta el punto de que, en la conferencia inaugural para la Cátedra de Estadística en la Universidad de Michigan por el profesor Jeff Wu en 1997, se sugiera la supresión de la estadística en favor de la CD, de manera que los estadísticos pasasen a llamarse científicos de datos.

Esta confusión de términos es abordada dos años después por Jacob Zahavi, profesor en la Universidad de Pennsylvania, que, comenzando a palpase ya la era digital en la que nos encontramos, establece la diferencia entre las disciplinas en los conjuntos de datos que estas trabajan. Los métodos estadísticos convencionales funcionan con pequeños conjuntos de datos, mientras que los grandes conjuntos de datos requieren de herramientas especiales que sean capaces de detectar relaciones no lineales e interacciones entre variables (Zahavi, 1999). Así, en 2001, Leo Breiman en su artículo titulado *Modelos estadísticos: las dos culturas*, explica que existen dos tipos de modelos de datos a través de los cuales obtener conclusiones: modelos estocásticos, empleados por los estadísticos, y modelos algorítmicos, empleados por científicos de datos (Breiman, 2001). Los modelos algorítmicos están especialmente preparados para enfrentarse a bases de datos masivas, con tipos de datos complejos y relaciones no lineales.

En este punto, podemos apreciar una primera separación de la estadística respecto de la matemática y una segunda separación de la CD respecto de la estadística. El desarrollo de las teorías de la información y el aumento en la generación de datos durante la segunda mitad del siglo XX, hicieron imposible mantener la estadística englobada en una ciencia como la matemática cuyo objeto de estudio ya no compartía. De igual manera, el progresivo desarrollo de herramientas de análisis de datos y métodos algorítmicos, apoyados en la computación,

dividen la estadística en las dos escuelas que Breiman expone. La CD, aunque comparta con la estadística la misión de obtener conocimiento desde conjuntos de datos, los materiales que trabaja y las herramientas que emplea ya no son las mismas.

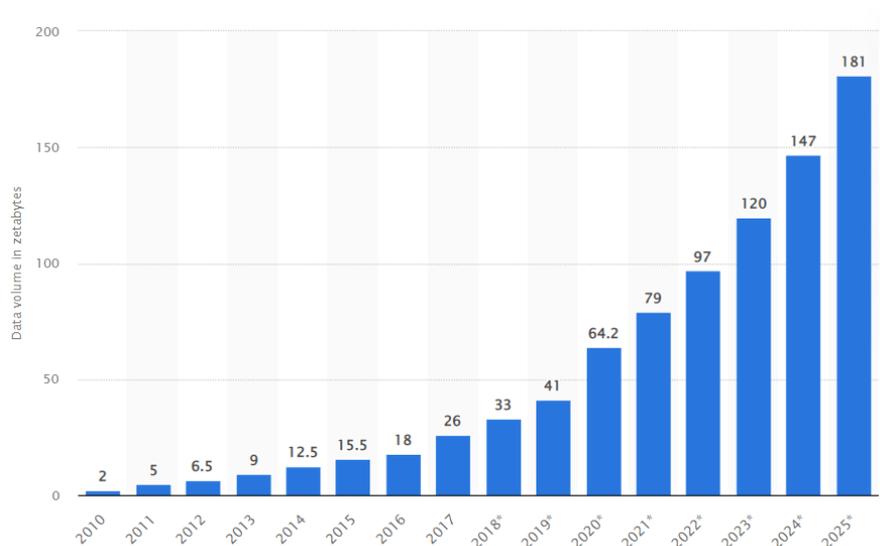
En el siglo XXI podemos encontrar menciones de la CD en los medios continuamente. El dato se ha convertido en el principal motor de la economía mundial y los empleos dedicados a su tratamiento son cada vez más demandados por empresas de todo tipo. Pero el impacto de la CD no se queda en el aspecto económico, ni mucho menos. En el año 2009, el científico de la computación Jim Gray publica póstumamente en la recopilación titulada *The Fourth Paradigm: Data-Intensive Scientific Discovery* un ensayo en el cual propone la idea de la CD como el cuarto paradigma científico en la historia de la ciencia. Según Jim Gray, el paradigma se define como el método en el que se lleva a cabo el desarrollo científico y a lo largo de la historia se han sucedido tres paradigmas diferentes. El primer paradigma corresponde al empirismo, a través del cual los científicos describían los fenómenos naturales en base a la observación. El segundo paradigma corresponde a la rama teórica de la ciencia, a través de la cual los científicos establecían modelos para generalizar los distintos fenómenos que se estudian. El tercer paradigma es el computacional, mediante el cual podemos crear simulaciones complejas que brindan información imposible de obtener en el mundo físico. A estos tres paradigmas Jim Gray añade la CD como la nueva forma de hacer ciencia. Desde todas las ciencias se originan miles de millones de datos que son empleados por los científicos de datos para obtener patrones de comportamiento, realizar predicciones y clasificar fenómenos, entre muchas otras posibilidades (Gray, 2009).

De esta manera, regresamos a la definición ofrecida anteriormente para la CD, como un campo interdisciplinar y transdisciplinar, que trabaja directamente con y sobre los datos generados por un amplísimo abanico de fenómenos, con el objetivo de encontrar patrones y relaciones que puedan ser explotados en beneficio del proyecto al que se dirijan.

## 2. La era del Big Data.

La era del Big Data se refiere al período en el que la CD experimentó un crecimiento significativo y una rápida evolución debido al aumento exponencial en la generación, almacenamiento y análisis de grandes volúmenes de datos. Este fenómeno se atribuye en gran medida a la digitalización de la información, el auge de las redes sociales, el Internet de las cosas (IoT) y la disponibilidad de tecnologías de almacenamiento y procesamiento de datos más económicas y eficientes.

Volumen de datos/información creada, capturada, copiada, y consumida en todo el mundo desde 2010 hasta 2020, con predicción desde 2021 hasta 2025(en zettabytes). Fuente: <https://www.statista.com/statistics/871513/worldwide-data-created/>



El término "Big Data" se utiliza para describir conjuntos de datos que son tan grandes, diversos y complejos que requieren técnicas y herramientas de procesamiento de datos avanzadas para su análisis y extracción de información útil. Las características principales del Big Data se describen mediante las "5 V": Volumen, Variedad, Velocidad, Veracidad y Valor.

- **Volumen:** Se refiere a la cantidad de datos generados y almacenados. El Big Data implica el manejo de terabytes, petabytes o incluso exabytes de datos.
- **Variedad:** Se refiere a la diversidad de tipos de datos, que incluyen datos estructurados, semiestructurados y no estructurados, como texto, imágenes, videos, registros de sensores y datos de redes sociales.
- **Velocidad:** Se refiere a la rapidez con la que se generan, procesan y analizan los datos. El Big Data a menudo implica la necesidad de procesar y analizar datos en tiempo real o casi en tiempo real.
- **Veracidad:** Se refiere a la calidad y precisión de los datos. Dado que los datos pueden provenir de diversas fuentes y en diversos formatos, garantizar su calidad y confiabilidad es un desafío importante en el análisis de Big Data.
- **Valor:** Se refiere a la capacidad de extraer información útil y valiosa de los datos para tomar decisiones informadas y obtener ventajas competitivas.

El análisis de Big Data presenta una serie de desafíos, como la necesidad de desarrollar técnicas y herramientas de procesamiento de datos más eficientes, garantizar la calidad y confiabilidad de los datos, proteger la privacidad y seguridad de los datos y garantizar la interpretabilidad y explicabilidad de los modelos y resultados de análisis.

La era del Big Data ha impulsado el desarrollo de una serie de tecnologías y técnicas de análisis de datos, como:

- Sistemas de almacenamiento y procesamiento de datos distribuidos: Hadoop y Apache Spark son ejemplos de sistemas de código abierto diseñados para almacenar y procesar grandes volúmenes de datos distribuidos en múltiples nodos o servidores. Estas tecnologías han permitido a las organizaciones procesar y analizar Big Data de manera eficiente y escalable.
- Bases de datos NoSQL: Las bases de datos NoSQL, como MongoDB, Cassandra y Couchbase, han sido desarrolladas para manejar la variedad y complejidad de los datos en entornos de Big Data. Estas bases de datos son particularmente adecuadas para almacenar y gestionar datos semiestructurados y no estructurados.
- Aprendizaje automático y técnicas de inteligencia artificial: El análisis de Big Data ha impulsado el desarrollo y la adopción de técnicas avanzadas de aprendizaje automático e inteligencia artificial, como las redes neuronales profundas, el aprendizaje por refuerzo y el aprendizaje federado. Estas técnicas han permitido a los investigadores y profesionales construir modelos altamente efectivos para abordar problemas complejos y desafiantes en una amplia gama de campos y aplicaciones.
- Visualización y análisis de datos: La era del Big Data ha llevado al desarrollo de herramientas y técnicas de visualización y análisis de datos más sofisticadas, como Tableau, D3.js y ggplot2. Estas herramientas han permitido a los científicos de datos explorar, analizar y comunicar resultados de manera más eficaz y atractiva.

Al mismo tiempo, el análisis de Big Data ofrece oportunidades significativas en una amplia variedad de campos y aplicaciones, como la medicina personalizada, la optimización de la cadena de suministro, la detección y prevención de fraudes, el monitoreo y predicción de eventos climáticos extremos y la toma de decisiones basada en datos en organizaciones y gobiernos.

El Big Data también ha tenido un impacto significativo en la sociedad, ya que ha permitido a las organizaciones y gobiernos tomar decisiones más informadas y basadas en datos en una amplia variedad de campos y aplicaciones. Sin embargo, también ha planteado preocupaciones sobre la privacidad, la seguridad y la ética en el uso y análisis de datos personales y sensibles.

Sobre el impacto del Big Data sobre los diferentes aspectos de la sociedad actual nos extenderemos más adelante. Por ahora, basta con conocer en qué consiste el fenómeno de dicho nombre y las principales herramientas que se han desarrollado para abordar su tratamiento. También es importante resaltar la forma en que una disciplina como la CD está en constante evolución, hecho de especial importancia cuando su campo de trabajo son una cantidad de datos que en esta sociedad de la información está experimentando un crecimiento exponencial.

### 3. Desarrollo de herramientas y algoritmos.

#### Lenguajes de programación.

Los lenguajes de programación han sido fundamentales en el desarrollo y la evolución de la CD. Algunos de los lenguajes de programación más populares y ampliamente utilizados en la CD incluyen:

- R: R es un lenguaje de programación y un entorno de software para el análisis estadístico y la visualización de datos (R Core Team, 2020). R se ha convertido en una herramienta esencial en la CD debido a su amplia gama de paquetes y funciones que permiten a los científicos de datos realizar análisis estadísticos, visualizar datos y desarrollar modelos de aprendizaje automático.
- Python: Python es un lenguaje de programación de alto nivel que ha ganado popularidad en la CD debido a su simplicidad y facilidad de uso (Van Rossum & Drake, 2009). Python también cuenta con una gran cantidad de bibliotecas, como NumPy, pandas y scikit-learn, que proporcionan funcionalidades para el análisis y la manipulación de datos, la visualización y el aprendizaje automático.
- Julia: Julia es un lenguaje de programación de alto nivel y alto rendimiento para la informática técnica (Bezanson et al., 2017). Julia ha sido diseñado específicamente para abordar los desafíos del procesamiento de datos y la computación numérica y se ha vuelto cada vez más popular en la CD debido a su capacidad para combinar la simplicidad y la productividad de Python y R con el rendimiento de lenguajes de programación de bajo nivel como C y Fortran.

#### Algoritmos de aprendizaje automático y minería de datos.

El desarrollo y la mejora de algoritmos de aprendizaje automático y minería de datos han sido fundamentales para el avance de la CD. Algunos de los algoritmos más influyentes y ampliamente utilizados en la CD incluyen:

- **Regresión lineal y logística:** Estos algoritmos son técnicas estadísticas básicas que se utilizan para modelar relaciones entre variables y realizar predicciones (Freedman, 2006). La regresión lineal y logística han sido fundamentales en la CD como métodos iniciales para abordar problemas de predicción y clasificación.
- **Árboles de decisión y ensambles de árboles:** Los árboles de decisión son algoritmos que permiten dividir los datos en subconjuntos basados en reglas basadas en las características de los datos (Quinlan, 1986). Los ensambles de árboles, como Random Forest y Gradient Boosting (Friedman, 2001), combinan múltiples árboles de decisión para mejorar la precisión y la robustez de las predicciones. Estos algoritmos se han vuelto populares en la CD debido a su interpretabilidad y capacidad para manejar datos de alta dimensión y complejidad.
- **Máquinas de vectores de soporte (SVM):** Las SVM son algoritmos de aprendizaje supervisado que pueden ser utilizados para problemas de clasificación y regresión (Cortes & Vapnik, 1995). Las SVM han sido ampliamente utilizadas en la CD debido a su capacidad para manejar datos de alta dimensionalidad y separar clases no lineales a través del uso de funciones kernel.
- **Redes neuronales artificiales y aprendizaje profundo:** Las redes neuronales artificiales son algoritmos inspirados en el funcionamiento del cerebro humano que pueden aprender y adaptarse a partir de datos de entrada (McCulloch & Pitts, 1943). El aprendizaje profundo, que implica el uso de redes neuronales con múltiples capas ocultas, ha revolucionado la CD al permitir la construcción de modelos altamente efectivos para tareas como el reconocimiento de imágenes, el procesamiento del lenguaje natural y la generación de texto (LeCun et al., 2015).
- **Algoritmos de clustering:** Los algoritmos de clustering, como K-means (MacQueen, 1967) y DBSCAN (Ester et al., 1996), son técnicas de aprendizaje no supervisado que permiten identificar grupos o clusters en los datos basados en la similitud entre las observaciones. Estos algoritmos son fundamentales en la CD para tareas como la segmentación de clientes, la detección de anomalías y el análisis de redes sociales.

#### Herramientas y plataformas de análisis de datos:

Además de los lenguajes de programación y algoritmos, el desarrollo de herramientas y plataformas de análisis de datos ha sido crucial para el avance de la CD. Algunas de las herramientas y plataformas más populares incluyen:

- **Jupyter Notebook:** Jupyter Notebook es una aplicación web de código abierto que permite crear y compartir documentos que contienen código en vivo, ecuaciones, visualizaciones y texto narrativo (Kluyver et al., 2016). Esta herramienta se ha vuelto esencial en la CD como una plataforma para la exploración interactiva de datos, el desarrollo de modelos de aprendizaje automático y la comunicación de resultados.
- **Apache Hadoop y Spark:** Como hemos mencionado en la sección 2.3, Apache Hadoop y Spark son sistemas de código abierto para el almacenamiento y procesamiento de datos distribuidos que han permitido a las organizaciones escalar y acelerar sus análisis de datos.
- **TensorFlow y PyTorch:** TensorFlow y PyTorch son bibliotecas de código abierto para el aprendizaje automático y el aprendizaje profundo que han ganado popularidad en la CD debido a su flexibilidad, rendimiento y facilidad de uso. Estas bibliotecas han facilitado el desarrollo y la implementación de modelos de aprendizaje profundo en una amplia gama de aplicaciones y campos.
- **Plataformas en la nube:** Las plataformas en la nube, como Amazon Web Services (AWS), Google Cloud Platform y Microsoft Azure, han permitido a los científicos de datos acceder a recursos de computación y almacenamiento escalables y de bajo costo. Estas plataformas también ofrecen herramientas y servicios específicos para la CD, como el almacenamiento y procesamiento de datos en tiempo real, el aprendizaje automático y la inteligencia artificial.

Todo este conjunto de herramientas, plataformas y algoritmos, entre muchas otras, son lo que componen el cuerpo de la CD. Podemos ver, de esta manera, ese carácter práctico y aplicado con que hemos definido a la disciplina previamente. La CD apenas cuenta con un desarrollo teórico propio, pues reutiliza y reinterpreta los materiales pertenecientes a los campos con los que trabaja, siendo estos los ya citados (estadística, computación, ciencia de la información) y aquellos sobre los que se trabaje directamente (biología, economía, psicología...). Son estos avances que hemos enumerado los que han permitido a los científicos de datos abordar problemas complejos y desafiantes en una amplia gama de campos y aplicaciones, al tiempo que han impulsado la innovación y el crecimiento en la CD. A medida que la CD continúa evolucionando, es probable que sigamos viendo nuevos desarrollos en herramientas y algoritmos que permitirán a los científicos de datos abordar desafíos emergentes y aprovechar nuevas oportunidades en el análisis de datos.

### 3. Ontología de la Ciencia de Datos

Ya conocemos los antecedentes de la CD y el estado actual en que esta disciplina se encuentra. El siguiente paso que debemos dar para llegar a un conocimiento lo suficientemente exhausto del fenómeno como para poder identificar y valorar el impacto que este tiene sobre las distintas capas de nuestra sociedad actual es saber a qué nos estamos refiriendo cuando hablamos de CD. A lo largo del anterior apartado hemos podido apreciar los distintos componentes que, agrupados, dan forma a la CD: herramientas, tecnologías, algoritmos, teorías... Pero, ¿sabemos qué es lo que todo esto forma? ¿Es la CD una ciencia, como su nombre parece indicarnos? ¿Es quizás un conjunto de tecnologías? A lo largo del presente epígrafe trataremos de responder a la pregunta ¿Qué es la CD?, primeramente, exponiendo la idea de ciencia que aquí suscribimos. En función de cómo la CD se adapta a nuestra definición de Ciencia, podremos determinar si lo es o no.

#### ¿Qué es la ciencia?

La pregunta que nos ocupa en el presente apartado ha ido adquiriendo a lo largo de la historia diversas respuestas en razón del ámbito o el momento en que se haya abordado. Una concepción de ciencia que podemos oír en determinados contextos cotidianos es la idea de ciencia como "saber hacer", también sinónima de arte en su sentido más vulgar, que escuchamos en ocasiones como "la ciencia del zapatero" o "el arte del navegante" para referir al conjunto de capacidades de unos y otros profesionales. Esta es una concepción mundana, no académica, de la ciencia como conjunto de saberes, que se opone, por ejemplo, a una segunda concepción de la ciencia como sistema ordenado de proposiciones derivadas de principios. Esta segunda concepción, al contrario que la primera asociada al taller, se asocia a la palabra, los libros y las lecciones impartidas en academias para desarrollar una ciencia basada en las deducciones lógicas a partir de principios inamovibles de carácter geométrico, teológico o filosófico. Esta concepción equivaldría a la impulsada por Aristóteles, en el siglo IV a.C., y desarrollada por los escolásticos hasta la revolución científica en el siglo XVI, cuando surge el laboratorio como simbiosis del taller y la academia. En el laboratorio se llevan a cabo estudios fenomenológicos, se realizan modelos, observaciones y se desarrollan teorías a través de las cuales las ciencias avanzan en su propósito de conocer el mundo. Esta será la concepción que tomará la delantera a partir del siglo XVIII hasta nuestros días y la que tomaremos como referencia en nuestro estudio (Bueno, 1995).

Siguiendo a Gustavo Bueno y su obra "¿Qué es la ciencia?", al centrarnos en esta tercera concepción de la ciencia, las conocidas como ciencias positivas, nos encontramos con que de nuevo no tenemos una única idea de ciencias positivas, sino tres que se clasifican en función de la configuración de materia y forma en la construcción de la verdad científica. Para Bueno, la materia corresponde al campo de una ciencia, y la forma a la estructura formada por las unidades que conforman el campo. Así, materia y forma desempeñan papeles holóticos en la construcción de las ciencias, y las distintas configuraciones entre estos dos elementos que se han desarrollado a lo largo de la historia comprenden las distintas teorías de la ciencia que Bueno expone.

Se suele establecer el inicio de la reflexión sobre la naturaleza del conocimiento científico a inicios del siglo XX, con la formación del Círculo de Viena, que enfrentándose a la visión adecuacionista, vigente desde Galileo hasta Einstein, proponen una visión descripcionista. Mientras que el adecuacionismo sostiene que la verdad científica se logra cuando las teorías (forma) se ajustan a la realidad empírica (materia), asumiendo una relación de isomorfismo entre materia y forma; el descripcionismo establece como verdad científica a la materia misma, quedando las formas de la ciencia con un papel descriptivo o representativo. Los descripcionistas rechazan las hipótesis como parte de las ciencias, que son constituidas únicamente por las constataciones. Como alternativa a estas dos visiones, Karl Popper, en lo que llamaríamos la visión teoreticista, propone que la verdad científica siempre se basa en lo hipotético y esto solo puede ser refutado, pero no demostrado definitivamente. De esta manera, la forma adquiriría un papel principal en la definición de la verdad científica (Bueno, 1995).

Gustavo Bueno analiza estas tres corrientes de pensamiento y subraya la imposibilidad de mantener una dicotomía definida entre hechos y teorías. No existen hechos desprovistos de teoría, ni teorías que no estén basadas en observación, medición y experimentación. La práctica y la teoría están intrínsecamente vinculadas, y no hay hechos ni teorías "puras". Ante esto, Bueno propone una teoría circularista de la ciencia que desarrollará a lo largo de su obra *Teoría del cierre categorial*, que será nuestra referencia para determinar si la CD es o no una ciencia.

### 1. Circularismo.

El circularismo se opone a las concepciones adecuacionista, descripcionista y teoreticista de la ciencia, negando la posibilidad de conectar materia y forma con el fin de alcanzar la verdad científica, por el hecho de que no concibe estas partes como distintas y separadas. Para el circularismo, las partes formales de la ciencia son los propios nexos concatenados de las partes materiales, siendo estas conexiones el contenido mismo de la ciencia. De esta manera, las formas

no se crean a posteriori de la materia tratando de explicarla, como entendían los descripcionistas; tampoco a priori subordinando la materia a las construcciones formales previamente elaboradas, como pensaban los teoreticistas; para el circularismo, las partes formales (los teoremas, por ejemplo) son el resultado de la concatenación entre las partes materiales, que cuando forman uniones objetivas y necesarias crean la verdad científica.

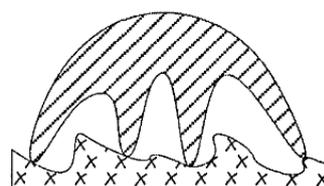
De esta manera, a través de los sucesivos nexos entre partes materiales, se va creando lo que Bueno denomina el cierre categorial, que delimita el campo de una ciencia. Este cierre categorial no es gratuito, pues no se establece en función de una teoría que no tiene en cuenta la materia, sino que es interno a la constitución de la ciencia. A medida que van formándose verdades científicas o teoremas, al concatenarse las partes materiales, la ciencia va organizándose alrededor de estos núcleos, que a su vez se relacionan entre ellos.

Estos campos categoriales deben entenderse como cerrados, en tanto que los materiales que los conforman se relacionan con formas del mismo campo categorial y no de otros; pero no clausurados, en tanto que nuevas formas dentro del mismo campo categorial pueden surgir merced a la aparición de sujetos materiales nuevos. De esta manera la teoría circularista de Bueno tiene en cuenta el desarrollo histórico de las ciencias.

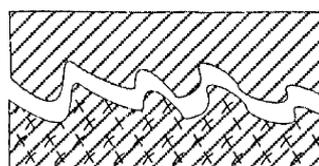
También destaca Bueno la importancia de los instrumentos y operaciones científicas en la construcción del campo categorial de una ciencia, pues solo mediante el trabajo técnico y tecnológico se pueden establecer relaciones entre determinadas partes materiales. Además, el desarrollo del campo categorial también propicia nuevos avances tecnológicos, que pueden exceder los límites del mismo campo categorial condicionando otras partes del mundo. La teoría circularista, por tanto, explica la ciencia como constructivista, pues no solo explica el mundo, sino que lo crea a medida que avanza.



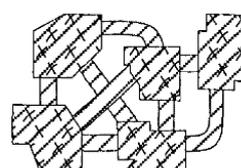
*Descripciónismo*



*Teoreticismo*



*Adecuacionismo*



*Circularismo*

## 2. Teoría del Cierre Categorial.

### Las partes formales y las partes materiales en la TCC

Una vez aceptamos la acepción circularista como la más apta para explicar las ciencias tal y como hoy se nos presentan, podemos adentrarnos en la teoría de la ciencia desarrollada por el filósofo Gustavo Bueno en su obra *¿Qué es la ciencia?* (1995) y a lo largo de los cinco tomos de su obra *Teoría del cierre categorial* (1978).

En su Teoría del cierre categorial, Gustavo Bueno, asumiendo una concepción circularista de la ciencia, señala dos clases constituyentes de los cuerpos de las ciencias, a saber, las partes materiales y las partes formales, que se definen de la siguiente manera:

- Las partes formales son aquellas que conservan y reproducen el todo de la ciencia a través de la ordenación de las partes materiales de modo que la categoría científica pueda mantener su unidad y recursividad.
- Las partes materiales son partes genéricas previas a la formalización de las ciencias, que son incorporadas al cuerpo de una ciencia cuando los científicos crean conexiones y relaciones entre materialidades.

De esta manera, tenemos que existen una serie de materialidades que solo pasan a formar parte de un campo científico cuando comienzan a establecerse relaciones entre ellas a través de verdades o teoremas, que inevitablemente deben surgir sobre la marcha. Estas materialidades preexistentes conforman lo que la TCC denomina el "contexto determinante", y solo cuando los fenómenos que lo conforman son formalizados pasan a formar parte del campo categorial de la ciencia en lo que se conoce como su "contexto determinado". Los teoremas, por tanto, no son realidades preexistentes, estos solo pueden crearse en momentos precisos del campo categorial, en contextos determinados específicos. El teorema de Pitágoras, por ejemplo, no es una realidad matemática preexistente, pues necesitó que el campo categorial se encontrara en un contexto determinado preciso, es decir, necesitaba que el campo categorial incluyera el concepto de triángulo para ser formulado. He aquí donde encontramos el carácter constructivista de la ciencia.

La idea de preexistencia de las realidades materiales implica también que podamos encontrarnos que los campos categoriales de dos o más ciencias incorporen un mismo objeto de estudio. El cuerpo humano puede ser estudiado desde la anatomía, si tenemos en cuenta sus partes; desde la medicina, si tenemos en cuenta las enfermedades que lo afectan; e incluso desde la física, si lo consideramos un objeto sometido a cierto grado de aceleración. Serán entonces las partes

formales de cada una de estas ciencias las que aborden el objeto a escalas diferentes en cada caso, descomponiendo el objeto en partes materiales que se incorporan a los campos categoriales. Gustavo Bueno, por tanto, entiende las ciencias como construcciones que incorporan parcelas de la realidad, las cuales no pueden ser entendidas por otras ciencias. Desde la anatomía no puede explicarse el comportamiento de un cuerpo humano sometido a aceleración, igual que desde la física no puede explicarse la constitución del cuerpo humano. Sin embargo, sí pueden establecerse relaciones ente teoremas de uno y otro campo categorial, pero esto no puede llevarnos a pensar que la anatomía pueda reducirse a la física o viceversa. Por tanto, también hay que tener en cuenta que la construcción de teoremas actualiza los contextos determinantes no solo de su campo categorial, sino de todos los demás, que quizás puedan establecer una relación entre esa nueva realidad y las partes materiales de su campo.

#### Espacio gnoseológico:

Una vez comprendida la organización de las ciencias en partes materiales y partes formales que se relacionan construyendo un campo categorial a partir de contextos determinantes que van sucediéndose y desarrollándose, podemos continuar con nuestra exposición analizando la composición del campo de la ciencia.

Según la TCC, los cuerpos de las ciencias, lo que Gustavo Bueno denomina el espacio gnoseológico, pueden clasificarse según una estructura trimembre formada por: 1) contenidos de carácter objetual, 2) contenidos de carácter subjetual y 3) contenidos simbólicos. Para establecer esta clasificación, Gustavo Bueno se basa en las teorías del lenguaje de Karl Buhler y Charles Morris, que, aunque independientes, exponen tres funciones propias del lenguaje, a saber, una función semántica cuando los signos se relacionan con los significados, una función pragmática cuando los signos se relacionan con los sujetos que los usan, y una función sintáctica cuando los signos se relacionan con otros signos. Gustavo Bueno establece una analogía entre las funciones del lenguaje y la clasificación de los contenidos del campo científico.

En los campos científicos encontramos estos tres órdenes: signos que se refieren a objetos, los objetos a los que se refieren los signos y sujetos que emplean los signos para referirse a los objetos. Entre estos tres ejes pueden establecerse relaciones binarias, siempre mediadas por el tercer componente sobrante. Decimos que estas relaciones deben estar mediadas porque si no la relación sería imposible. Volviendo a la analogía con el lenguaje, vemos que cuando se establece una relación entre un sujeto y un objeto es necesario que el sujeto emplee un signo para referenciar al objeto, y cuando la relación se da entre objeto y signo también será necesario un sujeto que emplee el signo para relacionar ambos. En ambos casos las relaciones son binarias,

pero están mediadas. Las únicas relaciones que no necesitan de mediación serían las que ocurren dentro de un mismo eje: signos con signos, objetos con objetos y sujetos con sujetos. El resultado de estas relaciones da lugar a nueve figuras.

Antes de exponer las nueve figuras que componen estos ejes, es importante tener en cuenta dos aspectos de ellos. En primer lugar, estos ejes son internos al espacio gnoseológico de la ciencia, no se aplican desde fuera tratando de buscar una explicación a los componentes de la ciencia (como haría el adecuacionismo), sino que estos mismos son los componentes que conforman el campo. En segundo lugar, los ejes son independientes esencialmente, pues tienen estructuras compositivas diferentes, pero no independientes existencialmente, pues los tres ejes están presentes en todo momento en las ciencias.

Ahora bien, ya sabiendo en qué consisten los ejes gnoseológicos y la forma en que construyen un espacio gnoseológico de tres dimensiones, debemos ver las figuras que componen los tres ejes:

- Eje sintáctico: Relación entre signos.
  - Términos: Relaciones entre signos mediadas por objetos.
  - Operaciones: Relaciones entre signos mediadas por sujetos.
  - Relaciones: Relaciones entre signos sin mediación.
- Eje semántico: Relación entre objetos.
  - Referenciales: Relaciones entre objetos mediadas por signos.
  - Fenómenos: Relaciones entre objetos mediadas por sujetos.
  - Estructuras: Relaciones entre objetos sin mediación.
- Eje pragmático:
  - Autologismos: Relaciones entre sujetos mediadas por signos.
  - Dialogismos: Relaciones entre sujetos mediadas por objetos.
  - Normas: Relaciones entre sujetos sin mediación.

Presente el esquema de los distintos ejes que conforman el espacio gnoseológico y las figuras que los constituyen, continuamos ofreciendo una definición de cada una de estas:

Figuras del eje sintáctico. Énfasis en los símbolos.

- **Términos:** Son la formalización de las partes objetuales y corpóreas de un campo científico sobre las que operan los sujetos. Estos términos pueden ser simples si no se componen de términos más pequeños dentro de su campo, o complejos si se componen de términos más pequeños. De esta manera, las ciencias establecen clases que identifican y delimitan los términos para hacerlos susceptibles de operaciones y relaciones semánticas, lo cual también implica asignar a los términos nombre mediante signos. Por supuesto, los términos son múltiples, no puede existir una ciencia con un solo término, por eso la TCC rechaza la idea de objeto de la ciencia por la idea de campo de la ciencia, este último formado por varios objetos. Los términos, además, se asocian a definiciones que delimitan y los constituyen. Ejemplos de términos son los elementos químicos en la Química, los doce sonidos de la escala musical en Teoría de la Música o el ADN en Genética.
- **Operaciones:** Son las transformaciones que un sujeto operatorio realiza sobre uno o más términos conceptualizados de un campo científico dando como resultado términos, más simples o más complejos, siempre dentro del mismo nivel. Estas operaciones siempre implican una manipulación física de los términos, ya sea por un sujeto humano o por instrumentos científicos controlados o programados por sujetos humanos. Un ejemplo de operaciones en el campo de las matemáticas serían las divisiones y las multiplicaciones, la primera de ellas creando términos más simples y la segunda más complejos.
- **Relaciones:** Son relaciones que se establecen entre términos definidos simbólicamente y que se predicen proposicionalmente. Esto es, las relaciones permiten conectar de manera racional y lógica los términos del campo categorial. Al contrario que las operaciones, las relaciones no crean otros términos, sino que crean materialidades de nivel superior como son las fórmulas. En física, la fórmula ( $F = m \cdot a$ ), que establece una igualdad entre la fuerza de un cuerpo y la multiplicación de su masa por su aceleración es una relación que, como vemos, está en un nivel superior que los términos que relaciona porque un sujeto no puede operar una fórmula con un término.

Figuras del eje semántico. Énfasis en los objetos.

- **Referenciales:** Son los contenidos corpóreos sobre los que los términos hacen referencia. La importancia de los referenciales radica en la necesidad de una realidad física sobre la que operar, pues operar sobre términos que no referencian a nada es un sinsentido. En la teología católica, por ejemplo, el término Dios no hace referencia a ningún objeto

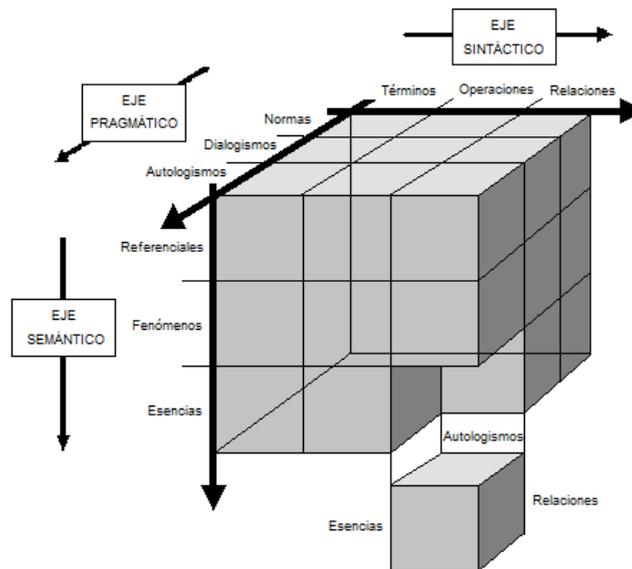
físico, por lo que esta no puede considerarse una ciencia, sino una metafísica. Un ejemplo de un referencial es el agua, que se formaliza en la química como H<sub>2</sub>O.

- Fenómenos: Son las situaciones en las que los objetos del campo se sitúan ante los sujetos. Los fenómenos son recurrentes, lo cual permite que puedan ser formalizados mediante estructuras. Los fenómenos son hechos que tienen que ser explicados por los sujetos que los observan creando nuevos términos que se incorporan al campo categorial. Este protagonismo del sujeto en los fenómenos, al igual que pasaba en las operaciones, conlleva un aspecto relativo del fenómeno, pues este no será interpretado siempre de la misma manera. Un fenómeno puede ser la Luna, pero esta no es igualmente interpretada en el siglo IV a.C. que, en la actualidad, donde contamos con instrumentos de observación y muestras de su composición. El fenómeno interpretado por el sujeto define la morfología y la textura de los términos que construyen.
- Estructuras: Son el resultado de la neutralización del sujeto en los fenómenos. Las estructuras refieren a las relaciones necesarias, objetivas y universales entre términos del campo categorial, las verdades científicas. Continuando con el ejemplo anterior de fenómenos, solo cuando hemos llegado a un conocimiento innegable de nuestro satélite podemos decir que la astronomía incorpora la estructura lunar a su campo categorial.

Figuras del eje pragmático. Énfasis en los sujetos.

- Dialogismos: Son las relaciones que se establecen entre los sujetos que intervienen en la construcción y el desarrollo científico. Se da cuando distintos sujetos operan sobre los mismos objetos en forma de debates, explicaciones, comunidades científicas, enseñanza universitaria... Las ciencias no pueden existir sin dialogismos porque un único sujeto no puede abarcar la complejidad completa de una ciencia. El proyecto logicista impulsado por Frege a finales del siglo XIX fue un dialogismo que se dio entre los matemáticos del momento.
- Autologismos: Son las relaciones que establecen los sujetos consigo mismos en el proceso de investigación científica. El científico, al estar operando sobre términos, lleva a cabo operaciones lógicas que no involucran a otros sujetos. Charles Darwin, en el desarrollo de su teoría de la evolución, presenta autologismos al estar empleando situaciones empíricas vividas por él para deducir nuevos términos que incorporar al campo de la biología.

- Normas: Son las series de pautas que los sujetos deben seguir a la hora de operar sobre los términos de un campo científico. En algunas ocasiones, las normas son impuestas por la definición de los mismos términos y en otras, las normas las imponen los sujetos sobre los objetos para conseguir unos resultados específicos. El método científico es un ejemplo de norma que los sujetos operatorios han seguido a la hora de operar con términos de los distintos campos científicos.



Representación del espacio gnoseológico desarrollado por Gustavo Bueno en su Teoría del Cierre Categorical. Fuente: <https://izquierdahispanica.org/2014/defensa-de-la-tesis-doctoral-de-santiago-armesilla/>

Estas son las nueve figuras formales que constituyen los campos categoriales. Sin embargo, saber qué compone un campo categorial no nos lleva a su total comprensión, pues es necesario conocer cómo se compone el campo como conjunto de figuras, es decir, en qué punto comienza un campo, en qué punto se logra el cierre de este y de qué maneras va desarrollándose. Estos aspectos los estudia Gustavo Bueno en lo que denomina la perspectiva sintética de las ciencias, que, a través de los principios y los modos, ordena de una manera sistemática el conjunto de materialidades que conforman los campos científicos.

#### Perspectiva sintética:

La importancia de la perspectiva sintética de la ciencia es tanta o más que la de sus figuras y ejes gnoseológicos, pues solo a través de los principios y modos podemos construir un sistema capaz de explicar la realidad desde criterios científicos. Estos dos conceptos, como se verá, marcan el desarrollo desde los contextos determinantes hasta los contextos determinados. De alguna manera, pueden entenderse como la forma en que una ciencia es capaz de convertir las realidades físicas en sistemas formales.

- Principios: Los principios de las ciencias son los criterios particulares que ordenan la variedad de normas esenciales que constituyen los términos y las relaciones del campo científico. Estos criterios son específicos de cada ciencia y son fundamentales para construir los contextos determinados a través de las operaciones realizadas por los sujetos gnoseológicos. Los principios no son dados a priori, solo pueden ser determinados una vez se han realizado ciertas operaciones con los términos. Al ser los principios la hoja de ruta por la cual una ciencia pasa de los cuerpos del contexto determinante a las formas del contexto determinado, es necesario que estos principios estén ligados al eje semántico, que es donde los sujetos pueden operar. Sin embargo, los principios no afectan solo a este eje, sino que por las relaciones que se producen también podemos hablar de principios sintácticos y principios pragmáticos.
  - Principios sintácticos:
    - Principios de los términos: Son los propios términos organizados en clases. El triángulo serían un principio de los términos de la geometría que, proveniente del eje semántico (contexto determinante), se incorpora al contexto determinado de la geometría a partir del cual se desarrollan nuevos teoremas.
    - Principios de las relaciones: Hacen referencia a los axiomas, que son aquellas relaciones entre términos que no necesitan de demostración, pero a las que solo se llega por su infabilidad.
    - Principios de las operaciones: Ponen límites a las operaciones de la categoría científica y, por tanto, cierran el campo categorial en un punto. Un principio de las operaciones matemáticas sería que la suma de dos números naturales siempre da otro número natural.
  - Principios pragmáticos: Los principios pragmáticos son más difíciles de definir, pues están determinados no solo por la ciencia sino por los sujetos que la desarrollan. Encontramos principios autológicos como el *cogito, ergo sum* de Descartes, principios dialógicos como la sustituibilidad de los sujetos en física y principios normativos como el principio lógico "dos cosas iguales a una tercera son iguales entre sí".
- Modos: Los modos científicos son los métodos con los que desde una ciencia se opera con los términos y relaciones de los campos determinados. Son las unidades sintéticas

mínimas del cierre categorial y dan lugar a lo que conocemos como teoremas o verdades científicas. La TCC considera los modelos, las clasificaciones o las definiciones, entre otros conceptos, como teoremas, que serán de una tipología o de otra en función del modo que emplee el sujeto operatorio para combinar figuras del eje sintáctico.

- Modelos: Establecen relaciones a partir de términos presentes en el campo categorial.
- Clasificaciones: Partiendo de relaciones, crean términos nuevos.
- Definiciones: Da lugar a términos nuevos a partir de términos preexistentes. Han de coordinarse con el resto de definiciones del campo.
- Demostraciones: Crea relaciones nuevas a partir de relaciones previas existentes en el campo.

Desde luego, no hay que pensar que todas las ciencias empleen estos modos de igual forma, pues existen ciencias como la Botánica que ponen especial atención a las clasificaciones. Lo que sí debe quedar claro es que la ciencia, desde la perspectiva sintética, es un tejido heterogéneo de modelos, clasificaciones, definiciones y formas.

#### Cierre categorial:

Conociendo ya los componentes que forman los campos categoriales de las distintas ciencias, solo nos quedaría entender la forma en que una categoría consigue cerrarse, distinguiéndose del resto de categorías. El punto de cierre categorial es lo que hace que consideremos a una categoría como autónoma y podamos hablar de una ciencia autónoma.

Para entender el concepto de cierre categorial, partimos de una perspectiva primeramente léxica. El concepto está formado por las palabras cierre y categorial, que definimos ahora:

- Cierre: Es el concepto, en lógica o matemáticas, según el cual, dada una multiplicidad de términos y una serie de operaciones definidas, el resultado que obtenemos de las posibles operaciones entre los términos dan términos pertenecientes al mismo campo. Por ejemplo, dado el conjunto de los números naturales y las operaciones de suma y resta, todas las operaciones posibles siempre nos darán un resultado dentro del conjunto de los números naturales.

- **Categorial:** Responde al concepto de categoría que hemos explicado previamente. Lo categorial subraya el carácter del universo dividido en dominios irreductibles los unos a los otros.

Por tanto, la idea de Gustavo Bueno de cerrar las categorías en que se divide el mundo implica la necesidad de que las ciencias, que como hemos visto son aquellas instituciones que categorizan las distintas materialidades que conforman el mundo, lleguen a un punto en el cual las operaciones (eje sintáctico) realizadas por los sujetos (eje semántico) den lugar a términos (eje sintáctico) que pertenezcan al mismo campo categorial. De esta manera, entendemos que las categorías en que se divide el mundo no son previas al cierre categorial, sino que son construidas por las ciencias a través de estas operaciones, y que solo a través de las ciencias podemos distinguir el mundo.

### ¿Es la Ciencia de Datos una ciencia?

En el presente apartado, trataremos de abordar la pregunta sobre si la CD constituye una ciencia capaz de establecer un cierre categorial sobre un conjunto de materialidades, identificando, en la medida de lo posible, los distintos componentes que hemos expuesto en el anterior apartado.

Recapitulando lo expuesto sobre la TCC, decimos que una ciencia tiene un dominio específico de objetos y relaciones, y que se basa en una serie de principios y metodologías o modos que la diferencian de otras disciplinas. A continuación, expondremos los siguientes puntos:

- **Dominio de objetos y relaciones:** Corresponde a lo que Gustavo Bueno denomina el espacio gnoseológico, formado por un eje sintáctico, un eje semántico y un eje pragmático.
- **Principios generales:** Estos principios se desarrollan a medida que los sujetos realizan operaciones sobre los términos, dando lugar a conceptos que definen la ciencia.
- **Metodologías:** Son las formas en que la ciencia avanza a través de la construcción de nuevos términos y verdades científicas.

#### Perspectiva analítica

Comenzando con la definición de un posible espacio gnoseológico para la CD, la mejor forma de proceder será distinguiendo desde los más simple hasta lo más complejo, y así empezaremos

buscando una primera distinción entre las partes materiales y las partes formales que conformarían el campo de la CD.

En cuanto a la parte material, encontramos que la CD parte de una serie de materialidades preexistentes, que han sido creadas por ciencias como la computación o la estadística. El teorema de Bayes, por ejemplo, es una verdad científica propia de la estadística que la CD aplica, a través de algoritmos de aprendizaje automático que son operaciones propias del campo categorial de la computación, sobre un conjunto de datos para realizar predicciones. Ambos conceptos son materialidades preexistentes que la CD reorganiza para emplearlos con un nuevo propósito, que será distinto en cada caso que sea empleado.

La parte formal, por su parte, correspondería a la formalización de una serie de referenciales, como el Teorema Central del Límite o el algoritmo de K-Means en unos símbolos específicos del campo de la CD. Pero, ¿esto se da? No es tan seguro, pues al llevar a cabo esta formalización se crean términos sobre los que posteriormente trabajar, pero la CD no pretende crear términos, sino trabajar a través de los creados por otras ciencias. Por mucho que hoy puedan identificarse los algoritmos de predicción más con la CD que con la computación, esto son construcciones informáticas, no de la CD, que simplemente los usa con objetivos específicos.

De hecho, si el objetivo de las partes formales es, como hemos indicado más arriba, mantener la unidad y recursividad de la ciencia, vemos que eso no ocurre en la CD, pues esta no muestra ninguna unidad o recursividad. Los científicos de datos no actúan de una manera uniforme ni cuando el conjunto de datos sobre el que trabajan es el mismo. No existen pautas ni medios específicos que aplicar sobre un determinado conjunto de datos porque estos los define el científico de datos en función de las necesidades y objetivos del tratamiento que quieran hacer.

Sí que podemos encontrar operaciones, que recordamos pertenecen al eje sintáctico del espacio gnoseológico, pero estas no se dan entre términos, sino entre referenciales, por lo que son operaciones de otro tipo. El conjunto de operaciones que, programadas por el científico de datos, se desarrollan sobre un conjunto de datos a través de un algoritmo, en ningún momento opera sobre términos formales, sino que directamente actúa sobre el objeto en cuestión. El procedimiento es más parecido al que realiza el zapatero al unir la suela y el zapato que a lo que hace el matemático al sumar dos más tres, porque el matemático sí está operando sobre términos para obtener otro nuevo, mientras que el científico de datos opera sobre un objeto con un determinado fin ya planeado, sea este una predicción o una recomendación, al igual que el zapatero, cuyo fin es la elaboración del zapato. Así, operando sobre los objetos materiales, no sobre los formales, tampoco pueden crearse relaciones que generen fórmulas. De hecho, estas

relaciones no serían posibles en la CD por su dinamismo, no pueden generalizarse relaciones porque los objetos sobre los que actúan nunca son los mismos.

Tampoco podemos encontrar, dentro del eje semántico, la figura de la estructura, pues esta solo se da cuando se excluye a los sujetos operatorios, lo cual en la CD resulta imposible. Un buen ejemplo de estructura sería el sistema solar, que, sin la mediación de sujetos operatorios, es decir, una vez han sido neutralizados, puede ser considerado como una verdad científica innegable. En cambio, en la CD, estas estructuras no pueden llegar a darse, porque neutralizar el sujeto operatorio implica no aplicar ninguna transformación sobre los datos, ya que estos no actúan por sí mismos, sino que si no se despliegan una serie de tecnologías que los recopilen, ordenen o transformen estos ni siquiera se generan. Cualquier aplicación de la CD lleva consigo un objetivo o una serie de objetivos desarrollados por una entidad, individuo o asociación, siendo estos los sujetos que operan sobre los distintos objetos. Si los neutralizamos no queda nada.

Por su parte, las figuras del eje pragmático, los autologismos, los dialogismos y las normas son fácilmente definibles en la CD, y es que, como hemos dicho, cada científico de datos emplea autologismos, pues cada conjunto de datos y cada situación pueden ser abordados desde perspectivas diferentes en función de las capacidades y tecnologías que emplee el individuo, pero también se generan dialogismos entre los profesionales del sector en los distintos foros de internet, trabajos académicos, documentación de herramientas... Las normas, en este caso, las encontramos, no aplicadas como un método científico que todos los sujetos que trabajen con datos deben seguir, que eso en realidad no existe en la CD, sino, más bien, como una serie de normas que impiden ciertos tratamientos objetivados en la legislación que más adelante veremos.

Ahora bien, encontrar fácilmente las figuras del eje pragmático no debe inducirnos a pensar que la CD es efectivamente una ciencia. Podemos encontrar estas figuras en muchas otras disciplinas que no son ciencias, como puede ser, de nuevo, la labor del zapatero, que, por supuesto, establece autologismos, sigue y participa de ciertos dialogismos, como son las modas, y trabaja según unas normas como puede ser que el zapato debe ser para el pie.

Para identificar a la CD como una ciencia, en el sentido en que las concibe la TCC, esta debe adaptarse al molde establecido. Por el contrario, estaremos hablando de otra cosa, que será necesario definir.

### Perspectiva sintética:

Desde la perspectiva sintética, la localización de los principios y modos es todavía más ardua tarea que la de localizar las figuras gnoseológicas, pues estos son los que marcan el camino desde los contextos determinantes a los contextos determinados, de alguna manera estableciendo relaciones entre materialidades que cierran el campo categorial constituyendo una ciencia.

Por la parte de los principios, vemos que la CD sí respeta una serie de principios propios de las ciencias en las que se apoya, como la estadística, de donde toma conceptos como población y muestra, o la computación, cuyo principio de "eficiencia computacional" es fundamental en el uso de cualquier algoritmo. Sin embargo, como hemos visto anteriormente, una ciencia debe ser capaz de establecer una serie de principios propios, como resultado de la sucesión de operaciones entre términos, las cuales no se dan realmente en la CD. Los principios no pueden ser dados a priori, pues si la base de una ciencia es la misma que la de otras, ¿cómo íbamos a distinguirlas? Las ciencias, si lo son efectivamente, establecen una serie de principios que son la base de todas las relaciones que se dan entre las figuras gnoseológicas que las conforman.

En cuanto a los modos, resulta muy complicado encontrar en la CD el uso de los conceptos de modelo, clasificación, definición y demostración. La razón de esto es que el científico de datos, cuando trabaja sobre un conjunto de datos, este siempre es único y debe ser tratado como tal. Aunque dos estudios traten sobre las mismas variables, estas pueden ser completamente diferentes y la definición, clasificación o modelos que emplee uno muy probablemente no sea de ninguna utilidad para el otro.

Sí es cierto que existen distintos modelos que son muy utilizados por la CD, como pueden ser los modelos de regresión o los de clasificación, pero cuando hablamos de modos de las ciencias nos referimos a las formas en que los sujetos hacen avanzar la ciencia, construyendo nuevos teoremas y verdades científicas, pasando del contexto determinante al contexto determinado. Desde la CD no se crea un modelo, sino que se emplean con unos determinados fines una serie de modelos creados desde las matemáticas o la computación.

### Cierre categorial

Llegados a este punto, habiendo visto la dificultad para establecer algunas de las figuras gnoseológicas fundamentales, no siendo capaces de encontrar unos principios generados a partir de las operaciones y resolviendo en que no existen unos modos propios que hagan avanzar la CD, ¿cómo podremos establecer un cierre categorial para la CD?

Recordamos que el cierre categorial se da cuando las operaciones generan términos nuevos que pertenecen al mismo campo categorial, pero en una disciplina en que no se generan términos nuevos a partir de las operaciones, el cierre categorial es imposible. No se puede cerrar un campo en el que los únicos componentes que se incorporan provienen de otras ciencias, sería necesario que la CD, a través de unos modos que no hemos logrado definir, fuera capaz de crear nuevos conceptos que se incorporaran directamente al campo categorial de la CD, pero sin modos con los que actuar y sin principios sobre los que apoyarse, estas operaciones son imposibles.

Además, el hecho de la transdisciplinariedad de la CD implica la imposibilidad de establecer un contexto determinado, pues continuamente este debe modificarse en función de los objetivos del estudio alterando todos los conceptos asimilados previamente. No funciona de la misma manera un algoritmo de k-means sobre un conjunto de datos biométricos que sobre los datos recogidos por una empresa en su página web.

### ¿Qué es la Ciencia de Datos?

Hemos determinado, entonces, que la CD no es una ciencia en el sentido en que Gustavo Bueno considera las ciencias positivas en su TCC. Su incapacidad para construir nuevos conceptos y para cerrar un campo categorial autónomo impide que la consideremos como tal, por lo que, a continuación, nuestro esfuerzo debe centrarse en determinar qué es lo que la CD es.

Para ello, nos retrotraemos al momento en que la CD surge como una disciplina propia, recordamos que primeramente derivada de la estadística, en 1962, con la obra de John W. Tukey *The Future of Data Analysis*. Tukey se da cuenta de que la estadística y la CD, desde el desarrollo de la computación deben separarse, porque tienen campos de estudio diferentes. La estadística, al contrario que la CD, sí trabaja con términos, y es por eso que, como indican Zahavi en 1999 y Breiman en 2001, esta no puede trabajar con conjuntos de datos excesivamente masivos como lo hace la CD, porque ese dinamismo no es formalizable desde una ciencia. La CD, al no tener pretensión de formalizar nada, es capaz de trabajar con cualquier conjunto de datos haciendo uso de cuantas herramientas sean necesarias, provengan estas de la estadística, la computación, la economía o la botánica, si fuera necesario.

Entonces, si concluimos en que la CD no tiene esa pretensión de formalizar y su actividad se centra en el uso de herramientas y tecnologías cuyo origen y fundamento provienen de otras ciencias, podemos determinar que, en realidad, la CD es una ciencia en su sentido de saber-

hacer. Al hablar de la idea de ciencia en general, al principio del epígrafe 2.3.1, definíamos una primera idea de ciencia como saber-hacer de un profesional sobre una serie de objetos sobre los que trabaja, como es la ciencia del zapatero sobre los zapatos o la del navegante sobre la navegación. El navegante se centra en la navegación, pero necesita de unas herramientas y conocimientos que no se realizan desde su "ciencia", sino desde la ingeniería o la astronomía. Nosotros, en este punto, equiparamos al científico de datos a este tipo de profesionales, en tanto que es capaz de trabajar sobre una serie de objetos, que son los datos, a través de una serie de herramientas y tecnologías procedentes de otras ciencias. Es por eso que surge esa separación entre la estadística y la CD. Surge porque la estadística no se puede dedicar al cumplimiento de objetivos de instituciones o entidades privadas, pero sus conceptos y conocimientos propios sí pueden ser adaptados a esas necesidades mediante las herramientas que la computación estaba desarrollando en esos momentos.

El científico de datos, por lo tanto, es el profesional que tiene los conocimientos y las capacidades de trabajar sobre conjuntos de datos mediante conceptos y herramientas propias de las distintas ciencias que sean necesarias. Y, en este punto, recordamos también la idea de la CD como cuarto paradigma científico, propuesto por Jim Gray, que propone que la CD es la nueva forma de hacer ciencia, porque, con la capacidad de recoger información, de analizarla y aprovecharla que nos brinda la CD, se puede llegar a conclusiones científicas de una manera más rápida, eficiente y robusta. La CD, por tanto y según la teoría de la ciencia expuesta por Gustavo Bueno, es también un modo de hacer ciencia, pues ofrece un nuevo enfoque sobre el que construir nuevos conceptos.

Entonces, si entendemos la CD como conjunto de conceptos y tecnologías adoptadas desde otras ciencias para trabajar sobre los datos, debemos tener una idea clara de lo que es la tecnología, que, siguiendo la lógica empleada hasta el momento, tomaremos del Materialismo Filosófico, sistema de pensamiento desarrollado por el filósofo Gustavo Bueno. Así, para Bueno, las técnicas son producto de la producción humana en contextos sociales, culturales e históricos precisos, que producen una progresiva transformación del entorno de los seres humanos (Bueno, 2000). De la misma manera son las tecnologías, pero estas se diferencian de las técnicas en que suponen alguna ciencia en marcha que las haya desarrollado. La idea que aquí nos interesa es esa capacidad de la tecnología para transformar el entorno, que es algo que hemos visto muy señaladamente con la computación recientemente.

En conclusión, entendemos la CD como ciencia en su sentido de saber-hacer, en este caso de saber-hacer con los datos, a través de un conjunto de conocimientos científicos de la estadística,

la computación, la ciencia de la información y cualesquiera campos de los que procedan los datos sobre los que trabaja. Además, se apoya sobre una serie de tecnologías, cuya naturaleza es fundamentalmente informática, para llevar a cabo su actividad, que, debido a la magnitud que ha alcanzado en los últimos años, tiene profundas implicaciones políticas, económicas y sociales que a continuación estudiaremos.

### 3. Impacto de la Ciencia de Datos en la actualidad

En el anterior apartado, tras estudiar detenidamente el concepto de ciencia y los requisitos necesarios que dan forma a una disciplina como tal, hemos conseguido determinar que la CD, a pesar de lo que el nombre parece indicar, no es en realidad una ciencia positiva, capaz de construir nuevos conceptos que den como resultado tecnologías propias, sino más bien una disciplina capaz de trabajar sobre grandes conjuntos de datos procedentes de cualquier ámbito, a través de herramientas y conceptos propios de, en primer lugar, la computación, la estadística y la ciencia de la información y, en segundo lugar, cualesquiera disciplinas sobre las que remitan los datos.

Esta nueva definición nos hace poner el foco en el abanico de herramientas tecnológicas, creadas específicamente para trabajar sobre los datos, que se emplean desde esta disciplina. Como hemos visto, las tecnologías son creadas desde las ciencias en contextos sociales, culturales e históricos precisos, y su aplicación produce una transformación progresiva del entorno. Este contexto en que las tecnologías empleadas por la CD son desarrolladas lo hemos visto a lo largo del apartado 2.2. Genealogía de la CD, donde hemos expuesto todas las contribuciones que, construidas desde otras ciencias, han pasado a formar parte del corpus de lo que conocemos como CD. A continuación, nos centraremos en esa transformación de la que habla Gustavo Bueno como consecuencia directa del uso de la tecnología, que, como veremos, al tratarse de unas herramientas cuyo uso se extiende a todos los ámbitos, ha producido profundas transformaciones en el contexto en que nos desenvolvemos.

Para desarrollar un análisis completo y exhaustivo de las implicaciones de la CD nos hemos centrado en tres aspectos que, sostenemos, definen las sociedades de nuestro tiempo: economía, política y sociedad. En cada uno de los siguientes puntos, se llevará a cabo una revisión de la forma en que la disciplina en cuestión ha transformado las relaciones entre los agentes económicos, entre las esferas políticas y los ciudadanos, y de los ciudadanos entre sí.

El orden que hemos seguido de ninguna manera es arbitrario. Realizamos el estudio comenzando desde el aspecto económico, porque vivimos en una sociedad en la que todas las decisiones de los otros dos aspectos deben estar y están subordinadas al aspecto económico. En este punto, coincidimos con Marx en que las relaciones de producción y las fuerzas productivas determinan

la estructura social, política y cultural de la sociedad (Marx, 1867). La política está sometida a una subordinación absoluta al poder económico y, de hecho, su principal labor es conducir y paliar los posibles desajustes que se den en el mercado en virtud de unos u otros grupos ideológicos. Por su parte, las relaciones sociales deben resignarse a una tercera posición por el simple hecho de que estas siempre han estado condicionadas por el contexto político y económico en que se encontraban, que, de alguna manera, establece un marco de libertad sobre el que las personas pueden desarrollar su vida social.

## 1. Economía

La CD ya está presente en todos los sectores económicos. Las posibilidades que la minería, análisis y modelado de datos brinda a las empresas es ya un hecho ineludible para cualquier agente económico que quiera mantener su competitividad en un mercado cada vez más exigente. La capacidad de tomar decisiones basadas en datos, de predecir escenarios, analizar tendencias, optimizar procesos productivos, segmentar clientes para ofrecer experiencias personalizadas, son todas ellas ventajas competitivas tan determinantes que ignoradas pueden suponer para cualquier empresa la pérdida completa de una cartera de clientes en favor de aquellas entidades mejor preparadas. Las distintas herramientas, que englobamos dentro de la disciplina de la CD, han adquirido en los últimos años cada vez más atención por parte de empresas de todos los sectores, promoviendo a su vez, el perfeccionamiento de estas y la total accesibilidad para cualquier tipo de usuario. Tanto es así que la CD ha llegado a empresas con niveles de actuación muy dispares, de tal manera que podemos encontrar estas herramientas aplicadas tanto en multinacionales de la talla de Facebook o Google, como en empresas más pequeñas o familiares como bodegas de vino o restaurantes.

Ya en 1945, en su ensayo *El uso del conocimiento en la sociedad*, Friedrich Hayek abordó la importancia en la economía del conocimiento y la información. Según Hayek, el conocimiento está disperso entre los distintos individuos, cada uno de los cuales poseedor de información específica y localizada sobre sus propias circunstancias y preferencias. Este conocimiento disperso es esencial para la coordinación y eficiencia de una economía, pero, en el momento en que Hayek escribía, era inaccesible para cualquier individuo o entidad, por lo que eran los precios en una economía de mercado las señales que las empresas debían interpretar para conocer las condiciones de oferta y demanda de los agentes económicos. Con la llegada de la CD, los precios dejan de ser el principal transmisor de conocimiento y las empresas pueden extraer la

información directamente de sus clientes, de sus procesos internos, de sus proveedores... Este cambio de paradigma no solo supone el surgimiento de nuevas fuentes de conocimiento para aquellas empresas que se guiaban por los precios de mercado para reducir costes o ampliar producción, sino que también empiezan a surgir empresas de servicios capaces de, a través del procesamiento de la información, perfeccionar las relaciones con los clientes hasta el punto de realizar recomendaciones personalizadas, publicidad dirigida...

También Herbert Simon, en su obra *Administrative Behaviour* de 1947, supo ver esta incapacidad de los individuos y las organizaciones para procesar la información y tomar decisiones debido a las restricciones cognitivas y a la disponibilidad limitada de información. Esta incapacidad la denominó "racionalidad limitada", término que relaciona directamente con la idea de "satisfacción", también propuesto en la misma obra, y que remite a la necesidad de las entidades de buscar siempre soluciones "suficientemente buenas", aceptando que la solución óptima solo podría llevarse a cabo si tuviéramos toda la información que rodea a la decisión, lo cual es imposible en virtud de nuestra racionalidad limitada. Más adelante, en su *The Sciences of the Artificial* de 1969, Simon explora las posibilidades de la inteligencia artificial en la búsqueda de la solución que ofrezca la mayor satisfacción, paliando de esta manera las limitaciones cognitivas y de información inherentes al ser humano. Hoy, más de cuarenta años después, vemos que las ideas de Simon ya se aplican a todos los niveles y a todas las industrias, que, de alguna manera, han aceptado su racionalidad limitada en favor de herramientas de inteligencia artificial y procesamiento de datos que asesoran y, en ocasiones, directamente ejecutan la toma de decisiones empresariales.

De esta manera, la información se ha convertido en el activo más importante en la mayoría de empresas, no solo por el hecho de combatir esa racionalidad limitada de la que Simon hablaba, sino porque la absoluta digitalización de todos los aspectos de la vida de los individuos genera una cantidad de información tan masiva que obliga a las empresas a hacer uso de ella. Para ilustrar esta situación recurrimos a la infografía realizada en 2022 por la consultora DOMO (anexo 1), donde se ilustra la cantidad de información generada en cada minuto. Algunos de los datos interesantes que la infografía ilustra son las seis millones de búsquedas en Google, las casi setenta mil imágenes compartidas en Instagram, el millón de horas de consumo de contenido en streaming o los más de doscientos millones de correos electrónicos enviados. Todas estas cantidades en el tramo de un minuto. La cantidad de información generada es abrumadora, tanto más si tenemos en cuenta los metadatos generados en cada interacción, que, de hecho, son el principal atractivo para las empresas que han adoptado el dato como forma de trabajo.

La infografía también nos ofrece información sobre el porcentaje de población que, a fecha de abril de 2022, ya es usuaria de internet, situado este porcentaje en el 63%, lo que supone casi cinco mil millones de personas. Hay que tener en cuenta que cada acción, desde visualizar un video hasta realizar una transacción, genera una serie de información que las empresas pueden utilizar en su beneficio. Son cifras imposibles de ignorar para ninguna empresa y es lo que estamos viendo. Todas las empresas, todos los sectores, quieren su cacho del pastel e invierten en CD, externalizan estas funciones o tratan de acercarse a este mundo desde la modestia de las pequeñas empresas, pero dar la espalda a un activo que puede generar tanto valor para las empresas es un acto suicida en una economía cada día más competitiva, donde las tendencias se suceden a gran velocidad. La única manera de mantenerse competitivo es medir la información procedente de todos los frentes a los que se tenga acceso para poder hacer un uso inteligente de los medios.

A continuación, veremos algunos aspectos económicos que se han visto afectados por la irrupción de la CD en las relaciones entre los agentes económicos y trataremos de vislumbrar las consecuencias que estas tecnologías han provocado o pueden llegar a provocar, que es el objetivo final del presente estudio.

## Teoría del valor

A lo largo de la historia se han sucedido distintas teorías del valor que han intentado explicar el valor de cambio de los bienes y servicios que se ofertan en los mercados. Las primeras teorías económicas que trataron de explicar este valor de cambio fueron las desarrolladas entre los siglos XVIII y XIX por Adam Smith, en su obra *La riqueza de las naciones* (1776), y por David Ricardo en su obra *Principios de la economía política y tributación* (1817), en las que proponían la teoría del valor-trabajo, basada en la idea de que el valor de un bien se basa principalmente en la cantidad de trabajo necesario para producirlo. Karl Marx, en su obra *El Capital* (1867), se basa en las ideas de Smith y Ricardo para elaborar una teoría del valor propia, que a la cantidad de trabajo socialmente necesario para producir un bien añade la plusvalía, el excedente de valor creado por el trabajo de los obreros sobre el salario que reciben, como los factores determinantes del valor de cambio de un producto.

Posteriormente, en la segunda mitad del siglo XIX, economistas pertenecientes a la escuela neoclásica, como William Stanley Jevons o Léon Walras, proponen una teoría del valor marginal en oposición a la teoría clásica del valor-trabajo. Esta teoría niega que el valor de un bien esté

determinado por la cantidad de trabajo necesario para producirlo, sino que, más bien, es la utilidad marginal la que lo determina, es decir, el beneficio adicional que un consumidor obtiene al consumir una unidad adicional de ese bien o servicio. Esta teoría neoclásica se basa en el principio de la utilidad marginal decreciente, que sostiene que a medida que un individuo consume más unidades de un bien o servicio, la utilidad que obtiene de cada unidad adicional tiende a disminuir.

La idea de utilidad es incorporada por la teoría del valor subjetivo, desarrollada desde la escuela austriaca primeramente por el economista Carl Menger en su obra *Principios de economía* (1871), pero no se concibe de la misma manera, porque mientras que los neoclásicos defienden el principio de utilidad marginal decreciente, la escuela austriaca no cree que la utilidad que los individuos esperan de un bien esté determinada por el número de unidades que adquieran, sino por las necesidades y preferencias de cada individuo en relación a ese bien. Para Menger, la utilidad marginal de un bien es determinada por cada individuo y, por tanto, los precios se asignan en función del consenso de todos los individuos sobre los que se establece un equilibrio entre la oferta y la demanda. La teoría del valor subjetivo también niega la teoría del valor-trabajo al rechazar la idea de que los bienes o servicios tengan un valor intrínseco u objetivo determinado por la cantidad de trabajo necesario para producirlo o la plusvalía generada por el capitalista. Este valor no se determina objetivamente sino subjetivamente por la percepción de cada individuo sobre la utilidad de ese bien.

En nuestro tiempo, la teoría del valor subjetivo ha demostrado ser la más capaz de explicar el valor de los bienes y servicios que consumimos, que, por supuesto, no se definen como consecuencia del trabajo necesario para producirlos ni tienen por qué responder al principio de utilidad marginal decreciente. Un ejemplo que ilustra perfectamente la potencia de esta teoría por encima de las otras es el mercado de NFT (Non Fungible Tokens), que son productos digitales con los que en los últimos años ha surgido una gran polémica por la posibilidad de especular con ellos. Los "CryptoPunks", por ejemplo, son una colección de NFT de imágenes relativamente simples que podrían ser creadas en cuestión de minutos mediante un programa de diseño de lo más rudimentario, pero que han llegado a venderse por millones de dólares, valiendo veintitrés millones de dólares la pieza más cara hasta la fecha. El valor de estas imágenes no está determinado por la cantidad de trabajo empleado como



CryptoPunk #5822, vendido por 8.000 ETH (23 millones de dólares).

Fuente:

<https://www.revistagq.com/noticias/articulo/venden-nft-23-millones-dolares-punk>

sostendría Marx y compañía, ni tampoco por la utilidad marginal objetiva de la obra, la cual es nula. La única manera de explicar este valor es la percepción subjetiva de un conjunto de individuos que generan un fenómeno mediático alrededor de estos productos, potenciado, cómo no, por las redes sociales. Por supuesto, existen más ejemplos, no solo de productos digitales, tales como los productos de lujo, los objetos de coleccionismo o los productos y servicios personalizados, todos ellos muy presentes en la economía actual.

Ahora bien, ¿de qué manera ha afectado la CD en la teoría del valor? Pues lo cierto es que la CD no modifica para nada la teoría del valor propuesta por la escuela austriaca, pero, de alguna manera, a través de esta teoría podemos comprender la potencia de la CD en la actualidad.

Tenemos, por tanto, una economía casi completamente digitalizada (hemos visto que el 63% de la población ya actúa a través de internet), donde los individuos que la conforman se comunican a través de redes sociales masivas, se entretienen en plataformas de streaming o vídeos, trabajan y consumen a través, o con ayuda, de internet, y en el proceso de todas estas actividades generan una cantidad masiva de datos, que es lo que conocemos como Big Data. Toda esta información generada, una vez procesada por algoritmos de inteligencia artificial en tiempo real o por herramientas de análisis estadístico, ofrecen conocimientos a las empresas sobre todo tipo de variables: sentimientos hacia determinados productos, preferencias de contenido, métodos de actuación ante determinados procesos... Todo esto sumado a información de carácter demográfico: lugar de residencia, edad, sexo... Es aquí, ante este conjunto de información, donde entran las empresas y la CD, y donde las palabras de Ludwig Von Mises, en su obra *Human Action: A Treatise on Economics* (1963), cobran la mayor importancia: "*Economics is not about things and tangible material objects; it is about men, their meanings and actions*". Mises, también perteneciente a la escuela de economía austriaca, sigue la estela marcada por Menger y Hayek, y con sus esclarecedoras palabras parece hacer una radiografía a la economía del siglo XXI, en la que el producto o el servicio ha perdido toda la importancia en el mercado, en el que lo único que importa es lo que la gente percibe y expresa a través de sus actos. Las empresas lo saben y por ello han centrado la atención en la CD como el medio de conocer esas percepciones subjetivas y alterarlas en su beneficio.

## Economía del comportamiento

La economía del comportamiento es la rama de la economía que, apoyándose en la investigación de las tendencias cognitivas y emocionales de los seres humanos en sociedad, estudia cómo los agentes económicos actúan en los mercados y cómo sus actuaciones son determinantes en los precios de los bienes y servicios y en la toma de decisiones de las empresas. El pensamiento económico conductual nace de la mano de Daniel Kahneman y Amos Tversky con su obra *Prospect theory: Decision Making Under Risk* (1979), en la que emplean técnicas de psicología cognitiva para explicar ciertos fenómenos económicos. Para los estudiosos de la economía del comportamiento, las personas a menudo toman decisiones económicas basadas en sesgos cognitivos, heurísticas y factores emocionales, que, guardando perfecta coherencia con la teoría del valor subjetivo, les alejan de la concepción de la economía neoclásica, que considera que los individuos son agentes racionales que actúan siempre en busca de maximizar la utilidad.

Esta economía del comportamiento, aunque teorizada ya a finales del siglo XX, era algo que ya estaba en práctica desde los inicios de la publicidad convencional, que siempre se centró en exaltar esos factores emocionales de los compradores, esas preferencias por un estilo de vida en particular, pero, con la llegada del siglo XXI, de la CD y el Big Data, la capacidad de las empresas de localizar exactamente qué es lo que quieren sus compradores, cuáles son sus características específicas, cuáles las técnicas que mejor funcionan con ellos, expande significativamente las posibilidades de la economía conductual, que ya no es una mera teoría sobre cómo funcionan las mentes de los compradores en los mercados, sino una forma de actuación por parte de las empresas.

Richard H. Thaler y Cass R. Sunstein, en su obra *Nudge: Improving Decisions About Health, Wealth, and Happiness* (2008), en el contexto de la economía del comportamiento, definen dos conceptos que nos merecen una especial atención:

- Elección arquitectónica: Se refiere a la organización y presentación de opciones e información en un contexto específico que puede influir en la toma de decisiones de las personas.
- Empujón (Nudge): Es una intervención sutil en el proceso de toma de decisiones que altera el comportamiento de las personas de manera predecible sin prohibir opciones ni cambiar significativamente los incentivos económicos. Los empujones pueden ser parte de la elección arquitectónica, ya que modifican el contexto en el que se toman las decisiones para influir en el comportamiento de una manera más deseable.

Si bien es cierto que Thaler y Sunstein proponen estas ideas como mecanismos ideados para ayudar a las personas a tomar mejores y más informadas decisiones, en respuesta a la irracionalidad innata de las personas, dando forma a lo que denominan "libertarismo paternalista", a través del cual las personas mantendrían su libertad de elección mientras desde las políticas públicas podrían preocuparse y potenciar el bienestar de las personas, estos mecanismos son difíciles de implantar a nivel político y, por tanto, más difíciles aún de localizar actualmente, mientras que en los mercados podemos ver como las empresas los emplean continuamente. Las plataformas de streaming, por ejemplo, como *Netflix* o *HBO*, utilizan algoritmos de recomendación para presentar a sus usuarios una selección personalizada de programas y películas basada en sus hábitos de visualización y preferencias, además de empujones como la reproducción automática de contenidos. También Amazon y AliExpress emplean la CD para personalizar la experiencia de compra de cada uno de sus clientes en función de sus búsquedas, sus compras anteriores...

Estas prácticas han demostrado ser tremendamente efectivas, pero también plantean algunas cuestiones éticas que deberían ser tenidas en cuenta. En el caso de Netflix o Amazon, en principio, parece no haber ningún problema en cuanto a la privacidad de los usuarios porque se supone que desarrollan una elección arquitectónica basada en los datos que recogen desde la plataforma, los cuales no deberían suponer ninguna violación a la privacidad y sí una mejora en la experiencia de usuario y la usabilidad de las funcionalidades. Sin embargo, no podemos reducir el problema a la privacidad, pues existen otras cuestiones como la adicción que determinados empujones pueden generar en los individuos. La reproducción automática en Netflix, por ejemplo, es un empujón que genera adicción al contenido. Este tipo de empujones los vemos continuamente, sobre todo en redes sociales, donde ciertos tipos de contenidos o ciertos aspectos presentes en estos funcionan mejor que otros por razones puramente psicológicas. Véase la cantidad de flechas y círculos rojos presentes en la mayoría de videos virales en la plataforma YouTube. Esto son empujones que los creadores de contenido han percibido que generan más audiencia de lo que el video quizás merecería y, por tanto, comienzan a proliferar en la plataforma. Este aspecto lo veremos en el último punto de esta revisión cuando veamos las redes sociales y sus mecanismos de distribución del contenido.

Aunque parezca banal la cuestión de la adicción en plataformas de streaming, desde luego no lo es tanto en una aplicación de inversión o en casas de subastas online, donde se emplean mecanismos de diseño y gamificación, como animaciones y sonidos atractivos, para hacer que la inversión sea más emocionante y, por tanto, menos racional. Además, una elección arquitectónica excesivamente simple e intuitiva en este tipo de plataformas puede suponer un

peligro para usuarios inexpertos. A todo esto, hay que sumar la publicidad en redes sociales de este tipo de productos, que se manifiesta directa, a través de anuncios pagados y dirigidos especialmente a los usuarios más vulnerables, e indirectamente, a través del debate generado alrededor de fenómenos de esta índole.

### El dato como mercancía, el usuario como productor

Hemos visto cómo las empresas emplean la información de sus usuarios y compradores para diseñar y desplegar herramientas y soluciones de CD con el fin de aprovechar sus debilidades cognitivas y emocionales, generando de esta manera beneficios económicos y más información que analizar. Así, se puede entender la empresa de esta economía digital como el agente económico que recibe como input información desordenada y de distintas tipologías, la procesa a través de sus herramientas de CD, y genera como output nuevas soluciones y plataformas que, de nuevo, recopilan más y más información. En este proceso, en el que ya hemos visto cómo actúan las empresas durante ese procesamiento, nos queda por ver la información como mercancía, como materia prima que es sometida a transformación para generar un valor añadido.

En su obra *Digital Capitalism: Networking the Global Market System* (1999), Dan Schiller examina cómo, en el capitalismo digital, la información se ha convertido en una mercancía central que se produce, se vende y se intercambia en el mercado, cuyo valor económico se deriva de su capacidad para ser utilizada en la producción de bienes y servicios para generar beneficios. En este contexto, el usuario, como generador de ese contenido, se convierte en un elemento clave del sistema: "los datos de los usuarios se han convertido en un recurso crucial para las corporaciones, que los utilizan para refinar sus estrategias de marketing, mejorar sus productos y servicios, y aumentar su participación en el mercado" (p. 135).

En la misma línea encontramos a Shoshana Zuboff con su obra *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (2019), en la que introduce el concepto de "capitalismo de vigilancia", como el modelo económico basado en la idea de que los usuarios son el producto y que sus datos personales son una nueva forma de capital. Esta forma de capitalismo se nutre de la recopilación masiva de datos y la inteligencia artificial para analizar, predecir y, en última instancia, influir en el comportamiento humano.

En este nuevo capitalismo digital o capitalismo de vigilancia, el usuario se convierte en el productor de la materia prima más importante, la información, y, por tanto, es el principal

generador de capital de las grandes empresas. Sin embargo, el usuario no recibe ninguna compensación económica y en ocasiones es él el que paga por usar los servicios que son mejorados gracias a su actividad. De alguna manera, es como si los usuarios, generadores constantes de contenido, por tanto, de capital y valor para las empresas, fueran empleados de estas a coste cero, que, bajo la apariencia de ofrecer un servicio fundamental de comunicación, en el caso de las redes sociales, o de entretenimiento, en el caso de las plataformas de contenido bajo demanda, dan por pagada la labor del usuario, que, por supuesto, para las empresas es mucho más valiosa.

La naturaleza de la información generada por los usuarios es uno de los puntos clave a la hora de analizar las cuestiones éticas relativas al uso de los datos por parte de las empresas. El conocimiento de que las empresas emplean nuestra información como parte principal de sus negocios es ya casi general entre la población, a la que, sin embargo, no parece importarles lo más mínimo, pues el uso de los productos que se van desarrollando es cada vez mayor. El asunto aquí es poner en duda el carácter íntimo y privado de esta información, porque ¿realmente estamos dando información íntima? Debemos analizarlo.

Cuando las personas navegamos por internet, interactuamos en redes sociales o nos desenvolvemos en el mundo real estamos generando información por el simple hecho de estar conectados a una red de internet. Pero, en el contexto en que estamos actuando, es un contexto público, no privado, en el que las herramientas de medición se han hecho más sofisticadas. Antes de internet y de las TIC, también generábamos cierta información al salir a la calle a comprar el pan, por ejemplo, una información que, aunque no se recogía en forma de datos, se recogía como información factual por el panadero que nos vendía, por el vecino que nos encontrábamos... De ninguna manera diríamos que el panadero está atentando contra nuestra privacidad, pero sí lo pensamos cuando una red social nos muestra un anuncio de un producto del que hemos estado hablando. No hay tanta diferencia, en realidad, porque al desenvolvernos en lugares sociales, e internet lo es de cierta manera, lo que estamos mostrando es nuestra "extimidad", no nuestra intimidad.

El concepto de "extimidad" fue introducido por el psicoanalista Jacques Lacan en su seminario *La ética del psicoanálisis* (1958), y refiere a la externalización de la intimidad y la personalidad de los seres humanos al desenvolverse en entornos sociales. Para Lacan, el núcleo íntimo de una persona no se encuentra en su interior, sino que es parte de su relación con el mundo exterior. La extimidad sugiere que la intimidad y la identidad no pueden ser separadas del entorno social y cultural en el que se encuentran, y en un entorno social digital mediado por elecciones

arquitectónicas específicas que fomentan que los usuarios se sientan cómodos ofreciendo su información, ha llevado a una cultura en la que las personas comparten abierta y públicamente aspectos de sus vidas privadas que antes se consideraban reservados para círculos íntimos. El hecho de que los productos se modifiquen en función de los gustos de los usuarios fomenta que estos compartan qué es lo que les gusta, lo que les interesa, lo que quieren seguir... También, por supuesto, obtienen información no ofrecida directamente por el usuario, como el tiempo de visualización de un contenido, pero que también forma parte de esa intimidad, porque son acciones que se realizan en entornos sociales o bajo condiciones de uso aceptadas. Este proceso de "extimización" de la intimidad crea una tensión entre la necesidad de mantener ciertos aspectos de nuestras vidas privadas y la presión de compartir y exponer nuestras experiencias y emociones en línea.

El problema de la intimidad y la privacidad de los usuarios se fundamenta en la capacidad de las empresas de personalizar las experiencias hasta un punto excesivamente preciso. Por ejemplo, es bastante común que una pareja que vaya a tener un bebé comience a encontrar anuncios de puericultura, lo cual puede resultar inquietante al tratarse de un asunto tan privado, pero, en realidad, esa pareja ha tenido que declarar previamente ese hecho a través de sus acciones en buscadores, redes sociales, tiendas online, etc. Lo cierto es que la forma en que las empresas recopilan y procesan la información de los usuarios, más que llevar a la personalización, siempre lleva a la despersonalización de la experiencia en beneficio del agrupamiento de usuarios en segmentos específicos. Cuando los datos se recopilan, y la información es tan masiva, las herramientas no pueden centrarse en cada uno de los usuarios, sino que detectan patrones de comportamiento que identifican a unos con otros y crean perfiles que asignan en función de las características de estos colectivos, asignándoles una elección arquitectónica específica que se prevé desembocará en ciertos comportamientos deseables. Los anuncios de puericultura no están dirigidos directamente a la pareja en cuestión, sino que se ha detectado que los integrantes de esa pareja pertenecen a un segmento de usuarios especialmente interesados en ese tipo de productos.

### **Distribución de la riqueza. Monopolio de la información y las Big Tech**

Otra cuestión importante que debemos estudiar en relación a las implicaciones de la CD en la economía es el de la distribución de la riqueza. Habiendo visto el proceso completo por el cual los usuarios o clientes generan los datos y las empresas los recopilan, procesan y generan un

valor añadido en miles de formas diferentes, debemos ahora centrarnos en el conjunto de infraestructuras que sostienen todo esto, que será la principal causa de distinción entre unas pocas grandes empresas muy poderosas y el resto.

Dan Schiller, de cuya obra *Digital Capitalism: Networking the Global Market System* (1999) hablamos anteriormente, habla de estas infraestructuras de red como un componente crucial para el funcionamiento del capitalismo digital, ya que han permitido una mayor integración económica y una mayor acumulación de capital a través de las fronteras geográficas. Según Schiller, las empresas que controlan la infraestructura de red tienen el poder de controlar y manipular la información que fluye a través de estas redes. Al hacerlo, pueden moldear los mercados y la opinión pública, lo que les otorga una posición de poder en la economía global.

Es interesante apreciar cómo, aunque en 1999, Schiller, sin poder saber hasta dónde llegaría la CD, ya detectaba el peligro de que el poder de las infraestructuras de red cayera en manos de unas pocas empresas. La posibilidad de controlar una gran cantidad del flujo de datos ya preocupaba a Schiller cuando la población de internet era mucho menor que en la actualidad, pero hoy, sin duda, el peligro es aún mayor. Ya hemos visto la población activa y la cantidad de información que se genera diariamente, lo cual sumado a las técnicas y herramientas de CD, que cada vez son más precisas, ponen a disposición de las organizaciones que dominen las infraestructuras de red unas oportunidades de control y manipulación que nunca antes en la historia de la humanidad se habían dado.

Las infraestructuras de red, para dejar claro el concepto, refieren a todos los componentes físicos, como cables, servidores y centros de datos, y aspectos intangibles, como protocolos de comunicación, software y estándares tecnológicos, que en conjunto crean una infraestructura sobre la que se sustentan las redes de comunicación y las tecnologías de la información en la sociedad en red. Estas infraestructuras facilitan la interconexión de personas, organizaciones y dispositivos, lo que permite a los actores sociales y económicos colaborar, competir e innovar en tiempo real y a escala global (Schiller, 1999).

Ante esta definición es imposible no pensar en las grandes compañías tecnológicas conocidas como "Big Tech": Google, Amazon, Facebook, Apple y Microsoft. Son empresas que han marcado el desarrollo del capitalismo digital del que habla Schiller y que, sin duda, son las que mayor cantidad de información generan, tratan y emplean en su actividad. Estas empresas son de sobra conocidas y cualquier intento de definición será menos exhaustivo que las ideas que toda la población tiene de ellas, tal es su presencia en la sociedad. Algunas características comunes entre las empresas que conforman este exclusivo grupo son las siguientes:

- **Tamaño y alcance global:** Las big tech suelen ser empresas con una gran capitalización de mercado y presencia en todo el mundo. Han expandido sus operaciones y servicios a nivel global, llegando a millones, si no miles de millones, de usuarios.
- **Innovación y tecnología:** Estas empresas son líderes en innovación y desarrollo tecnológico. Invierten grandes sumas de dinero en investigación y desarrollo (I+D) para crear y mejorar productos y servicios, lo que les permite mantenerse a la vanguardia de la industria.
- **Dominio de mercado:** Las big tech tienen una posición dominante en sus respectivos mercados y, en algunos casos, cuentan con cuotas de mercado cercanas a un monopolio. Esta posición les permite tener un control significativo sobre sus industrias, estableciendo estándares y prácticas que otros deben seguir.
- **Recopilación y uso de datos:** Las big tech son conocidas por recopilar y utilizar grandes cantidades de datos de sus usuarios para alimentar sus algoritmos y mejorar sus productos y servicios. Estos datos también les permiten ofrecer publicidad dirigida y personalizada, generando ingresos significativos.
- **Integración vertical y diversificación:** Muchas de estas empresas han expandido sus operaciones a través de la integración vertical, adquiriendo o desarrollando componentes y servicios clave en su cadena de suministro. También han diversificado sus negocios al adentrarse en nuevos mercados y ofrecer una amplia gama de productos y servicios.
- **Influencia política y social:** Dada su importancia económica y presencia global, las big tech tienen un impacto significativo en la política y la sociedad. En algunos casos, han enfrentado críticas y escrutinio por temas como la privacidad de datos, el poder monopolístico, la desinformación y las prácticas laborales.
- **Adquisiciones y fusiones:** Las big tech a menudo adquieren empresas más pequeñas para expandir sus capacidades tecnológicas, acceder a nuevos mercados o eliminar competidores potenciales. Estas adquisiciones y fusiones les permiten mantener su posición dominante y crecimiento en el mercado.
- **Altos niveles de rentabilidad:** Las big tech suelen ser altamente rentables, con márgenes de beneficio superiores a los de otras empresas en la industria. Esto les permite generar un flujo constante de ingresos y reinvertir en sus operaciones y crecimiento.
- **Cultura corporativa y atracción de talento:** Las big tech suelen tener una cultura corporativa centrada en la innovación, la creatividad y la colaboración. Estas empresas atraen y retienen a algunos de los profesionales más talentosos y cualificados del

mundo, lo que les permite mantenerse a la vanguardia en términos de desarrollo tecnológico e innovación.

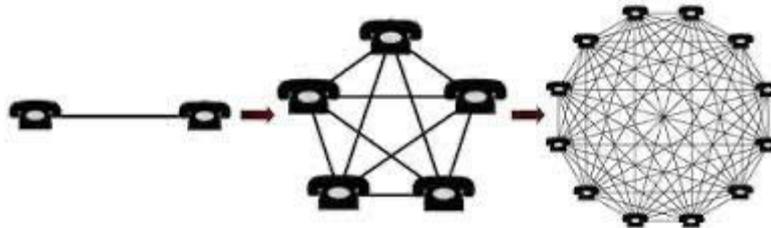
- **Redes de efecto:** Las big tech se benefician de las redes de efecto, lo que significa que su valor y utilidad para los usuarios aumenta a medida que más personas utilizan sus productos y servicios. Este fenómeno les permite crecer rápidamente y mantener a raya a la competencia.
- **Barreras de entrada:** La posición dominante de las big tech en sus respectivos mercados, junto con sus recursos y capacidades tecnológicas, crea barreras de entrada para las empresas emergentes y competidores más pequeños. Esto puede dificultar la innovación y competencia en la industria y generar preocupaciones antimonopolio.

A través de estas características comunes podemos discernir el poder que estas empresas pueden llegar a tener sobre economía, política y sociedad mediante el uso de la información. Son empresas presentes en todos los rincones del mundo, allí donde haya una conexión a internet lo más probable es que alguna de estas empresas esté recogiendo información. Participan en casi todos los mercados digitales en los que haya posibilidad de encontrar rentabilidad, adquieren empresas más pequeñas si demuestran ser útiles para sus fines, acaparan la atención de la mayoría de los perfiles profesionales tecnológicos y se mantienen en un estado de constante desarrollo y mejora sostenido por el constante flujo de información que generan y trabajan a través de las herramientas y los profesionales más sobresalientes de la CD.

Todo parece indicar que estas empresas son la quintaesencia del capitalismo, empresas que a través de la innovación y el desarrollo tecnológico han adquirido una posición dominante en el mercado, cuya actividad consiste en la producción de bienes y servicios que facilitan la vida de la gente. Sin embargo, hay que ser críticos. Estas empresas mantienen un monopolio, o un oligopolio si se quiere, en el que las barreras de entrada son insuperables, no ya solo por incapacidad económica de otras empresas, sino porque cualquier pequeña organización que comience a despuntar es rápidamente adquirida por una de las grandes. También debemos sumar el hecho de que la competitividad tan celebrada por el capitalismo, en el ámbito digital, solo se da entre estas seis empresas, que son las que marcan el camino del desarrollo tecnológico, el cual todas las demás deben seguir por no quedar perdidas en proyectos que no generaran el interés de una sociedad dominada por lo que dictan desde las big tech.

El aspecto que más nos debe interesar en este estudio es el de las big tech como redes de efecto, pues es aquí donde reside la principal fuente de poder de estas empresas. Decimos que una empresa o un servicio es una red de efecto cuando su valor aumenta a medida que más gente lo

utiliza. El Marketplace de Amazon es una red de efecto que es más atractiva y eficiente cuantos más compradores y vendedores interactúan en ella. Con las redes sociales también se ve claramente que cuantos más usuarios, mayor contenido y, por tanto, más entretenimiento. La labor de la CD aquí es fundamental, porque su misión es encargarse de que ese incremento de los usuarios y, en consecuencia, de contenido e información no degenera en un caos incomprensible para los usuarios, consiguiendo que cada uno de ellos pueda desenvolverse cada vez mejor en un océano de contenidos. Como podrá verse este proceso se retroalimenta continuamente: los nuevos usuarios generan más contenido, la información generada se emplea para optimizar la experiencia, lo cual atrae a más usuarios que, de nuevo, generan más contenido. La forma en que estas plataformas mejoran la experiencia de los usuarios ya la hemos revisado en el anterior apartado. Lo que no hemos abordado es la posibilidad de que las plataformas, haciendo uso de esta información, influyan en la opinión pública y en la toma de decisiones a través de la distribución y promoción de contenido específico. Al controlar los algoritmos que determinan qué contenido se muestra a los usuarios, las big tech tienen un poder significativo para influir en las percepciones y creencias de la sociedad, y, como veremos en el siguiente punto, ya se han dado ocasiones en que la CD se ha empleado con estos propósitos.



Ejemplo de red de efecto. Dos teléfonos = 1 conexión, 5 teléfonos = 10 conexiones, 12 teléfonos = 66 conexiones. Fuente: <https://7about.fr/la-loi-de-metcalf/>

La retroalimentación de la que hablamos tampoco acaba ahí, porque todos estos usuarios que emplean las infraestructuras de red de estas empresas les generan una rentabilidad generalmente alta, lo que permite a las empresas diversificar su actividad ocupando sectores que aún no están ocupados por las grandes tecnológicas, pero que una vez estas llegan a la competencia se vuelve imposible. Aquí el problema es el que comentábamos al principio a propósito de la obra de Schiller: estas empresas controlan una cantidad de información tan masiva y tan sumamente valiosa, que tienen la libertad de ejercer cualquier actividad empresarial que se propongan. La libertad, en términos del Materialismo Filosófico de Gustavo Bueno, refiere a la capacidad de hacer contra los demás. La libertad de este tipo de empresas, en ocasiones inmunes a cualquier legislación, es muy diferente a la de una pequeña o mediana empresa a la que le arrebatan parte del mercado al que se había estado dedicando. La libertad,

sustentada sobre el dinero y el poder, de las big tech se ha construido en parte mediante el uso de una información extraída de unos usuarios que en la mayoría de ocasiones no se dan cuenta, o por lo menos no se dan cuenta de las consecuencias que esta puede tener.

Ante estos desafíos y preocupaciones, algunas propuestas sugieren la creación de un mercado descentralizado de datos como una solución para romper el monopolio de la información y redistribuir el poder y la riqueza en la economía digital (Lanier, 2013). Jaron Lanier, un pionero de la realidad virtual y filósofo tecnológico, propone esta idea en su libro *Who Owns the Future?* (2013). Lanier argumenta que, en lugar de permitir que las big tech acumulen y controlen grandes cantidades de datos de usuarios, los datos deberían ser propiedad de los individuos que los generan y ser compensados cuando estos datos son utilizados por otras empresas o plataformas.

En un mercado descentralizado de datos, los usuarios tendrían el control de sus propios datos y podrían decidir con quién compartirlos y en qué términos. Las empresas interesadas en utilizar estos datos para mejorar sus productos, servicios o publicidad tendrían que compensar a los usuarios de manera justa. Este enfoque no solo permitiría una distribución más equitativa de la riqueza generada a través de los datos, sino que también reduciría el poder que las big tech tienen sobre la información y, por ende, sobre la economía y la sociedad en su conjunto.

## Conclusiones

Habiendo visto las principales cuestiones problemáticas del uso de la CD en la economía, podemos proceder a enumerar una serie de puntos a tener en cuenta cuando revisemos la legislación vigente en materia de datos, con el objetivo de determinar si estos aspectos ya están siendo abordados por la ley y, en caso contrario, cómo podrían ser legislados. Es importante decir que, a pesar de que nos estamos centrando en las problemáticas derivadas del uso de la CD, esta disciplina, como cualquier otra basada en tecnologías tan potentes, presenta muchos beneficios potenciales que, sin embargo, no son materia de este trabajo, cuyo objetivo principal es desentrañar esos posibles abusos del poder que estas herramientas ofrecen.

Procedemos, así, a enumerar los aspectos económicos que abordaremos en la última parte del presente estudio:

- Protección de datos y privacidad: Implementación de políticas de protección de datos más estrictas para proteger la privacidad del usuario y limitar la recopilación y el uso de datos por parte de las empresas.
- Transparencia y responsabilidad: Crear leyes que obliguen a las empresas a ser transparentes en sus prácticas y responsables de sus acciones, especialmente en lo que respecta a la manipulación de la información y la influencia sobre la opinión pública.
- Propiedad de los datos: Legislación que defina y proteja la propiedad de los datos, posiblemente siguiendo el modelo de un mercado descentralizado de datos en el que los usuarios tengan el control de sus propios datos.
- Control de la infraestructura de red: Regulación de las empresas que controlan la infraestructura de red para evitar que tengan el poder de manipular la información y moldear los mercados.
- Dominio de mercado y prácticas anticompetitivas: Creación de leyes para controlar las prácticas monopolísticas y fomentar una competencia justa en la economía digital.

## 2. Política

Las posibilidades que la CD presenta en una sociedad completamente digitalizada han quedado más que demostradas en el anterior apartado sobre la economía. Las empresas la emplean para conocer las preferencias de los clientes, para optimizar sus procesos productivos, para realizar predicciones o modelos que apoyen la toma de decisiones... Todas estas funciones son trasladables al ámbito político, pero las consecuencias que se derivan de esto son, si cabe, más delicadas que las anteriores.

Hemos visto los peligros a los que un mal uso de estas tecnologías puede conducir, pero estos, como hemos visto en las conclusiones, son abordables desde una legislación coherente y trabajada que, se supone, debe llegar desde los poderes políticos y legislativos de los Estados o las organizaciones políticas internacionales. En el caso de que los abusos de estas tecnologías procedan directamente de los poderes que deben encargarse de castigarlos, las soluciones se antojan más complicadas.

La política, a grandes rasgos, consiste en la organización y dirección de las sociedades humanas, las cuales están determinadas por las condiciones materiales que las rodean, a saber, condiciones económicas, sociales, culturales y tecnológicas. Aunque será más adelante cuando nos centraremos en analizar la sociedad resultante de un capitalismo basado en la explotación de los datos, para entender de qué manera desde las esferas políticas hacen uso de la CD, debemos considerar a la sociedad en su conjunto como entendía Platón a la clase de los productores en su *República* (S.IV a.C).

Para Platón la República ideal debía estar formada por tres clases sociales con las que establecía una analogía con las partes constituyentes del alma individual:

- La clase de los gobernantes, cuya tarea es liderar al Estado, la equipara con la razón, la encargada de conducir el cuerpo hacia la verdad y el conocimiento.
- La clase de los guerreros, encargados de asegurar que las decisiones de los gobernantes se cumplan, la compara con el espíritu, la parte del alma que busca el honor y es responsable de las emociones y las pasiones.
- La clase de los productores, encargados de trabajar para satisfacer las necesidades físicas del Estado, es comparada con la parte más baja del alma que busca satisfacer los deseos físicos y materiales, el apetito.

Digo que tenemos que comparar a la sociedad actual en su conjunto con la clase productora porque, creo, es la clase que en la realidad más se ha acercado al ideal platónico. La sociedad del capitalismo digital, el más exigente y de mayores dimensiones hasta la fecha, es una sociedad guiada por el apetito, por la satisfacción de unas pasiones efímeras ofrecidas con asombrosa precisión por empresas cada vez más grandes e inteligentes. La sociedad, obligada a trabajar más que nunca, a generar más que nunca, trabaja incluso cuando piensa que está en su tiempo libre, empleada por las grandes empresas como generadora de información. La implicación política es necesariamente nula, pues la labor política y el consecuente uso de la razón necesitan un tiempo y dedicación del que no disponemos, por lo que nos limitamos a consumir entretenimiento, bienes y servicios, incentivados por campañas basadas en la CD. Nos guiamos por el apetito porque no tenemos más opciones y porque este es constantemente estimulado por algoritmos de inteligencia artificial capaces de determinar qué es lo que va a despertarlo.

En cuanto a la labor política, dirigida necesariamente sobre la sociedad existente, el ideal platónico está perdido en el horizonte. La política actual se basa en la generación y explotación de conflictos que enfrentan primeramente a los grupos políticos y como consecuencia a los miembros de la sociedad que ven reflejados sus intereses en uno u otro bando político. Pero esta representación es solo superficial, la labor diaria del pueblo no permite ahondar en los problemas presentes en una sociedad y los políticos, conocedores de esto, lo explotan con discursos grandilocuentes que despiertan las emociones de los ciudadanos y los enfrentan entre ellos. El político de derechas, al igual que el de izquierdas, ya no defiende intereses de una u otra clase social, sino que genera un espectáculo que mantiene al ciudadano frente a la pantalla injuriando o aplaudiendo. Los políticos, esto sí lo supo ver Platón, son una clase social aparte que no tiene ninguna relación con el pueblo más que de sustento, por lo que el político de derechas tiene más dependencia del político de izquierdas que de cualquiera de sus votantes. Es una visión maquiavélica, si se quiere, de la labor política, en la que los miembros de esta clase social han determinado que la mejor forma de mantener el poder es generar conflictos entre los ciudadanos para que estos no puedan organizarse contra ellos.

Bajo esta determinación, la irrupción de la CD en la labor política aparentemente presenta más problemas que beneficios. Si bien, como señala Beth Simone Noveck en *Smart Citizens, Smarter State: The Technologies of Expertise and the Future of Governing* (2015), la CD puede ayudar a los responsables políticos a entender mejor a sus electorados, a predecir el resultado de las políticas y a optimizar la entrega de servicios, resulta muy optimista pensar que las intenciones políticas vayan a ser tan loables, por lo que, desde este estudio coincidimos más con el enfoque proporcionado por Jamie Bartlett en su obra *The People Vs Tech: How the Internet is Killing*

*Democracy (and how we save it)* (2018) o Jamie Susskind en *Future Politics: Living Together in a World Transformed by Tech* (2018), donde exploran cómo la recopilación masiva de datos, junto con el poder de los algoritmos para influir en la opinión pública y el comportamiento electoral, está socavando la autonomía individual y la toma de decisiones informadas, que son esenciales para el funcionamiento de una democracia saludable.

## 1984 y las sociedades de control

En el año 1949, inmediatamente después del final de la Segunda Guerra Mundial y en el comienzo de una Guerra Fría que amenazaría durante más de cuarenta años a las democracias occidentales, George Orwell escribe *1984*, donde presenta una visión aterradora de una sociedad totalitaria en la que cada aspecto de la vida de los ciudadanos está bajo el control del Estado.

En la novela, el estado totalitario de Oceanía, liderado por el omnipresente Gran Hermano, mantiene un control exhaustivo sobre sus ciudadanos a través de la vigilancia constante y la manipulación de la información. Las "telepantallas", dispositivos que permiten tanto la transmisión como la recepción de información, se utilizan para espiar a los ciudadanos en sus hogares, mientras que la "policía del pensamiento" vigila cualquier signo de deslealtad o pensamiento subversivo.

Además, el Estado cuenta con una institución encargada de reescribir la historia para que se ajuste a la narrativa del Partido. Orwell basa esta idea en la práctica habitual de regímenes como el de la Unión Soviética que, a través de la propaganda y la censura, creaban un relato sobre el que cimentaban y justificaban sus actos. También la experiencia del autor en España, donde participó en la Guerra Civil del lado del bando republicano, fue fuente de inspiración para esta parte central de la historia. En su libro *Mi guerra civil española* dice "Ya de joven me había fijado en que ningún periódico cuenta nunca con fidelidad cómo suceden las cosas, pero en España vi por primera vez noticias de prensa que no tenían ninguna relación con los hechos, ni siquiera la relación que se presupone en una mentira corriente. (...) Estas cosas me parecen aterradoras, porque me hacen creer que incluso la idea de verdad objetiva está desapareciendo del mundo. A fin de cuentas, es muy probable que estas mentiras, o en cualquier caso otras equivalentes, pasen a la historia. ¿Cómo se escribirá la historia de la guerra civil española? (...) El objetivo tácito de esa argumentación es un mundo de pesadilla en el que el jefe, o la camarilla gobernante, controla no solo el futuro sino también el pasado. Si el jefe dice de tal o cual acontecimiento que

no ha sucedido, pues no ha sucedido; si dice que dos y dos son cinco, dos y dos serán cinco. Esta perspectiva me asusta mucho más que las bombas, y después de las experiencias de los últimos años no es una conjetura hecha a tontas y a locas."

En su novela, Orwell plantea un argumento asfixiante en el que los protagonistas van descubriendo las tramas de un gobierno totalitario donde la libertad de expresión o pensamiento es completamente nula, donde la moral es impuesta desde los poderes políticos y la vigilancia y el temor convierten la vida en sociedad en imposible. Para ello, Orwell emplea referentes de su tiempo: utiliza pantallas de televisión y cámaras como mecanismos de vigilancia, propaganda bélica para el control de la población, cuerpos de seguridad para la detención del disidente. Todos estos mecanismos ya han quedado anticuados en nuestro tiempo, pero los propósitos de cada uno de ellos son cada vez más evidentes.

En nuestra sociedad contemporánea, no necesitamos telepantallas para ser vigilados: nuestros teléfonos móviles, ordenadores, tarjetas de crédito y redes sociales proporcionan una cantidad de información sobre nosotros que podría hacer que incluso el Gran Hermano se sonrojara. Las herramientas de la CD permiten recopilar, almacenar, analizar y utilizar esta información de formas que Orwell apenas podría haber imaginado.

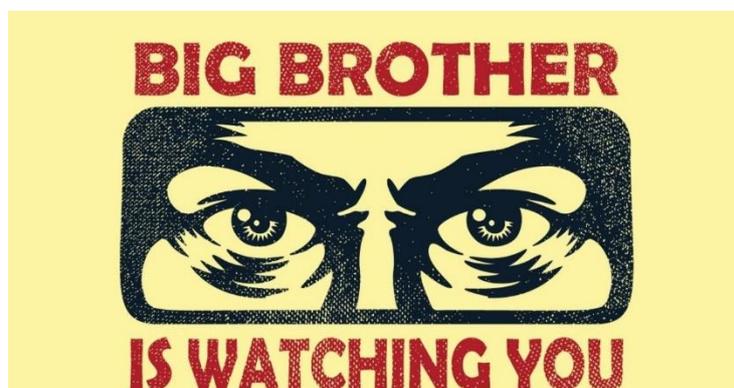
Los mecanismos de control social que fabula Orwell se materializan en nuestro tiempo en forma de algoritmos que determinan lo que vemos en las redes sociales, que directamente influyen en nuestra percepción de la realidad, y en el miedo a ser "cancelados" o condenados en las redes sociales, que obliga a las personas a autocensurarse.

Gilles Deleuze, en su ensayo corto *Postscript on the Societies of Control* (1992), examina esta problemática y argumenta que, al momento de su redacción, la sociedad estaba pasando de una fase de "sociedades disciplinarias", como se describen en las obras de Michel Foucault, a lo que él llama "sociedades de control".

En las sociedades disciplinarias, según Foucault, el poder se ejerce en espacios cerrados, como escuelas, fábricas, prisiones y hospitales, que se organizan en torno a un sistema de encierro y vigilancia. Cada uno de estos espacios tiene su propio conjunto de reglas y normas, y los individuos pasan de un espacio a otro a lo largo de sus vidas (Foucault, 1975). Deleuze argumenta que esta estructura está dando paso a las sociedades de control, donde el poder no se ejerce en espacios cerrados, sino que se extiende por todo el tejido social, penetrando en la vida cotidiana en formas mucho más sutiles y difusas. En las sociedades de control, según Deleuze, los mecanismos de control son continuos y sin límites, y los individuos son rastreados y monitoreados en tiempo real. Las tecnologías de la información, como las bases de datos y los

algoritmos, juegan un papel central en este nuevo tipo de control, permitiendo un tipo de vigilancia constante y omnipresente. Deleuze también sugiere que, en las sociedades de control, la identidad del individuo se vuelve más fluida y flexible. Los individuos ya no son moldeados y formados por las instituciones disciplinarias (como la escuela o la fábrica), sino que están sujetos a un constante proceso de "modulación" o ajuste, en función de una variedad de factores, como sus hábitos de consumo, su historial de salud o su perfil de riesgo.

Una sociedad estática, muda y sin ambiciones, como la que presenta Orwell, solo podía funcionar en términos literarios, más aún después de la victoria del bando capitalista en la Guerra Fría. Con el capitalismo y el auge de las tecnologías de la información se instaura una sociedad líquida, como la que define el sociólogo Zygmunt Bauman en *Modernidad Líquida* (1999), de naturaleza cambiante, volátil y efímera, en la que los individuos están constantemente incitados a consumir y a definirse a sí mismos a través de sus elecciones de consumo, y en la que las personas deben aprender a vivir en un mundo de cambio constante y riesgo incierto, que genera una sensación de inseguridad y ansiedad. Este es el terreno perfecto para establecer las sociedades de control de las que habla Deleuze, que requieren de esta incertidumbre de los individuos, para desde los poderes económicos y políticos establecer un marco de libertad que es al mismo tiempo una forma de control: los individuos son libres para reinventarse a sí mismos, pero sólo dentro de los parámetros establecidos por los sistemas de control.



Representación del personaje del Gran Hermano en 1984. Fuente: <https://www.culturagenial.com/es/novela-1984-de-george-orwell/>

## Vigilar y castigar

En 1975, el filósofo francés Michel Foucault publica *Vigilar y castigar. Nacimiento de la prisión*, obra en la que examina la transformación del castigo criminal en las sociedades occidentales

desde el siglo XVII hasta el siglo XIX, a la vez que desarrolla una teoría del poder y la vigilancia que sigue siendo influyente en la filosofía, la sociología, la criminología y otras disciplinas.

De acuerdo con Foucault, durante la Edad Media, el suplicio se utilizaba como un estricto método de demostración penal, cuyo objetivo era exhibir la verdad que se había obtenido a través del proceso penal, haciendo que el culpable se convirtiera en el anunciador de su propia sentencia al soportar el castigo físico en su cuerpo (recorrido por las calles, cartel, lectura de la sentencia en los cruces de caminos...). Adicionalmente, el suplicio también se consideraba un ritual político, amparado en el derecho de la Antigüedad, para el cual el delito era principalmente un ataque al soberano, la fuente de la ley. Por lo tanto, la pena no sólo debía reparar el daño causado, sino también servir como una venganza por la ofensa al rey.

A finales del siglo XVIII y principios del XIX, el castigo cambia dramáticamente. En lugar de castigos físicos públicos, las sociedades occidentales comenzaron a encarcelar a los criminales, sujetándolos a un régimen estrictamente regulado de trabajo y aislamiento. Foucault argumenta que este cambio fue parte de una transformación más amplia del poder social, de un régimen de "poder soberano" a uno de "poder disciplinario".

Según Foucault, este poder disciplinario, del que hablamos más arriba como eslabón previo a las sociedades de control de Deleuze, es un tipo de poder que no se ejerce simplemente a través de la fuerza o la prohibición, sino a través de la observación constante y la normalización. Este tipo de poder se ejerce no sólo en las prisiones, sino también en una variedad de otras instituciones, como las escuelas, los hospitales y los cuarteles. En todos estos lugares, las personas son observadas y evaluadas constantemente, y su comportamiento es corregido y normalizado de acuerdo con ciertos estándares.

Si bien es cierto que parece anacrónico hablar de sociedades disciplinarias cuando, supuestamente, nos encontramos ya en sociedades de control, como se verá, las características de una y otra organización son muy similares, dándose la principal diferencia en el uso de las tecnologías de la información en las descritas por Deleuze. Los conceptos que Foucault nos brinda en su obra son perfectamente adaptables a nuestro tiempo y al uso de las nuevas tecnologías, por lo que extenderlos en su desarrollo será de utilidad a la hora de buscar ejemplos concretos en los que la CD haya afectado a las relaciones entre el poder político y el pueblo.

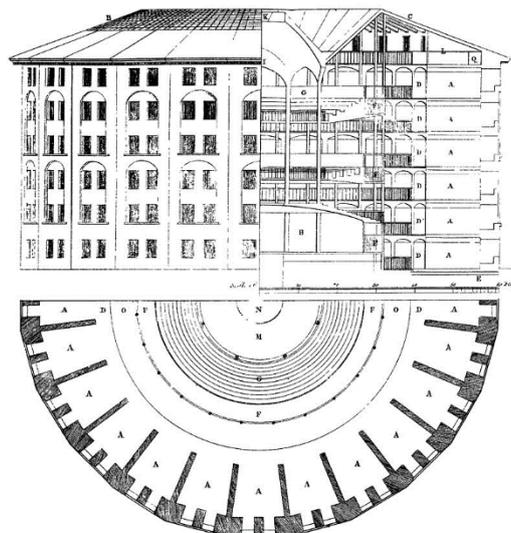
### Panóptico

Foucault retoma el concepto de "panóptico" del filósofo y reformador social británico Jeremy Bentham, que lo desarrolló en su obra *El Panóptico* (1791) como una forma de diseño

arquitectónico de prisiones, que tenía como finalidad permitir a un guardián, resguardado en una torre central, la capacidad de observar a todos los prisioneros, los cuales se alojan en celdas individuales dispuestas alrededor de la torre, desconociendo si están siendo vigilados o no. La idea de Bentham era que el Panóptico generara en el recluso una sensación constante y consciente de estar bajo observación, lo que aseguraría la operatividad autónoma del poder, incluso cuando no se esté ejerciendo de manera explícita y constante, dado que el prisionero no puede discernir cuándo está siendo vigilado y cuándo no.

Como decimos, Foucault retoma y expande este concepto, el cual no ve sólo como una innovación arquitectónica, sino también como una metáfora de la moderna sociedad disciplinaria y su preferencia por la vigilancia constante y la normalización. Así, Foucault argumentaba que las sociedades modernas, a diferencia de las sociedades anteriores que ejercían su poder a través de espectáculos públicos de violencia, como la tortura y las ejecuciones, habían evolucionado hacia formas más sutiles y psicológicas de control. Las prisiones, las escuelas, los hospitales, las fábricas, e incluso la familia, podrían ser vistas como instituciones panópticas en las que las personas se disciplinan a sí mismas por miedo a la vigilancia.

El Panóptico, para Foucault, representa la forma más pura de poder, ya que es un poder que se ejerce constantemente y cuya existencia siempre se infiere. Los individuos, conscientes de que podrían estar bajo vigilancia en cualquier momento, se convierten en sus propios carceleros, modificando su comportamiento para ajustarse a las expectativas sociales y normativas.



Diseño del panóptico de Bentham. De Jeremy Bentham - The works of Jeremy Bentham vol. IV, 172-3, Dominio público, <https://commons.wikimedia.org/w/index.php?curid=3130497>

Como explica Deleuze, las sociedades disciplinarias han desaparecido, la normalización de la sociedad y la represión en espacios cerrados ha dejado de ser posible con la consolidación absoluta del sistema capitalista, de la globalización y de las tecnologías de la información, que han restado valor e identidad a los espacios donde anteriormente se disciplinaba (familia, escuela, fábricas...) y han ampliado significativamente los campos de interacción de la sociedad. Ante esta insuficiencia nacen las sociedades de control, sustentadas por tecnologías que ya no necesitan de espacios disciplinarios, pues actúan y están presentes en la totalidad de las interacciones de los individuos. Podría decirse que ha cambiado el modo en que se ejerce la vigilancia, pero el panóptico no solo se mantiene inalterable como una estructura político-social en la que los guardianes vigilan sin ser vistos, sino que se ha perfeccionado siendo capaz de vigilar más y mejor.

Hasta aquí, hemos hablado de sociedades de control, sociedades líquidas y panópticos, todos ellos conceptos teóricos que tratan de describir la sociedad de nuestros días. Podemos pensar, y sería lógico, que todas estas ideas tienen una vertiente distópica cuyo germen se encuentra en obras de ciencia ficción como la ya mencionada *1984* de George Orwell, lo cual, en parte, tiene algo de razón si consideramos que las condiciones de vida, las posibilidades económicas y las libertades individuales, que forman el marco de actuación de los ciudadanos de los países más desarrollados, han alcanzado la mayor amplitud en la historia de la humanidad. En este contexto, la búsqueda de estas sociedades de control, o como queramos llamarlas, parece un intento por desenmascarar al sistema del capitalismo digital que tanto bien parece que nos ha hecho, darle la vuelta y ver todos aquellos trapos sucios que hacen que el sistema pueda continuar y expandirse continuamente. Sin embargo, desde aquí, aplaudimos y trabajamos por llevar a cabo esta labor de desvelamiento, no por el hecho de desacreditar al sistema o a las tecnologías sobre las que se soporta, sino porque solo desde una perspectiva crítica se pueden apreciar los peligros y fallos en fenómenos ya tan arraigados como al que nos referimos aquí, y solo una vez hecha esta aproximación se puede desarrollar una legislación coherente y responsable.

De esta manera, continuamos nuestra revisión alejándonos de las teorías y conceptos para centrarnos en unos hechos que, cada vez más, nos harán volver a las ideas de Foucault, Deleuze y Bauman.

## El caso de la República Popular China

En el año 2014, la República Popular China inicia un proyecto gubernamental con la intención de establecer un esquema nacional para rastrear la confiabilidad de los ciudadanos, las empresas y los funcionarios gubernamentales. Este sistema, conocido como Sistema de Crédito Social y gestionado por el planificador económico de China, la Comisión Nacional de Desarrollo y Reforma (NDRC), el Banco Popular de China (PBOC) y el sistema judicial del país, recopila información de fuentes tradicionales como registros financieros, criminales y gubernamentales, y de la recopilación de datos a través de la videovigilancia y la transferencia de datos en tiempo real.

La implantación de este sistema, aún pendiente de completarse a nivel nacional, tiene como objetivo recompensar las acciones que generan confianza en la sociedad y castigar lo contrario a través de un sistema de puntos que beneficiará o restringirá en determinados aspectos a los sujetos. El crédito social de cada individuo u organización se divide en dos aspectos, en principio, diferenciados:

- Solvencia financiera: Este indicador no se diferencia demasiado de los empleados en los países de Occidente por las entidades bancarias para clasificar a los usuarios en función de su capacidad para devolver dinero.
- Solvencia social: Este indicador es el que mayor controversia genera porque, desde el gobierno chino, se incide en la necesidad de aumentar el nivel de confianza en la sociedad a través de su oposición contra la corrupción, el fraude fiscal, las estafas, el plagio, la contaminación... La lista sigue y los criterios que exponen no parecen fijos, en ocasiones entrometiéndose en aspectos de la moral que dudosamente deben ser impuestos desde el poder político.

Se entiende, por tanto, que, igual que la calificación crediticia evalúa la solvencia financiera, la solvencia social evalúa la fiabilidad de las personas como miembros de la sociedad. Si bien es cierto que se esperaba la concreción del proyecto para el año 2020, a mayo de 2023 este aún no ha sido materializado en su totalidad, pues, mientras que la parte financiera sí está en funcionamiento, de una manera muy similar a las que encontramos en los países de Occidente, la parte social ha sido fruto de experimentos y controversias.

### ¿Cómo funciona el sistema de crédito social de China?

El sistema de crédito social de China, en su vertiente puramente social, plantea una dualidad fundamental basada en la recompensa y el castigo. De este modo, los individuos y empresas con

un historial crediticio favorable en todas las áreas reguladas pueden gozar de privilegios otorgados por el gobierno, como la inclusión en listas prioritarias para recibir subsidios gubernamentales. En contraparte, aquellos con un historial crediticio insatisfactorio podrían enfrentar sanciones que incluyen la divulgación pública de su información y la prohibición de participar en licitaciones gubernamentales, adquirir productos de lujo y salir del país.

Desde Occidente, la idea de un sistema que, basado en la vigilancia y la recopilación de datos, pueda definir cuáles son las acciones moralmente aceptables y castigar o premiar a los ciudadanos y las empresas en función de su comportamiento, ha generado gran controversia entre aquellos que comparan la situación en China con la imaginada por Orwell en su novela y aquellos que llaman a la prudencia y piden que se analice con detenimiento este sistema antes de lanzar valoraciones infundadas.

Lo cierto es que la vertiente social del sistema no ha sido implantada a nivel nacional, aunque los medios occidentales más sensacionalistas hablen de China en su totalidad, sino que, como el gobierno no ha conseguido avanzar demasiado en cuanto a su puesta en marcha, ha permitido a las administraciones locales experimentar con ella, lo que ha llevado a extrapolar las medidas de ciudades como Rongcheng al resto del territorio. En Roncheng, localizada a unos ciento cincuenta kilómetros de Pekín y con una población de medio millón de habitantes, se asignó automáticamente a cada ciudadano una puntuación base de mil puntos que, en función de sus buenas o malas acciones, podía aumentar o disminuir. Entre estas buenas acciones podemos encontrar comportamientos como participar en actividades de servicio comunitario, realizar donaciones a la caridad, pagar las facturas a tiempo o respetar las normas de tráfico. Las acciones castigables por el sistema incluyen, por ejemplo, cruzar la calle imprudentemente, hacer ruido excesivo, ensuciar espacios públicos, publicar información falsa en internet...

Los beneficios a los que los ciudadanos chinos con una buena puntuación social pueden acceder son el acceso preferencial a préstamos y productos financieros, más oportunidades de trabajo, viajes más accesibles y beneficios sociales, como acceso preferencial a la vivienda pública, a las escuelas o a la atención médica, mientras que las desventajas a las que se enfrentan por una mala puntuación van desde la negación de todas las anteriores hasta un estigma social derivado de la exhibición pública de información privada. Como vemos, el sistema es controvertido cuanto menos, pues además de establecer unas directrices morales más allá de la ley que limitan la libertad de los ciudadanos, plantean un conjunto de beneficios y desventajas que se mezclan con aspectos financieros, como el acceso a préstamos, que, en principio, deberían estar regulados por la parte crediticia del sistema.

Ahora bien, el de Rongcheng fue solo un experimento que, además, ha recibido el rechazo del Gobierno central chino por castigar conductas triviales como no usar la mascarilla o no reciclar correctamente, acciones que la legislación actual no recoge, a lo que han instado a las administraciones locales a limitarse a castigar aquellas acciones consideradas ilegales.

Sin embargo, la controversia no termina aquí, porque el proyecto sigue en pie y su implantación a nivel nacional podría estar más cerca que nunca, habiendo publicado el gobierno un nuevo proyecto de ley a mediados de noviembre de 2022, en el que se da la información hasta ahora más fiable del sistema. En él, se recalca la necesidad de que los castigos deben estar en perfecta concordancia con la legislación vigente, pero esto no tranquiliza en absoluto, porque la legislación vigente también puede ser injusta e intrusiva (Feced, 2022) (Yang, 2022).

### ¿Un panóptico real?

Sumado a un más que posible sistema de crédito social, encargado de evaluar el comportamiento cívico de los ciudadanos y la confiabilidad que la sociedad puede depositar en cada individuo, tenemos que sumar el importante desarrollo e implementación de tecnologías de CD desplegado por el Gobierno chino para monitorear la conducta de los ciudadanos.

Sobre este creciente desarrollo en tecnologías de la información y los usos que el Gobierno chino está dando a las herramientas y tecnologías de la CD se han escrito numerosos artículos, trabajos académicos y libros en los que, en la mayoría de las ocasiones, se condenan las acciones de la esfera política del país oriental por atentar contra la privacidad y las libertades individuales de los ciudadanos chinos.

En un extenso artículo, publicado en *The Atlantic* en septiembre de 2020, titulado *China's Artificial Intelligence Surveillance State Goes Global* el periodista Ross Andersen discute como China está utilizando la inteligencia artificial (IA) y otras tecnologías para construir un sistema de vigilancia y control social sin precedentes, estableciendo una estrategia que busca convertir a China en líder mundial en IA para 2030. El texto menciona que China ya tiene cientos de millones de cámaras de vigilancia en funcionamiento y que el gobierno espera lograr una cobertura total de video de áreas públicas clave pronto. La IA se utiliza para analizar las imágenes recogidas por estas cámaras en busca de amenazas de seguridad.

Andersen también comenta el proyecto desarrollado por la empresa Alibaba, denominado "City Brain", que se propone como un centro de datos automatizado capaz de sintetizar flujos de datos de una multitud de sensores distribuidos en un entorno urbano. El City Brain tiene aplicaciones benignas, como contar personas y automóviles para ayudar con el tiempo de los semáforos rojos

y la planificación de líneas de metro o facilitar la recogida de residuos con sensores en los contenedores de basura, pero también podemos encontrar aplicaciones más sensibles para los ciudadanos, sobre todo cuando combinamos este sistema con tecnologías de reconocimiento facial, sensores en vehículos autónomos y dispositivos del Internet de las cosas.

Al artículo no le falta su dosis de catastrofismo, y sugiere que versiones futuras de City Brain podrían ser capaces de leer pensamientos no pronunciados a través de interfaces cerebro-computadora y podrían ser utilizados para monitorear las reacciones emocionales a los estímulos políticos, planteando, además, el escenario de un estado autoritario que podría alimentar cada detalle de la actividad neural de un ciudadano en una base de datos del gobierno, creando perfiles de datos detallados y en tiempo real para cada ciudadano, que podrían usarse para evaluar continuamente el "riesgo" que cada individuo representa para el poder del Partido, y el sistema podría tomar medidas automatizadas, como restringir los movimientos de una persona, si su "factor de riesgo" aumenta. Por supuesto, esto se acerca más a lo que diría a un futurólogo que a lo que diría un periodista reputado, pero intentaremos no tenerlo en cuenta y limitarnos a los hechos fehacientes.

Lo que es cierto es que China está centrando su atención en la vigilancia de los lugares públicos y está automatizando el análisis de la información a través de herramientas de IA. También en el entorno digital, China es de sobra conocida por la profunda censura y filtrado que realiza en los contenidos online a través del conocido Gran Cortafuegos (The Great Firewall), que es un sistema basado en técnicas de análisis de datos para examinar y bloquear sitios web, palabras clave y contenido que el gobierno considera inapropiado o subversivo.

La combinación de la vigilancia masiva en lugares públicos, la censura y filtrado en internet y el sistema de crédito social, nos lleva a un escenario en el que la totalidad de las interacciones sociales de la población china está, si no vigilada, por lo menos controlada por herramientas y tecnologías de CD, que responden directamente a los intereses del gobierno autoritario dirigido por Xi Jinping. Todas estas acciones, si el sistema de crédito social termina de implantarse a nivel nacional, son evaluables como positivas o negativas con las consecuentes consecuencias. La población china estaría en una situación en la que cada acción, sea en un lugar público o privado, está bajo una cierta vigilancia que obligará a respetar las indicaciones morales y legislativas de un estado completamente autoritario. Volviendo a la metáfora del panóptico, tenemos un vigilante al que no podemos ver, pero sabemos que vigila, situado en una posición preferente que permite observar todo el resto de estancias de la estructura. Los individuos vigilados se ven obligados a actuar conforme a las normas impuestas.

## El escándalo de Cambridge Analytica

Uno puede pensar que la situación que parece acercarse a la China de Xi Jinping no es para nada asombrosa, teniendo en cuenta la deriva política de su líder, que, entre otras muchas medidas autoritarias, mantiene campos de concentración de uigures. Es de esperar que, en una dictadura como la china, donde la censura y la violación de los derechos humanos es común desde hace varias décadas, las nuevas tecnologías contribuyan a ejercer una represión aún más dura y eficiente, pero extrapolar los usos que la CD está recibiendo en el país asiático a los países democráticos parece un sinsentido si tenemos en cuenta las significativas diferencias que encontramos entre los dos sistemas políticos.

Sin embargo, el uso de estas tecnologías por parte de los países democráticos, aunque están aún lejos del caso chino, ya muestran ciertas violaciones a la privacidad y la autonomía de los ciudadanos. El caso de la empresa de consultoría política Cambridge Analytica es uno de los más famosos en los que los poderes políticos, en el Reino Unido y Estados Unidos específicamente, han empleado la CD con el objetivo de influir y manipular la opinión de los ciudadanos en favor de unas ideas políticas concretas.

En su libro, *Mindf\*ck: Cambridge Analytica and the Plot to Break America* (2019), Christopher Wylie, antiguo director de investigación en Cambridge Analytica, cuenta su visión personal de la interferencia de la empresa en las elecciones de 2016 en los Estados Unidos y el referéndum del Brexit en Reino Unido. Así, Wylie argumenta que Cambridge Analytica recolectó y explotó grandes cantidades de datos personales para construir perfiles psicológicos detallados de los votantes, sobre los que dirigir campañas de desinformación y propaganda, buscando influir en su comportamiento de voto.

El libro detalla cómo Cambridge Analytica trabajó con la campaña de Trump en 2016 y con la campaña a favor del Brexit en el Reino Unido, desempeñando un papel crucial en la manipulación de la opinión pública y en la promoción de la desinformación durante estos dos eventos políticos trascendentales. Para ello, Cambridge Analytica, a través de una aplicación de encuestas, extrajo indebidamente los datos de millones de usuarios de Facebook, que posteriormente se utilizaron para perfilar y segmentar a los votantes, permitiendo a la compañía dirigirse a ellos con mensajes políticos personalizados.

El escándalo de Cambridge Analytica tuvo amplias y profundas implicaciones en varios niveles, desde la percepción pública hasta la regulación de la industria tecnológica:

- **Repercusiones políticas:** El escándalo de Cambridge Analytica tuvo importantes repercusiones políticas, porque ambas campañas apoyadas por la consultora salieron victoriosas en las votaciones, convirtiéndose Donald Trump en presidente de los Estados Unidos y saliéndose de la Unión Europea el Reino Unido. La revelación de que los datos de los usuarios fueron utilizados para influir en estas votaciones ha llevado a un mayor escrutinio de las tácticas de campaña y ha alimentado los debates sobre la interferencia en las elecciones y la legitimidad de los resultados.
- **Cambios en las políticas de privacidad de los datos:** El caso de Cambridge Analytica forzó a las empresas de redes sociales, y en particular a Facebook, a revisar y ajustar sus políticas y prácticas de privacidad de los datos. Facebook implementó cambios significativos en la forma en que maneja y protege los datos de los usuarios, y cómo comparte esos datos con terceros.
- **Regulación más estricta:** El escándalo ha llevado a una mayor regulación y supervisión de las empresas de tecnología. En particular, la Unión Europea implementó el Reglamento General de Protección de Datos (GDPR), que proporciona a los individuos un mayor control sobre sus datos personales y establece penas más severas para las empresas que violan estas regulaciones.

A través del escándalo de Cambridge Analytica, podemos ver un uso disimulado de la CD para influir y manipular la opinión pública. Mientras que la República Popular China opta por un control directo y explícito a través de un sistema de vigilancia con incentivos y castigos, los países democráticos ejercen una sutil manipulación a través de los contenidos en las redes sociales. Aunque la regulación haya sido revisada, el uso que desde los poderes políticos hacen de estos medios de comunicación de masas escapa de la legislación, pues es el propio funcionamiento de estas plataformas el que favorece la difusión de desinformación y la consecuente manipulación de la opinión.

## Posverdad

A partir de las campañas políticas mediadas por Cambridge Analytica y su posterior desvelamiento, el concepto denominado "posverdad" comenzó a crecer en popularidad hasta ser considerado uno de los rasgos definitorios de nuestro tiempo. El término tiene su origen en la década de 1990 de la mano del escritor, de origen serbio y residencia estadounidense, Steve Tesich en un artículo, publicado en la revista *The Nation*, donde reflexionaba sobre la guerra del Golfo Pérsico y la perspectiva del pueblo americano ante el conflicto. Desde entonces, el término

fue empleado en distintas ocasiones refiriéndose a los intentos desde los distintos gobiernos de enmascarar la realidad con fines políticos y propagandísticos.

No obstante, fue en el entorno de 2016 cuando el término "posverdad" cobró notoriedad. En este año, el Diccionario Oxford lo seleccionó como la palabra del año, definiéndola como "relativa a, o que denota, circunstancias en las que los hechos objetivos influyen menos en la formación de la opinión pública que los llamados a la emoción y la creencia personal". La definición del Diccionario de la Lengua Española incluye la consideración como acto deliberado: "Distorsión deliberada de una realidad, que manipula creencias y emociones con el fin de influir en la opinión pública y en actitudes sociales."

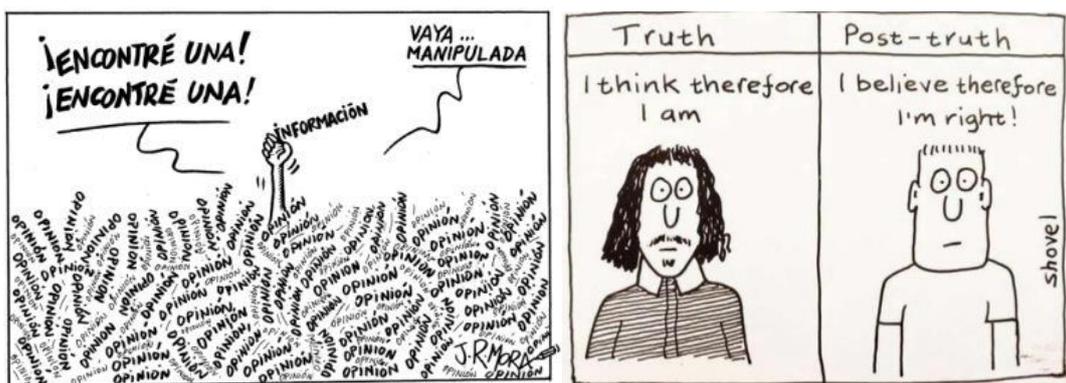
Sobre el concepto escribió en 2019 el filósofo y académico Lee McIntyre en su obra titulada *Post-Truth*, donde lleva a cabo una introducción al término y a las relaciones entre este y otras instituciones como la ciencia y la política. Para McIntyre, vivimos en una era de "posverdad", en la que la verdad objetiva se ha vuelto menos influyente en la formación de la opinión pública que las apelaciones a la emoción y la creencia personal, lo que conduce a la población a la negación de fenómenos como el cambio climático o la eficacia de las vacunas. McIntyre también se enfoca en cómo la posverdad ha influenciado la política, argumentando que los políticos y los medios de comunicación a menudo explotan esta circunstancia para avanzar en sus agendas.

También otros autores, como James Ball o Tom Nichols, se han pronunciado sobre la posverdad en sus obras. Ball, en *Post-Truth: How Bullshit Conquered the World* (2017), localiza en la tecnología y las redes sociales una de las causas más determinantes en la propagación de la posverdad. Las plataformas de redes sociales, argumenta Ball, permiten la rápida difusión de información, independientemente de su veracidad, y los algoritmos a menudo promueven el contenido más polémico y divisivo, lo que puede incluir noticias falsas. Esto, como hemos visto, ha sido aprovechado por los políticos que, utilizando la desinformación y estas noticias falsas, han conseguido ventajas políticas a costa de la polarización de la población. Para Ball, también los medios de comunicación han contribuido a esta situación, pues han tenido que adaptarse a las redes sociales sustituyendo el periodismo de calidad por el contenido sensacionalista y emocionalmente cargado.

Por su parte, en *The Death of Expertise: The Campaign Against Established Knowledge and Why it Matters* (2017), Tom Nichols explora cómo la sociedad contemporánea ha desarrollado un rechazo y desdén generalizado por los expertos y el conocimiento establecido. Nichols expone que, con el auge de Internet y las redes sociales, el conocimiento se ha democratizado, lo que ha llevado a que muchas personas creen que su conocimiento no experto es tan válido como el

conocimiento experto. Nichols sostiene que, aunque la democratización del conocimiento puede ser positiva, también puede conducir a una falta de respeto por la experiencia y a la propagación de desinformación. Así, Nichols señala que, aunque vivimos en la era de la información, también vivimos en la era de la confusión, pues la gran cantidad de información disponible puede ser abrumadora y puede llevar a las personas a confiar en fuentes de información no confiables o a rechazar la información que no se alinea con sus creencias preexistentes.

En este punto, es interesante volver a lo dicho en la introducción del presente epígrafe, donde definíamos a los políticos como generadores de conflictos y debates insustanciales, que, deliberadamente, enfrentan a la población. También recalcábamos la incapacidad de los ciudadanos normales de atender a las cuestiones políticas directamente, lo que provoca que, ante discursos más enfocados a la exaltación de las emociones que al ejercicio de la razón, estos se posicionen con aquellos que se identifiquen más con sus creencias irracionales. Así, este modo de conducta no es solo ante la política, sino que se extiende al resto de campos, llegándose a situaciones en las que hechos completamente ciertos e indiscutibles, como la pandemia de la COVID-19, son cuestionados infundadamente. Sin embargo, es en el campo de la política donde estos hechos cobran una mayor importancia, porque los distintos partidos, concedores ya del funcionamiento de la redes sociales y los algoritmos que hay detrás, basan las campañas electorales en mentiras o verdades a medias, inventan problemáticas para ocultar escándalos (pantallas de humo) y, en definitiva, debilitan la capacidad democrática de los pueblos, que se ven indefensos e inoperantes ante una política que basa su actividad en generar conflictos que buscan más el beneficio propio que el del grupo.



Viñetas cómicas sobre el concepto de posverdad. Fuente: <https://cbaglobal.com.ar/la-democracia-en-la-era-de-la-posverdad/>

## Conclusiones

Para concluir con nuestra revisión de las implicaciones políticas de la CD volvemos de nuevo a la novela de Orwell y las principales formas de represión que el Gran Hermano, el líder absoluto del Estado donde se desarrolla la acción, pone en funcionamiento. Principalmente, estos son:

- Dispositivos de vigilancia: Nos referimos a las telepantallas a través de las cuales el Gran Hermano observa y comunica la información.
- Manipulación de la información: Función llevada a cabo por el Ministerio de la Verdad, encargado de llevar a cabo una reescritura continua de la historia para adecuarla a los intereses políticos.
- Cuerpos de seguridad: Objetivada en la policía del pensamiento, cuyas funciones se centran en el mantenimiento de la verdad oficial del Partido mediante la detención inmediata de cualquier disidente.

Lo que podía parecer una trama de ciencia ficción distópica ha resultado, como hemos ido analizando, que tiene más similitudes con las sociedades contemporáneas de lo que hubiéramos imaginado. El mejor ejemplo lo expone la República Popular China, donde la vigilancia se extiende a todos los ámbitos de la sociedad, desde internet hasta los lugares públicos y privados, y las leyes y directrices morales se hacen cumplir mediante un sistema de puntos en el que, si sales malparado, puedes ser incapacitado para funciones tan elementales en nuestros días como pedir un préstamo o salir del país. China ha logrado alcanzar esa situación que se da en la novela en la que los ciudadanos ni siquiera se pueden permitir disentir del sistema, que es capaz de excluirlos si eso pasará. La policía del pensamiento de la obra de Orwell encuentra también sus referencias en las numerosas detenciones y desapariciones de ciudadanos que disienten del régimen, siendo la masacre de Tiananmén en 1989 el mejor ejemplo de esto ((Zhang & Gombeaud, 2020)), que, inmediatamente y hasta el día de hoy, ha sido ocultada por el Gobierno y los medios afines a él, en correspondencia con los actos del Ministerio de la Verdad en la novela.

También en las sociedades occidentales podemos encontrar similitudes, sobre todo en los dos primeros puntos, con una manipulación de la información que se realiza en tiempo real a través de la posverdad y unos dispositivos de vigilancia, como hemos dicho, más sutiles, que, coordinando las organizaciones políticas con las grandes empresas tecnológicas, son capaces de influir significativamente en la conducta y los pensamientos de la ciudadanía.

Siguiendo la práctica del anterior epígrafe, donde exponíamos los principales puntos en los que la legislación debe actuar, si es que aún no lo hace, procedemos a enumerar los relativos a la práctica política:

- Regulaciones sobre publicidad política en línea: La legislación debe regular cómo y cuándo se puede utilizar la publicidad política en línea y en las redes sociales, incluyendo la divulgación de quién paga por los anuncios y cómo se seleccionan los destinatarios.
- Derecho al olvido: Los individuos deben tener el derecho a que su información personal sea eliminada de las bases de datos, especialmente en el caso de que se utilice para fines políticos.
- Leyes contra la desinformación: Implementación de leyes que penalizan la creación y distribución de desinformación o noticias falsas, especialmente durante los periodos electorales.
- Regulación de los sistemas de crédito social: Las legislaciones deben proteger a los ciudadanos de sistemas de crédito social que pueden ser utilizados para la vigilancia masiva y el control social.
- Auditorías de algoritmos: Para garantizar que los algoritmos utilizados en la política (por ejemplo, para la segmentación de votantes) son justos y no discriminatorios, deberían ser obligatorias las auditorías de algoritmos.

### 3. Sociedad

Llegamos al último punto de nuestra revisión de las implicaciones de la CD, habiendo ya visto numerosos aspectos de la sociedad y los individuos que la conforman. Ya habíamos indicado al inicio que la sociedad se define y modela en función de los poderes económicos y políticos a los que está subordinada. Es por eso que hemos hablado ampliamente de cómo percibe la sociedad los distintos fenómenos económicos y políticos, y como se relaciona con ellos, porque sería imposible comprenderlos sin saber sobre quiénes se están desarrollando.

La distinción entre estos tres aspectos (economía, política y sociedad) es harto complicada, pues son, en realidad, un conjunto que funciona estableciendo relaciones continuas entre ellos. Por eso, desde un principio dijimos que nos centraríamos en exponer como la CD había modificado las relaciones entre los agentes económicos, en el punto sobre la economía; las relaciones entre las esferas políticas y los ciudadanos, en el punto sobre la política; y las relaciones entre los individuos como pertenecientes a una sociedad, que será de lo que nos ocuparemos a continuación. A través de esta división, creemos que es más sencillo localizar aquellos posibles abusos, injusticias o perjuicios que las tecnologías de CD puedan ocasionar, pero siempre debemos tener en cuenta que esta división es solo virtual, con fines analíticos, y que en la realidad no pueden aislarse unos de otros.

Así, en este último apartado, nos adentramos en el núcleo de las interacciones humanas: las relaciones entre los individuos en la sociedad, que desde la aparición y el rápido desarrollo de la CD se han reconfigurado de manera drástica y profunda. Lo que una vez fue el dominio exclusivo de los encuentros cara a cara y las interacciones personales se ha trasladado cada vez más a la esfera digital, creando nuevas formas de asociación y vinculación que todavía estamos tratando de comprender plenamente. Las redes sociales, por ejemplo, que serán tema de especial importancia, han permitido la rápida difusión de información y la interacción en tiempo real entre personas de todo el mundo, rompiendo las barreras geográficas y temporales que antes limitaban nuestra capacidad de comunicación. La CD, como principal motor de estas plataformas, ha permitido una interconexión global sin precedentes, uniendo a individuos de diferentes culturas, antecedentes y perspectivas en una única red global.

Sin embargo, esta conectividad también ha llevado a nuevas formas de alienación y desconexión, ya que los algoritmos y las interfaces digitales, diseñados para adaptarse a nuestros gustos, intereses y comportamientos, aunque tienen como objetivo mejorar la experiencia del usuario, también tienen implicaciones significativas para cómo los individuos se relacionan entre sí y con

la sociedad en general. En lugar de exponer a los usuarios a una diversidad de perspectivas, estos sistemas a menudo refuerzan las ideas preexistentes y crean cámaras de eco digitales, lo que puede polarizar aún más las opiniones y limitar nuestra capacidad para entender y empatizar con los demás.

A continuación, iremos viendo algunos conceptos que definen las relaciones desarrolladas entre los individuos de nuestro tiempo. Por supuesto, la importancia en este apartado de las redes sociales será vital, pues son las plataformas donde se desenvuelven la mayoría de interacciones entre individuos y es donde estos adoptan formas de comportamiento y de comunicación que proyectan más allá del entorno digital.

### Las redes sociales como campo de batalla

Facebook, Twitter, YouTube, Instagram, Pinterest, Snapchat, TikTok, son todos ellos nombres que a día de hoy son imposibles de ignorar. Es en estas plataformas, que también hemos trabajado desde el punto de vista económico, donde la gran mayoría de las interacciones sociales se desarrollan hoy en día. Cada una de estas ha sabido acaparar una parte del mercado de la comunicación masiva, desde compartir imágenes hasta realizar debates, pasando por el entretenimiento más simple y a la vez rentable. Desde la creación de Facebook en 2005, impulsados por el desarrollo de tecnologías de la información y digitalización, han ido apareciendo estos servicios, que, poco a poco, han ido adquiriendo una importancia cada vez mayor, hasta el punto de ser una de las piezas clave para analizar la sociedad actual.

Lee Rainie, director de la *Internet and Technology Research* en Pew Research Center, y Barry Wellman, profesor de la Universidad de Toronto, nos ofrecen una introducción perfecta al fenómeno de las redes sociales en su obra *Networked: The New Social Operating System* (2012), donde analizan cómo las redes sociales y el internet han cambiado nuestra forma de vivir, trabajar y comunicarnos. La tesis central del libro es que hemos evolucionado desde una sociedad de "grupos", en la que los individuos solían pasar la mayor parte de su tiempo y energía en grupos sociales cohesivos y bien definidos (como la familia, el lugar de trabajo, la iglesia, etc.), hacia una sociedad de "redes", en la que las personas mantienen y cultivan sus relaciones personales a través de redes sociales, lo que permite mantener conexiones con un mayor número de personas, pero también puede llevar a relaciones más superficiales.

La obra aborda la influencia de las redes sociales en el trabajo, en las relaciones personales, en el aprendizaje y en la política; siendo la parte de las relaciones personales la que más nos interesa

en este punto. Para los autores, el paso de la sociedad de grupos a la sociedad de redes ha llevado a la formación de "redes personales" en lugar de grupos sociales, siendo estas redes personales mucho más fluidas y flexibles que los grupos tradicionales, y formándose a menudo alrededor del individuo, en lugar de alrededor de un lugar o una institución. Este cambio ha tenido tanto ventajas como desventajas. Por un lado, ha permitido a las personas mantenerse en contacto con un mayor número de personas y ha proporcionado más oportunidades para la colaboración y el intercambio de ideas. Por otro lado, también ha llevado a un aumento en la superficialidad de las relaciones, ya que las personas pueden mantener relaciones con otras personas sin tener que comprometerse profundamente con ellas, lo cual puede aumentar la sensación de aislamiento y soledad, ya que las interacciones en línea a menudo carecen de la intimidad y la cercanía de las interacciones cara a cara.

Las ideas de estos dos autores son evidentemente valiosas en relación a la transformación desde una sociedad organizada en grupos definidos hasta una sociedad organizada en redes indefinidas, fluidas y superficiales, que correspondería a la "cibersociedad" que Manuel Castells define en su obra *La Galaxia Internet* (2001) una década antes, caracterizada por sustituir los espacios físicos tradicionales por un "espacio de flujos" donde la información, los recursos y el poder fluyen a través de las redes digitales. En este texto, Castells ya aborda la capacidad que la tecnología tiene para influir en nuestras identidades y, en última instancia, en la forma en que nos relacionamos con el medio.

Sin embargo, Rainie y Wellman, aunque aportan una valiosa visión sociológica de las redes sociales, no abordaron la cuestión de la organización de la información en estas plataformas, que es un punto clave para entender cómo funcionan y cuáles pueden ser sus efectos sobre la sociedad. P.W. Singer y Emerson T. Brooking, en cambio, sí centran su atención en este aspecto en su obra *LikeWar: The Weaponization of Social Media* (2011), donde explican que los algoritmos que impulsan las redes sociales están diseñados para maximizar la participación del usuario, ya sea exponiéndolo a contenido acorde a sus creencias preexistentes o a contenido que las niega completamente y genera una reacción negativa en él. De esta manera, el usuario, privado de posiciones intermedias, se refuerza en sus creencias, aislándose en lo que Eli Pariser, en su obra *The Filter Bubble: What the Internet Is Hiding from You* (2011), denomina "burbujas de filtro", donde los usuarios solo ven noticias e información que refuerzan sus puntos de vista existentes.

Los algoritmos que organizan la información en las redes sociales, por tanto, funcionan de manera que buscan siempre provocar una reacción emocional fuerte en el individuo. Para ello,

en base a toda la información recopilada y procesada en torno a cada usuario, diseñan una elección arquitectónica de la plataforma específica capaz de generar reacciones en ellos, ya sean positivas o negativas. Por supuesto, los algoritmos no pueden discernir entre información falsa o verdadera, sino que simplemente actúan en función de maximizar las interacciones, lo que provoca estas burbujas de filtrado, en las que los usuarios se polarizan y se convierten en blanco de las manipulaciones de los grandes discursos.

Esta forma de organizar la información en las redes sociales es duramente criticada por Jaron Lanier en *Ten Arguments for Deleting Your Social Media Accounts Right Now* (2018), donde explora cómo las redes sociales, además de llevar a la polarización a través de las burbujas de filtrado, o, como él las llama, "cámaras de eco"; estas son empleadas para la manipulación y el control de los usuarios. Lanier argumenta que los algoritmos de las redes sociales aprenden de nuestros comportamientos en línea y utilizan esa información para predecir y manipular nuestras acciones futuras, erosionando nuestra capacidad de autodeterminación. También retoma el concepto de "posverdad", explicando que las redes sociales han erosionado nuestra capacidad para discernir la verdad al inundarnos de información falsa y engañosa, y señala, además, que las redes sociales pueden tener efectos perjudiciales en la salud mental de los usuarios, citando estudios que muestran una correlación entre el uso de estas plataformas y el aumento de la ansiedad, la depresión y otros problemas de salud mental. Por último, Lanier explora la posibilidad de que las redes sociales, cuyo diseño se dirige a monopolizar nuestra atención con el fin de maximizar los ingresos publicitarios, puede llevar a una adicción a estas y a un detrimento de otras áreas de nuestras vidas.

Conociendo todo esto, podemos llegar al título que da nombre a este punto, "las redes sociales como campo de batalla", donde mi intención es dar a entender la deriva que estas plataformas, sobre todo aquellas orientadas a compartir opiniones, como Facebook o Twitter, terminan alcanzando por la estructura que les da forma. Una plataforma donde el objetivo es mantener a los usuarios pegados a la pantalla, los métodos solo pueden ser, como ya hemos explicado, generar la máxima atención posible al contenido, lo cual solo es posible cuando este concuerda perfectamente con lo que el usuario quiere ver o se distancia diametralmente provocando obstinación o repulsión al usuario. Así, los usuarios interesados en ciertos temas, como puede ser la política por ejemplo, solo verá contenido sobre política, y no sobre todas las opiniones o ideologías políticas, sino solo sobre aquellas que generen una reacción sobre él, buena o mala, dando la impresión de que en las redes sociales se mantiene una disputa constante entre esas dos posiciones, lo cual los usuarios extrapolan al mundo real, donde las opiniones son más diversas y fundadas, ya que la mayoría de las que se viralizan y acaban frente a nuestras pantallas

en las redes sociales suelen ser las más extremas y, por tanto, carentes de justificación racional. Una vez los usuarios se encuentran en esas cámaras de eco, la manipulación por parte de agentes capaces de aprovechar las cualidades de los algoritmos se vuelve realmente accesible, lo cual lleva a una mayor polarización, una mayor adicción y los consecuentes problemas de salud mental que, ante un clima de estrés y discusión continuo, pueden perfectamente aparecer.

## Ideología y tribalismo

Conociendo el funcionamiento de estos entornos de flujos de información, podemos abordar la forma en que distintas instituciones los han empleado en su beneficio. Las redes sociales, aunque nacieron con el objetivo de ofrecer plataformas donde compartir y consumir información, conectar a personas de todas partes del mundo y, de alguna manera, disolver esas diferencias, sean geográficas o culturales o idiomáticas, se han convertido en una herramienta perfecta para generar, difundir y reproducir ideas, que cargadas de ese factor emocional necesario del que antes hablábamos se convierten en ideologías, que antes que disolver diferencias las perpetúan.

El tema de las ideologías y la proliferación de los "-ismos" lleva acaparando la atención de filósofos, sociólogos y psicólogos ya desde el siglo pasado, pero es ahora, en la era de las redes sociales, cuando estos movimientos y agrupaciones están apareciendo cada vez más rápido y cada vez más fuertes en el momento de su concepción. Las características de las redes sociales promueven este tipo de agrupaciones, a través de los *hashtags* o los foros, que atraen usuarios afines a los movimientos y, por tanto, emocionalmente implicados.

La obra del filósofo marxista Louis Althusser, *Ideología y aparatos ideológicos del estado* (1970), es interesante porque introduce el concepto de "aparatos ideológicos", que, si bien él se refería exclusivamente a los que empleaban los Estados (educación, religión, familia, medios de comunicación), también concuerda con la función de las redes sociales en la actualidad. Las ideologías, para Althusser, son sistemas de creencias, valores y prácticas culturales que, a través de los aparatos ideológicos, son materializadas en acciones y prácticas cotidianas, que moldean las percepciones y los comportamientos individuales. Así, las redes sociales serían perfectos aparatos ideológicos que, mediante la generación de debates y la difusión del contenido, despiertan acciones y prácticas recurrentes en los usuarios, como revisar continuamente el *feed* de noticias o suscribirse a ciertos contenidos. La obra también introduce el concepto de "interpelación", que es el proceso a través del cual los individuos son reclutados en ideologías,

lo cual, en el contexto de las redes sociales, es automatizado en el funcionamiento de los algoritmos. Pudimos verlo cuando estudiamos el caso de Cambridge Analytica, que fue capaz de interpelar o llamar a los usuarios para apoyar a ciertos bandos políticos.

Es interesante tener en cuenta que, mientras Althusser solo concebía los aparatos ideológicos al alcance de los Estados, las redes sociales han cambiado este paradigma, porque ahora cualquier grupo bien organizado puede emplear esta herramienta para generar discusión e interpelar a usuarios. De alguna manera, las ideologías ya no están solo al servicio de los poderes políticos o de las oposiciones a estos, sino que la capacidad de generar un pensamiento alternativo a estos ha aumentado. No quiero referirme solo a pensamientos políticos o sociales, como puede ser el feminismo, el liberalismo o el socialismo, sino a pensamientos e intereses de cualquier tipo. La masividad de usuarios en las redes sociales y la democratización de la información ha permitido que los intereses de cualquier individuo encuentren en las redes sociales una asociación que comparta sus inquietudes, organizando así a la población en grupos específicos. Plataformas como Reddit cuentan con millones de foros que abarcan prácticamente la totalidad de los temas, desde la aeronáutica hasta los tebeos.

Esta necesidad de pertenecer a grupos con personas que compartan las creencias y valores de cada uno la explora el psicólogo social Jonathan Haidt en *The Righteous Mind: Why Good People Are Divided by Politics and Religion* (2012), donde sostiene que las personas son inherentemente tribales y buscan compañeros que, además de compartir sus ideas, las protejan junto a ellos. Esto lleva a la polarización, de los míos es lo mejor, y a la demonización de "los otros", porque sus ideas van contra las mías. En un contexto en el que los grupos ya no aspiran a tomar o mantener el control de una sociedad, la creación de estas tribus digitales, defensoras, por ejemplo, de las películas clásicas, entran en un enfrentamiento inútil e insustancial contra otras tribus digitales, defensoras, por ejemplo, de las películas de superhéroes. Esto se da a todos los niveles, aficionados del Real Madrid contra aficionados de Barcelona, fútbol contra baloncesto, películas contra series, comunismo contra capitalismo, hombres contra mujeres. Las redes sociales se convierten en ese campo de batalla del que hablamos en el anterior punto, donde la crispación y el conflicto no para de crecer, pero nunca explota debido a la ausencia de espacio físico en la plataforma.

Esta tendencia al tribalismo es natural, pero esta, potenciada con el aparato ideológico de las redes sociales, provoca una situación que el filósofo Jean Baudrillard denomina "hiperrealidad", en la que las representaciones de la realidad se perciben como más reales que la realidad misma. Esto lo aborda en su obra *Simulacres et Simulation* (1981) y, aplicado a las redes sociales,

podemos ver cómo los perfiles de los usuarios, las publicaciones, los me gusta y los seguidores se convierten en simulacros que representan a los individuos y a sus vidas. En este espacio, los debates y discusiones de las redes sociales a menudo se consideran más "reales" o más importantes que las experiencias y relaciones humanas fuera de las plataformas digitales. Las personas, que ya no han necesitado de espacios físicos y estables para establecer relaciones, se han servido de las redes sociales para moldear su vida social a su antojo, lo cual, como decían Rainie y Wellman, ha generado relaciones muy superficiales, basadas simplemente en gustos o intereses de cierto tipo. Esta superficialidad y la intención de los algoritmos de tribalizarnos degenera en continuos enfrentamientos hiperreales, contra representaciones que carecen de realidad operatoria, pero que acaparan toda la atención de unos usuarios que, a continuación, se ven incapaces de lidiar con el mundo real, carente de estas tribus sobre las que basan su personalidad y comportamiento.

Esta confusión que impide a los usuarios de las redes sociales discernir entre lo real y su representación y la posterior inadaptabilidad al medio, al mundo real, es un problema fundamentalmente ontológico. Para la ontología materialista desarrollada por Gustavo Bueno en sus *Ensayos materialistas* (1972) y posteriores obras, existen tres géneros de materialidad:

- Primer género de materialidad (M1): Este nivel engloba todas las realidades que existen fuera de nuestra conciencia, así como ciertas características asociadas a los cuerpos físicos que se presentan como objetivas en nuestra percepción.
- Segundo género de materialidad (M2): Este nivel incorpora todos los procesos reales que se encuentran en una dimensión temporal más que espacial, manifestándose en el mundo como "interioridad". Correspondería a las percepciones psicológicas de cada individuo, que no tienen por qué coincidir objetivamente con los objetos percibidos.
- Tercer género de materialidad (M3): Este nivel se refiere a los objetos abstractos que no son externos ni internos, como por ejemplo: el espacio proyectivo regulado, las líneas paralelas, el conjunto infinito de números primos, el "Langue" de Saussure, las relaciones morales, las identidades sintéticas que no se encuentran en un lugar o tiempo específicos (por ejemplo, el sistema de los cinco poliedros regulares no está ni en Francia ni en Alemania, y tampoco está en la mente de los franceses o alemanes: es atópico. No dura ocho años o seis días: es acrónico). M3 también incluye entidades no esenciales pero individuales y concretas que son irrevocables, como todas las realidades que existieron en el pasado pero que ya no forman parte de M1 (por ejemplo, Julio César no es una parte del mundo físico actual) ni de M2 (Julio César es diferente de los pensamientos psicológicos acerca de Julio César).

Mientras los entornos sociales eran exclusivamente analógicos, cada individuo debía adaptarse a los medios en los que se desarrollaban las relaciones sociales, el M2 de cada individuo se adaptaba al M1 y el M3, porque de no ser así era incapaz de adaptarse y sobrevivir. Sin embargo, con la digitalización de la vida social, los usuarios pueden decidir qué ver, qué hablar o con quién agruparse, creando un entorno social a medida, subordinando el M1 y el M3 al M2. El entorno se adapta a la subjetividad de la persona, haciéndole creer que ese es el mundo, que las representaciones con las que trata diariamente son tan reales o más que el mundo real. Cuando los usuarios de las redes sociales salen al mundo real, no encuentran los grupos a los que pertenecen digitalmente, tampoco a "los otros", que daban sentido a su actividad social, incluso aquellas verdades que creían inamovibles son refutadas, y es en ese momento cuando se generan enfermedades mentales derivadas de la inadaptabilidad al medio.

Un ejemplo perfecto de esta subordinación de M1 Y M3 al M2, a la subjetividad del individuo, nos la ofrece Miguel de Cervantes en su *Don Quijote de la Mancha* (1605). Don Quijote, de tanto leer libros de caballerías, acaba creyendo todo lo que en ellos hay escrito y, decidiéndose a salir a buscar unas aventuras imposibles, las va inventando a cada ocasión. Cualquier acontecimiento que le ocurre es reinterpretado desde los códigos de los libros de caballería, pero Don Quijote no está en un entorno digital donde sus acciones son inofensivas, sino que se mueve por el mundo real y a cada paso va generando más conflictos que los que trata de resolver. Apalea a hombres de iglesia, libera a delincuentes, es apedreado... Don Quijote proyecta imágenes subjetivas, pertenecientes al M2, sobre la materia primogénica M1, y donde objetivamente hay un molino o una bacía de barbero, él ve un gigante o el yelmo de Mambrino. Esto resulta divertido en la novela, donde la acción queda encerrada entre las páginas, pero en nuestra sociedad, donde las redes sociales tienen tanta influencia en el comportamiento de la población, es peligroso que la gente construya sobre el mundo sus propias fantasías, porque estas van a chocar contra una realidad que no las permite.

De hecho, la novela también ilustra cómo las fantasías de Don Quijote hacen al protagonista manipulable. En la segunda parte (1615), Don Quijote, ya famoso por la publicación de sus aventuras, es acogido por unos duques que despliegan todo un aparato de representaciones y juegos que se adaptan a las locuras del caballero andante, que, harto de que se rían de él y su escudero, termina dejando todo un palacio para volver a la austeridad del campo y el camino. Aquí no puedo evitar citar a Cervantes por boca de Don Quijote, recién liberado de su estancia en la casa de los duques: "La libertad, Sancho, es uno de los más preciosos dones que a los hombres dieron los cielos; con ella no pueden igualarse los tesoros que encierra la tierra ni el mar encubre; la libertad, así como por la honra, se puede y debe aventurar la vida, y, por el

contrario, el cautiverio es el mayor mal que puede venir a los hombres. Digo esto, Sancho, porque bien has visto el regalo, la abundancia que en este castillo que dejamos hemos tenido; pues en mitad de aquellos banquetes sazonados y de aquellas bebidas de nieve, me parecía a mí que estaba metido entre las estrechezas de la hambre, porque no lo gozaba con la libertad que lo gozara si fueran míos; que las obligaciones de las recompensas de los beneficios y mercedes recibidas son ataduras que no dejan campear al ánimo libre. ¡Venturoso aquel a quien el cielo dio un pedazo de pan, sin que le quede obligación de agradecerlo a otro que al mismo cielo!"



Don Quijote en casa de los duques. De Enrique Recio y Gil. Fuente: <https://www.museodelprado.es/coleccion/obra-de-arte/don-quijote-en-casa-de-los-duques/fd68135d-308d-4d2d-8132-a018ac1d1321>

Igual que Don Quijote, entre favores y con todo moldeándose a sus gustos y percepciones, la sociedad en las redes sociales no goza de toda la libertad que la democracia nos brinda. Antes que guiarnos por la razón y desarrollar nuestras propias ideas sobre las que construir nuestra experiencia vital, nos subordinamos a las ideologías y los discursos emocionales para pertenecer a grupos que, en vez de destacarnos por nuestras diferencias, nos normalizan por ellas mismas. Las ideologías, al fin y al cabo, establecen una serie de ideas que sustituyen las ideas propias para incorporarnos al grupo de manera que seamos fácilmente manipulables. Igual que Don Quijote no se sentía libre cuando las ideas eran las que lo controlaban a él, y no él a sus ideas - "no lo gozaba con la libertad que lo gozara si fueran míos"-, cuando pertenecemos a estas tribus digitales o nos suscribimos a determinadas ideologías, delegamos nuestra autonomía a la acción del grupo.

## Arte y entretenimiento

A partir de la última cita nos adentramos directamente al punto final de la revisión de las implicaciones de la CD en las relaciones entre los individuos pertenecientes a una sociedad, centrándonos ahora en el arte y el entretenimiento.

Jesús G. Maestro, en su obra *Crítica de la Razón Literaria: El Materialismo Filosófico como Teoría, Crítica y Dialéctica de la Literatura* (2016), donde desarrolla un sistema de interpretación de la literatura a partir del materialismo filosófico de Gustavo Bueno, asegura que la literatura, y podemos extenderlo al arte en general, indica el grado de racionalismo de una sociedad, lo cual podemos ver a lo largo de la historia. Fue en Grecia, en el siglo VIII a.C., de la mano de Homero, su *Iliada* y *Odisea*, cuando la literatura nace, y de igual manera el resto de artes alcanzan su mayor grado de desarrollo en la Grecia de ese tiempo. A continuación, la ciencia y la filosofía comienzan su historia, dando inicio al pensamiento continental. El arte del Siglo de Oro español, a su vez, demuestra el mayor grado de racionalismo del momento, abanderado por un imperio capaz de extender su dominio por cinco continentes. De igual manera lo haría el arte francés, con artistas como Jacques-Louis David, Victor Hugo o Auguste Rodin, y el alemán, con Goethe, Beethoven o Wagner, durante el periodo posterior a la Revolución Francesa.

Teniendo el arte, entonces, como un indicador de la capacidad de razonar de una sociedad, queda por determinar cómo es el arte en una sociedad que se desenvuelve en entornos mediados por la CD, en los que se potencia y busca el impulso y la emoción que genera el consumo y se evita la reflexión y contemplación previa. Este es un aspecto que explora la obra *The Shallows: What the Internet Is Doing to Our Brain* (2010), escrita por Nicholas George Carr, donde el autor argumenta que, aunque el acceso a la información se ha vuelto más fácil, nuestra capacidad para concentrarnos, contemplar y reflexionar se está viendo perjudicada debido a la naturaleza rápida y fragmentada del consumo de información en Internet. Para Carr, Internet nos está volviendo superficiales como pensadores porque promueve la exploración rápida y pasajera de información en lugar de la concentración y la contemplación. En lugar de leer y reflexionar profundamente sobre los textos, Carr argumenta que tendemos a escanear, hacer clic en enlaces y mover rápidamente nuestra atención de un elemento de información a otro. Todo esto erosiona nuestra capacidad para el pensamiento crítico, la reflexión y la contemplación, lo cual puede tener consecuencias graves para nuestra cultura y nuestras sociedades.

Ante una sociedad de este tipo, donde el razonamiento ha quedado apartado a los espacios científicos, el entretenimiento ya no requiere de obras que muestren gran complejidad o incorporen ciertas innovaciones formales o temáticas, sino que basta con contenido que

distraiga. Este concepto, el de la "distracción", lo analiza el filósofo Damon Young en su libro *Distraction: A Philosopher's Guide to Being Free* (2008), donde define la distracción como una falta de atención y concentración en lo que realmente importa. A diferencia de la relajación o la diversión, que son actividades conscientes y deliberadas, la distracción es una actividad inconsciente que nos aleja de nuestras metas y propósitos. Las redes sociales o las plataformas de streaming, ambas potenciadas por la ciencia de datos, son poderosos agentes de distracción que nos desconectan de nosotros mismos y del mundo que nos rodea, capturan nuestra atención y nos mantienen en un estado constante de reactividad, en lugar de permitirnos estar presentes y enfocados. Esto es lo que plataformas como Netflix o TikTok ofrece a sus usuarios, un entretenimiento banal y sencillo que permita a los individuos desconectar por completo de un mundo que, por lo general, no se adapta a las exigencias de estos.

Ante este predominio de la distracción sobre el arte, los proveedores de entretenimiento han sido tajantes, y sabiendo que cualquier contenido que exija un mínimo de razonamiento en el usuario será rechazado, han sepultado el contenido de calidad con todo el contenido basura que podemos encontrar en internet. Este contenido, que tan bien funciona, es reproducido en serie, podemos verlo claramente en el caso del cine de superhéroes, género en el que el argumento cambia mínimamente de una película a otra, pero que, sin embargo, cada vez funciona mejor. Aquí, por supuesto, la CD también está implicada, ya que los productores, pertenecientes a grandes empresas de entretenimiento, están muy atentos a las opiniones vertidas por el público en las redes sociales, las cuales son analizadas a través de herramientas de procesamiento del lenguaje natural (NLP) para determinar cuáles son los aspectos que más han gustado y cuáles los que menos, para aplicar estas mejoras en siguientes entregas. En plataformas de streaming, incluso se analizan los momentos en los que los usuarios abandonan el visionado del contenido, con el objetivo de eliminar esas partes ante las que el público se muestra más reacio.

Se puede pensar que, en este aspecto, la CD es todo un avance, pues permite que la producción de entretenimiento se adapte a los gustos del público, llegando al punto de que no sea posible hacer una película que no guste, en la que los deseos de los consumidores se vean reflejados en la pantalla. Esto es lo que se conoce como "fan-service", y lejos de dar obras de calidad, lleva, en la mayoría de ocasiones, a producciones innecesarias que su única razón de existir es complacer al público como si de un producto más se tratara.

El caso es que el arte no debe tratarse como un producto, sino como un transmisor de ideas originales, aunque el gran público no sea capaz de interpretarlas. La capitalización del arte lleva a la autocomplacencia y la mediocridad, porque los artistas, limitados por las exigencias de las

grandes empresas del entretenimiento, deben resignarse a crear obras que se adapten a los gustos de un público que lo único que demanda es una distracción. ¿Qué hubiera sido del *Quijote* o del *Ulises* (1922) de Joyce en una sociedad en la que un TikTok de más de veinte segundos ya es ignorado?

Las ideas expresadas en la obra coescrita por Theodor Adorno y Max Horkheimer, *La dialéctica de la Ilustración* (1944), pueden sernos de utilidad para comprender a dónde queremos llegar. Los autores desarrollan a lo largo de la obra una crítica cultural y filosófica de la Ilustración, que lejos de llevar a la sociedad a alcanzar las promesas de liberación y racionalidad, este movimiento histórico se había convertido en una "industria cultural" que promovía la conformidad y la pasividad, y, en definitiva, en una herramienta de dominación y opresión. Argumentan que la industria cultural produce bienes culturales - películas, música, televisión, etc. - que refuerzan las ideologías y las normas del status quo. Estos productos culturales se crean y se distribuyen en masa, y a través de su consumo, los individuos son socializados en patrones de pensamiento y comportamiento que sirven a los intereses del sistema capitalista. Adorno y Horkheimer también abordan el concepto de "falsa conciencia", una idea derivada del marxismo que sugiere que los individuos, a través de la ideología dominante, a menudo no son conscientes de sus propios intereses de clase. En el contexto de la industria cultural, sugieren que los individuos pueden no reconocer cómo los productos culturales que consumen pueden trabajar para perpetuar su propia opresión.

Por último, quería comentar la teoría del genio desarrollada por Jesús G. Maestro en su ya comentada *Crítica de la razón literaria*. En su obra, Maestro desarrolla una teoría de la genialidad, que es el factor capaz de hacer avanzar la historia del arte, a través de la creación de obras de arte geniales. Para el autor, la interpretación artística entiende la genialidad como una expresión de racionalismo nunca antes vista, una versión elevada de la razón, que solicita a sus receptores contemporáneos que mejoren o superen las premisas racionales existentes o previamente establecidas. De esta manera, establece que las cualidades de la obra de arte se analizan desde dos perspectivas fundamentales: ideas y formas. En función de la originalidad de las ideas expuestas y las formas empleadas por el autor tenemos cuatro posibles situaciones en la obra: Kitsch, recurrencia, recursividad y genialidad, siendo esta última creadora de obras de arte realmente valiosas.

Ideas Técnicas	<i>Consabidas</i>	<i>Originales</i>
<i>Preexistentes</i>	KITSCH	RECURSIVIDAD
<i>Nuevas o inéditas</i>	RECURRENCIA	GENIALIDAD

El problema de la capitalización del arte, cuando este solo se centra en satisfacer los requerimientos del público, es la caída en el Kitsch, en la reproducción continua y mecanizada de las formas y temáticas que funcionan con el público. Cuando detectan que las formas empleadas generan fatiga se puede llegar a las obras recurrentes, más comunes en nuestro tiempo por el avance de las tecnologías. Cuando las temáticas empleadas generan fatiga se da libertad a autores siempre respaldados por las formas y técnicas que saben que funcionan. La genialidad, definida por la concurrencia de ideas originales y formas nuevas o inéditas, suponen un riesgo inasumible para las empresas, que, al considerar cada proyecto como un negocio, no muestran interés por este tipo de producciones.

Aquí es donde entramos en la última y crucial pieza del puzzle: el papel de la CD en este escenario. En una era en la que somos, más que nunca, consumidores de contenido, la CD se ha vuelto una herramienta esencial para las corporaciones de entretenimiento para capitalizar nuestras distracciones. No sólo se utiliza para analizar y predecir nuestras preferencias, sino también para moldear y determinar el tipo de contenido que consumimos. Esta es una característica que va más allá de la mera oferta y demanda, y que, en realidad, tiene más que ver con la manipulación de nuestras preferencias y la influencia de nuestros comportamientos. La CD, en este sentido, se convierte en una poderosa herramienta de control social y cultural.

Por un lado, la CD nos da lo que queremos, o, mejor dicho, lo que creemos que queremos. Nos da superhéroes, nos da distracciones, nos da consumibles rápidos y fáciles. Pero, ¿a qué costo? ¿A qué costo viene esta satisfacción instantánea y esta gratificación inmediata? ¿Cuál es el precio de esta perpetua complacencia y conformidad?

El costo es, por supuesto, la pérdida de la originalidad, la autenticidad, y el pensamiento crítico. Cuando el contenido se reduce a una fórmula, cuando se basa en algoritmos y datos, perdemos la capacidad de sorpresa, de exploración, de descubrimiento. Perdemos la oportunidad de ser desafiados, de ser conmovidos, de ser transformados.

Por otro lado, la CD también nos despoja de nuestra agencia, de nuestra capacidad para elegir y determinar nuestro propio consumo cultural. Se convierte en una especie de profecía autocumplida: los algoritmos nos dan lo que quieren que queramos, y entonces, lo queremos. Esta es la esencia de la "falsa conciencia" que Adorno y Horkheimer describían: no somos conscientes de cómo nuestros gustos y preferencias están siendo moldeados y manipulados por las corporaciones de entretenimiento y sus algoritmos.

En este contexto, es importante recordar que la CD no es en sí misma el problema. Al igual que cualquier otra herramienta, puede ser utilizada para el bien o para el mal, para la liberación o para la opresión. El problema radica en cómo se utiliza la CD, en qué manos cae, y para qué propósitos se utiliza.

Por tanto, si queremos resistir a la capitalización del arte, a la distracción perpetua, y a la conformidad impuesta, tenemos que ser conscientes de cómo las grandes empresas de entretenimiento están moldeando nuestras vidas y nuestra cultura. Tenemos que cuestionar y desafiar las narrativas y las estructuras de poder que la CD refuerza. Y, por supuesto, tenemos que seguir valorando y promoviendo el arte y el pensamiento crítico, la originalidad y la autenticidad, y todas aquellas cosas que nos hacen humanos y que no pueden ser reducidas a un simple conjunto de datos.

## Conclusiones

Una vez vistas las principales problemáticas surgidas en las relaciones entre los individuos de la sociedad derivadas de la CD, podemos determinar los principales aspectos que requieren una profunda revisión desde la legislación. Estos aspectos son los siguientes:

- Privacidad de datos y consentimiento informado: Las leyes actuales deben reforzarse y actualizarse para garantizar la privacidad de los datos de los usuarios y asegurar que los mismos estén plenamente conscientes y den su consentimiento para cómo se utilizan sus datos. Esto incluye no solo los datos personales explícitamente proporcionados, sino también la información recopilada por el seguimiento de sus comportamientos y preferencias en línea.
- Regulación de algoritmos de redes sociales: Las leyes deben establecer normas claras para los algoritmos de personalización de contenido en las redes sociales, para evitar la creación de "cámaras de eco" y la propagación de desinformación, y para fomentar una mayor diversidad de contenido y puntos de vista.

- Control de la desinformación: La legislación puede tratar de abordar el problema de la desinformación y las "fake news" en las redes sociales, obligando a las plataformas a tomar medidas para detectar y eliminar dicho contenido. Este es un tema delicado, pues debe equilibrarse con la libertad de expresión.
- Efectos en la salud mental: Las leyes deben reconocer y abordar los posibles efectos negativos de las redes sociales en la salud mental, incluyendo la creación de normas para limitar ciertas prácticas dañinas y promover un uso saludable de estas plataformas.
- Promoción de la responsabilidad de las empresas de redes sociales: Las leyes deben exigir a las empresas de redes sociales que asuman una mayor responsabilidad por el impacto que tienen en la sociedad, incluyendo la manera en que influyen en las elecciones de consumo, el discurso público y la salud mental de los usuarios.
- Educación y alfabetización mediática y digital: La legislación debe buscar promover una mayor educación y alfabetización mediática y digital para que las personas estén más informadas y sean más conscientes de cómo su comportamiento y sus preferencias pueden ser influenciados por las plataformas en línea y los algoritmos.

Como podrá apreciarse, la cuestión del arte no se ha incluido entre los aspectos a regular. La razón es que los Estados, por lo general, no han hecho grandes favores al arte históricamente y, lo más probable, es que el entrometimiento de la legislación en la creación artística lleve a situaciones de censura o propaganda política. La cuestión artística no es abordable desde la legislación, pero la educación digital de la sociedad y la regulación de los algoritmos sí, lo cual podría ser suficiente para convertirnos en una sociedad que aprecie y demande arte cada vez más desafiante.

## 4. Análisis crítico de la legislación en materia de datos

Llegados al último epígrafe de nuestro estudio, solo nos queda revisar la legislación en materia de datos diseñada para regular la disciplina de la que nos hemos ocupado en este estudio, la CD.

A lo largo del anterior epígrafe hemos examinado detenidamente las posibles implicaciones negativas que un mal uso de esta disciplina puede ocasionar a nivel económico, político y social, lo cual es de vital importancia a la hora de poder regular una serie de tecnologías con tantas posibilidades como esta.

Sin embargo, se verá que tan o más importante que el estudio de sus implicaciones es la primera parte del estudio, donde, mediante un estudio genealógico y ontológico, hemos sido capaces de determinar a qué nos estamos refiriendo cuando hablamos de CD. Tengamos en cuenta que de ser la CD una ciencia positiva, como la matemática o la geografía, este último epígrafe carecería de sentido, pues no se puede regular una ciencia. Sería ridícula una ley que proscribiera el uso de las raíces cuadradas por ser peligrosas para las generaciones más jóvenes, por ejemplo. En cambio, sí es lógico regular el uso de algoritmos de segmentación de contenido que generen cámaras de eco. La única regulación que cabe en las ciencias positivas es en sus operaciones, que al realizarse en medios físicos deben estar de acuerdo a la legislación vigente. Por ejemplo, si la experimentación en animales está prohibida por ley, ciertas operaciones en las que se utilicen a animales no podrán realizarse, pero en este caso la prohibición no es consecuencia de los resultados de esa operación, que pueden ser útiles para desarrollar fármacos, sino de una consideración ética materializada en ley, mientras que si se regulara el uso de ciertos algoritmos sería directamente por el resultante de su empleo, cámaras de eco, por ejemplo.

La parte de la genealogía, por su lado, nos ha brindado un conocimiento fiable de las distintas tecnologías y herramientas que conforman la CD, lo cual es evidentemente necesario para poder regularlas. ¿Si no sabemos a qué nos referimos cómo vamos a regularlo correctamente?

Así, llegamos al último epígrafe y el procedimiento que vamos a efectuar consistirá en la determinación de siete bloques temáticos, basados en las conclusiones con las que poníamos final a cada uno de los apartados del anterior epígrafe, sobre los que examinar la legislación vigente en materia de datos. La revisión la resumiremos en tres puntos: aciertos, con aquellos aspectos de los bloques que la legislación regula correctamente; errores, con los aspectos donde

la legislación falla de alguna manera, siendo incapaz de resolver los peligros a los que hacen referencia los bloques; y carencias, con aquellos aspectos que no hayan sido contemplados aún por la legislación. De esta manera, pasamos a enumerar los distintos bloques:

- BLOQUE 1. Protección, privacidad y propiedad de los datos.
- BLOQUE 2. Monopolios de información.
- BLOQUE 3. Responsabilidad y transparencia de las empresas de redes sociales.
- BLOQUE 4. Regulación de publicidad y contenido sensible.
- BLOQUE 5. Control de la desinformación.
- BLOQUE 6. Control de contenido sesgado.
- BLOQUE 7. Alfabetización digital y mediática

Y, de igual manera, pasamos a indicar el conjunto de cada una de las leyes, regulaciones, normativas y directivas que establecen un marco regulador para la actividad de la CD:

- Reglamento General de Protección de Datos (GDPR): Este reglamento de la Unión Europea fue implementado en mayo de 2018. Tiene como objetivo principal proteger a los ciudadanos de la UE en lo que respecta al procesamiento de sus datos personales y a la libre circulación de estos datos. El GDPR contiene disposiciones y requerimientos para el procesamiento de datos personales de individuos que se encuentren en la UE, e impacta a todas las organizaciones que traten con dichos datos, independientemente de su ubicación.
- Ley de Servicios Digitales (DSA): Propuesta por la Comisión Europea en diciembre de 2020, la DSA es un conjunto de normas destinadas a crear un espacio de servicios digitales seguro y transparente, reforzando las obligaciones de las empresas digitales y protegiendo los derechos fundamentales en línea. Sus disposiciones incluyen regulaciones sobre contenido ilegal, publicidad y transparencia algorítmica. En enero de 2022 fue aprobada por el Parlamento Europeo y se espera que entre en vigor para enero de 2024.
- Directiva de Derechos de Autor en el Mercado Único Digital: Esta directiva de la UE tiene como objetivo armonizar aspectos del derecho de autor en el mercado único digital. La directiva, aprobada en 2019, ha sido objeto de controversia, particularmente en lo que respecta a los artículos 15 y 17, que tratan sobre los derechos de los editores de prensa y la responsabilidad de las plataformas de compartir contenido, respectivamente.
- Ley Española de Educación (LOMLOE): Esta ley, conocida como "Ley Celaá", establece el marco legal para la educación en España. Entre sus disposiciones, incluye referencias a

la alfabetización digital y mediática, reconociendo la importancia de preparar a los estudiantes para navegar de manera segura y eficaz en un entorno digital. La revisión de esta ley corresponde al análisis del último bloque temático, pues solo desde la educación pueden fomentarse aspectos como la alfabetización digital o el pensamiento crítico frente a la desinformación.

Así, comenzamos con el análisis de cada uno de los bloques temáticos.

## 1. Protección, privacidad y propiedad de los datos

Este bloque se centra directamente en el uso que terceros pueden hacer de los datos de los individuos. Respecto a la privacidad de los datos, es importante que la legislación asigne derechos a los individuos con respecto a su información personal, que incluyan el derecho a ser informado sobre cómo se recopilan, utilizan y comparten sus datos; el derecho a acceder a sus datos y a rectificar cualquier inexactitud; el derecho a que sus datos sean borrados u "olvidados"; y el derecho a limitar o revocar el consentimiento para el uso de sus datos. Por supuesto, garantizar la protección de estos datos es fundamental a través del establecimiento de normas y procedimientos para prevenir la pérdida, el acceso no autorizado, la divulgación, la modificación o la destrucción de los datos. Por último, en cuanto a la propiedad de los datos, la ley debe asegurar la propiedad de los usuarios tanto en los datos facilitados directamente, como en los generados indirectamente, como los datos de comportamiento.

### Aciertos

Los principales aciertos que encontramos en relación a las problemáticas derivadas de este bloque son los siguientes:

- **Reglamento General de Protección de Datos (GDPR):**
  - **Consentimiento:** El Artículo 7 del GDPR establece que las organizaciones deben obtener el consentimiento libre, informado e inequívoco de los individuos para procesar sus datos. Esto garantiza que los individuos tengan control sobre cómo y cuándo se utilizan sus datos.
  - **Derecho al olvido:** El Artículo 17 del GDPR proporciona el derecho al olvido, lo que permite a los individuos solicitar que sus datos sean eliminados cuando ya no sean necesarios para los fines para los que se recopilaron.
  - **Derecho de acceso y rectificación:** Los Artículos 15 y 16 del GDPR garantizan el derecho de acceso y rectificación. Los individuos pueden solicitar una copia de sus datos y corregir cualquier error en ellos.
  - **Portabilidad de datos:** El Artículo 20 del GDPR introduce el derecho a la portabilidad de datos, permitiendo a los individuos obtener y reutilizar sus datos personales para sus propios fines en diferentes servicios.

- Directiva sobre privacidad electrónica (2002/58/EC, conocida también como Directiva de cookies):
  - **Consentimiento para cookies:** La Directiva establece que los usuarios deben dar su consentimiento antes de que las cookies puedan ser almacenadas y recuperadas en su computadora o dispositivo móvil. Esto ha aumentado la transparencia y el control sobre cómo se rastrean las actividades en línea de los usuarios.
- Propuesta de Reglamento sobre la Gobernanza de los Datos (2020/0259(COD)):
  - **Fomento del intercambio de datos:** Este reglamento propone medidas para fomentar el intercambio de datos entre organizaciones, manteniendo al mismo tiempo los derechos de protección de datos. Esto podría abrir nuevas oportunidades para la innovación y el crecimiento en la economía digital.

## Errores

A pesar de los aciertos, existen aspectos donde la legislación establece unas directrices vagas o confusas, que, en ocasiones, no se acercan a la realidad del conjunto de empresas a las que hacen referencia, y tampoco son capaces de solucionar las problemáticas a las que los usuarios se enfrentan. Algunos de estos errores son los siguientes:

- Reglamento General de Protección de Datos (GDPR):
  - **Datos recogidos indirectamente:** GDPR define los datos personales como cualquier información relacionada con una persona física identificada o identificable (artículo 4), pero la interpretación de qué constituye la identificación puede ser debatida, especialmente en el caso de datos de comportamiento recogidos indirectamente.
  - **Aplicación y cumplimiento incoherentes:** Aunque el GDPR es uniforme en papel, en la práctica, su implementación puede variar de un estado miembro a otro. Esta inconsistencia (artículo 51-55) puede causar confusión y hacer que sea difícil para las empresas garantizar el cumplimiento total.
  - **Sanciones desproporcionadas (Artículo 83):** El GDPR establece sanciones significativas para las infracciones, incluyendo multas de hasta 20 millones de euros o el 4% de la facturación anual global de la empresa, lo que sea mayor. Aunque las sanciones deben ser disuasorias, existe la preocupación de que las

multas potencialmente enormes podrían afectar desproporcionadamente a las pequeñas y medianas empresas.

- **Carga administrativa (varios artículos):** Los requisitos del GDPR, como la obligación de mantener un registro de las actividades de procesamiento de datos (Artículo 30), pueden imponer una carga administrativa considerable, especialmente para las pequeñas y medianas empresas.
- Directiva sobre privacidad electrónica (2002/58/EC, conocida también como Directiva de cookies):
  - **Experiencia de usuario deteriorada:** Si bien la intención de pedir consentimiento para las cookies es proteger la privacidad del usuario, en la práctica, esto a menudo puede llevar a los usuarios a aceptar ciegamente las cookies para deshacerse de las molestas notificaciones, lo cual está lejos del consentimiento informado que la ley pretendía (artículos 5(3)).
- Propuesta de Reglamento sobre la Gobernanza de los Datos (2020/0259(COD)):
  - **Dificultades en la gestión de los derechos de propiedad de los datos:** Aunque la propuesta de reglamento busca fomentar el intercambio de datos, puede ser difícil en la práctica equilibrar esto con los derechos de propiedad y control sobre los datos, especialmente cuando los datos se crean y recogen a través de múltiples fuentes y por múltiples actores.

## Carencias

Por último, hay que resaltar que ciertos matices tocantes al motivo del presente bloque temático quedan sin abordar por la legislación vigente. Los aspectos sobre los que es necesario un acercamiento más profundo desde la legislación son los siguientes:

- Reglamento General de Protección de Datos (GDPR):
  - **Falta de claridad en el concepto de 'consentimiento informado':** El GDPR (Artículo 4, definición 11) establece que el consentimiento del titular de los datos debe ser "informado", pero hay cierta falta de orientación clara y específica sobre qué constituye un "consentimiento informado" en todos los contextos y escenarios.
  - **Ambigüedad sobre el tratamiento de datos biométricos y genéticos:** El GDPR (Artículo 9) prohíbe el tratamiento de datos biométricos y genéticos sin

consentimiento, pero hay una falta de claridad sobre cómo deben ser tratados estos tipos de datos en contextos específicos, como la investigación médica.

- Directiva sobre privacidad electrónica (2002/58/EC, conocida también como Directiva de cookies)
  - **Falta de adaptación a las nuevas tecnologías:** La directiva se ha quedado atrás en términos de las nuevas tecnologías y los métodos de seguimiento, y puede necesitar ser actualizada para abordar más eficazmente las nuevas formas de seguimiento en línea.
- Reglamento sobre la Gobernanza de los Datos (2020/0259(COD)):
  - **Falta de consideración para los modelos emergentes de propiedad de datos:** A medida que evolucionan las formas en que los datos pueden ser poseídos y compartidos (por ejemplo, a través de cadenas de bloques y contratos inteligentes), la regulación actual puede carecer de las consideraciones necesarias para estos modelos emergentes.
- Ley de Servicios Digitales (DSA):
  - **Necesidad de mayor detalle en la regulación de las plataformas de inteligencia artificial:** Aunque la DSA establece algunas normas para las plataformas de IA, todavía puede haber carencias en términos de cómo se deben regular y supervisar estos sistemas, especialmente en lo que respecta a los sistemas de toma de decisiones automatizados y a la transparencia de los algoritmos.

En el presente bloque temático, relativo a las problemáticas surgidas de la protección, privacidad y propiedad de los datos personales, encontramos que la legislación no ha hecho un mal trabajo al consagrar los derechos de los individuos sobre sus datos, pudiendo, según el reglamento, acceder a ellos, rectificarlos, eliminarlos de las plataformas u obtener y reutilizarlos. Sin embargo, de nada sirve garantizar estos derechos si las posibilidades reales de los individuos son luego tan limitadas. La falta de claridad sobre la identificación de los datos incita a pensar que estos derechos pueden quedarse en nada, pues si no está clara la diferencia entre datos recogidos directa e indirectamente, ¿de qué sirve reclamar la propiedad de tus datos, si van a poder seguir empleándolos una vez transformados? Además, al no haber directrices claras sobre cómo facilitar que los usuarios ejerzan estos derechos, estos pueden quedar en nada en la mayoría de las ocasiones, neutralizando los aciertos que creíamos haber encontrado. Tampoco el consentimiento obligatorio es un acierto en la práctica, pues de sobra es sabido que una molesta pantalla emergente que nadie se lee no es más que un trámite que, aunque las empresas se

hayan visto obligadas a incluir, no supone un consentimiento realmente "informado", y esto sin tener en cuenta el deterioro evidente de la experiencia de usuario, habiéndose convertido algo tan importante como ceder tu información en un obstáculo que tratas de superar lo más rápido posible aceptando todo lo que te pidan.

En lo tocante a las empresas, encontramos más errores que aciertos, pues la legislación no tiene en cuenta la inmensa diversidad de organizaciones que ya hacen uso de la CD en su actividad, e impone unas sanciones que no se adaptan a las características de esa diversidad, lo veremos también en el siguiente bloque sobre los monopolios de información.

Por último, resaltar el carácter desfasado que una legislación tan importante ha adquirido en tan cortos periodos de tiempo. Nuevas tecnologías, como las inteligencias artificiales de creación de contenido o la "Blockchain", no son contempladas aún por la legislación, y en un momento donde la innovación va a tanta velocidad y las tecnologías resultantes son tan potentes, la aparición de nuevas problemáticas es una constante que la legislación debe abordar rápidamente.

## 2. Monopolios de información

Continuando con el segundo bloque, debemos en primer lugar tener en cuenta que disminuir la presencia de las big tech en los entornos digitales es una tarea imposible para cualquier legislación. El objetivo de la legislación, por tanto, no es tanto hacer la guerra a estas empresas, sino tratar de disminuir el poder de estas en favor de los usuarios y las pequeñas empresas, diversificando el mercado de la información. En el epígrafe anterior, en el apartado sobre economía, se proponía la creación de un mercado descentralizado de datos, donde los usuarios, dueños de su propia información, recibieran una compensación por el uso de sus datos. De esta manera, las big tech no gozarían de la propiedad de estos de manera gratuita solo por el hecho de que los usuarios usen sus servicios, sino que tendrían que pagar por esa información. Así, las empresas más pequeñas también podrían participar en este mercado adquiriendo una cantidad de información, que de no ser así no podrían obtener por sus propios medios. Esto solo es una idea sugerente, no buscamos algo así en la legislación, pero sí motivos para pensar que algo así es posible, que la información sea propiedad de aquellos que la generan y contra los que se emplea es de cierta obligatoriedad si desde la legislación se quieren combatir los monopolios de información.

### Aciertos

Comenzamos así a buscar aquellos puntos en los que la legislación parece acertar, sentando bases para que los usuarios sean capaces de tomar y emplear sus datos en otros servicios y regulando ciertos aspectos sobre el funcionamiento de las plataformas en línea, encontramos lo siguiente:

- Reglamento General de Protección de Datos (GDPR):
  - **Derecho a la portabilidad de los datos (Artículo 20):** El GDPR garantiza el derecho a la portabilidad de los datos, lo que permite a los individuos recibir y transferir sus datos personales entre diferentes servicios. Esto puede ayudar a limitar la concentración de poder de información al permitir la competencia entre diferentes servicios y plataformas.
- Ley de Servicios Digitales (DSA):
  - **Regulación de las grandes plataformas en línea (Artículo 26):** La DSA introduce reglas específicas para las llamadas "plataformas de puerta de entrada", que son

grandes plataformas en línea que tienen un impacto significativo en el mercado interno. Esto puede ayudar a evitar que estas plataformas abusen de su posición y se conviertan en monopolios de información.

- **Transparencia de los algoritmos (Artículo 24):** La DSA requiere que las grandes plataformas en línea proporcionen información sobre la lógica, la importancia y las consecuencias de sus algoritmos. Esto puede ayudar a prevenir prácticas desleales y a garantizar que las empresas no puedan utilizar sus algoritmos para mantener un monopolio sobre la información.
- **Reglamento de Gobernanza de los Datos (2020/0259(COD)):**
  - **Promoción de la compartición de datos (Capítulo IV):** Este reglamento propone medidas para fomentar la compartición de datos entre organizaciones y sectores, lo que puede ayudar a prevenir la concentración de datos en unas pocas grandes empresas.

## Errores

Pasando con los errores, encontramos los siguientes aspectos que merecen una revisión:

- **Reglamento General de Protección de Datos (GDPR):**
  - **Cumplimiento complejo (Artículos 2-4, 12-22):** Aunque el GDPR es una ley de privacidad esencial y bien intencionada, su cumplimiento puede ser complejo y costoso, especialmente para pequeñas y medianas empresas. Esta barrera puede, paradójicamente, favorecer a las grandes empresas y reforzar inadvertidamente los monopolios de información.
  - **Dificultades en la aplicación del derecho a la portabilidad (Artículo 20):** Aunque el derecho a la portabilidad de los datos está destinado a permitir la competencia, en la práctica puede ser difícil de implementar. No todas las empresas utilizan los mismos formatos de datos, lo que puede obstaculizar la transferencia efectiva de los datos.
- **Ley de Servicios Digitales (DSA):**
  - **Falta de medidas de control de la fusión (Artículo 12):** Aunque la DSA aborda muchas prácticas anticompetitivas, carece de medidas explícitas para controlar las fusiones y adquisiciones que podrían dar lugar a monopolios de información.

## Carencias

Terminamos con las carencias en la regulación de estos monopolios de información:

- **Reglamento General de Protección de Datos (GDPR):**
  - **Falta de regulación explícita sobre la competencia (General):** Aunque el GDPR es principalmente una ley de privacidad, no tiene disposiciones explícitas que aborden el problema de los monopolios de información. Como tal, puede carecer de las herramientas necesarias para prevenir o dismantelar tales monopolios.
  - **Inconsistencias en la aplicación (Artículo 58):** El GDPR otorga a las autoridades de protección de datos el poder de imponer multas y sanciones. Sin embargo, ha habido algunas críticas sobre las inconsistencias en su aplicación, lo que puede socavar su efectividad.
- **Ley de Servicios Digitales (DSA):**
  - **Falta de enfoque en la interoperabilidad de datos (General):** Aunque el DSA incluye regulaciones destinadas a prevenir prácticas monopólicas por parte de las plataformas "porteras", se podría criticar por no tener disposiciones más fuertes que requieran interoperabilidad de datos. Tales medidas podrían ayudar a reducir los monopolios de información al facilitar que los usuarios cambien de servicio.
  - **Supervisión inadecuada de la toma de decisiones algorítmicas (Artículos 24-25):** Si bien estos artículos requieren ciertos niveles de transparencia en cuanto a la toma de decisiones algorítmicas, pueden carecer de requisitos robustos para auditorías externas o escrutinio público de estos sistemas, que juegan un papel crítico en la configuración del paisaje informativo en estas plataformas.

De nuevo, en este segundo bloque, la legislación cae en los mismos errores que en el anterior. Establece una serie de derechos que permiten a los usuarios la portabilidad de sus datos, pero la falta de unas indicaciones claras sobre cómo las empresas deben facilitar esta posibilidad a sus usuarios provoca discordancia de formatos y formas, que hacen que estos derechos queden como ideales inalcanzables.

También, como ya señalamos en el anterior bloque, falta cierto enfoque en las pequeñas y medianas empresas, pues si se establecen unas normas para evitar que las grandes empresas

abusen de la información de los usuarios, pero el cumplimiento de estas normas por sus costes solo es asumible por esas mismas grandes empresas, realmente se está favoreciendo la creación de monopolios de información. Desde la misma legislación se imponen barreras de entrada.

Por otra parte, es cierto que la nueva Ley de Servicios Digitales (DSA) sí parece acertar al regular estas grandes plataformas, imponiendo ciertas normas de funcionamiento y pidiendo más transparencia en el uso de algoritmos, aunque también la falta de directrices claras sobre cómo se debe abordar la auditoria de estas tecnologías puede hacer que esta transparencia sea solo una fachada. Lo veremos en el siguiente bloque.

### 3. Responsabilidad y transparencia de las empresas de redes sociales

Las redes sociales, a pesar de sus muchas ventajas, son un entorno potencialmente peligroso y dañino desde su constitución, por lo que estas empresas deben ser responsables de lo que su funcionamiento genera. Este bloque, ligado al bloque anterior por referirse a las grandes empresas digitales, se centra en la necesidad de pedir responsabilidad y transparencia por lo ocurrido en estas plataformas. Con esto, será necesario que las empresas de redes sociales se muestren transparentes con el funcionamiento de sus algoritmos, a través de, quizás, auditorías de algoritmos, y responsable por las facilidades que ofrecen para determinadas prácticas. Incluir avisos de los posibles peligros en la plataforma o limitar el acceso a contenidos para determinadas edades o grupos puede ser formas de responsabilizar a las empresas de redes sociales.

#### Aciertos

El éxito de la regulación en este bloque se refleja en cómo maneja los desafíos de responsabilidad y transparencia, así como en la protección de los derechos de los usuarios y la promoción de un entorno digital equitativo y seguro. Algunos de los aciertos que hemos encontrado son los siguientes:

- Reglamento General de Protección de Datos (GDPR):
  - **Principios relativos al tratamiento de datos personales (Artículo 5):** La ley establece que los datos personales deben ser tratados de manera transparente para el interesado.
  - **Transparencia de la información y comunicación de los derechos de los interesados (Artículo 12):** El GDPR exige a las empresas proporcionar información sobre el tratamiento de los datos de manera clara y transparente.
  - **Protección de datos desde el diseño y por defecto (Artículo 25):** Este artículo establece que las empresas deben implementar medidas técnicas y organizativas adecuadas para proteger los datos personales.
- Propuesta de Ley de Servicios Digitales (DSA)
  - **Transparencia en la moderación de contenidos (Artículos 14 y 15):** Estos artículos obligan a las plataformas a informar a los usuarios sobre cualquier

moderación de contenidos, proporcionando transparencia en la forma en que se aplican las políticas de las plataformas.

- **Sistema de gestión de riesgos (Artículo 24):** Este artículo exige a las plataformas muy grandes adoptar un sistema de gestión de riesgos para tratar los posibles riesgos sistémicos.
- **Auditoría independiente (Artículo 26):** Este artículo establece la obligación de las plataformas muy grandes de someterse a auditorías independientes.
- **Informes de transparencia (Artículo 30):** Este artículo exige a las plataformas proporcionar informes de transparencia en relación con el contenido ilegal y las medidas adoptadas para su gestión.

## Errores

A pesar de los aciertos notables en términos de responsabilidad y transparencia de las empresas de redes sociales, también hay algunas áreas en las que la legislación europea podría mejorar. A continuación, analizamos algunos de los errores más relevantes:

- Reglamento General de Protección de Datos (GDPR):
  - **Automatización de decisiones individuales, incluida la elaboración de perfiles (Artículo 22):** Este artículo ofrece protección contra la toma de decisiones basada únicamente en el procesamiento automatizado, incluido el perfilado. Sin embargo, su aplicación ha sido inconsistente y su interpretación varía, lo que ha llevado a una falta de claridad para las empresas.
  - **Condiciones para el consentimiento (Artículo 7):** Aunque el GDPR establece reglas estrictas para obtener el consentimiento, en la práctica, a menudo se producen situaciones en las que los usuarios dan su consentimiento sin entender completamente las implicaciones.
- Propuesta de Ley de Servicios Digitales (DSA):
  - **Sistema de gestión de riesgos (Artículo 24):** Aunque este artículo exige a las plataformas muy grandes adoptar un sistema de gestión de riesgos, no proporciona directrices claras sobre cómo se debe implementar este sistema.
  - **Informes de transparencia (Artículo 30):** Aunque esta disposición exige a las plataformas proporcionar informes de transparencia, no hay suficientes detalles

sobre qué información debe incluirse en estos informes y cómo se debe presentar.

- **Falta de sanciones efectivas (Artículo 41):** El artículo 41 de la DSA establece que los Estados miembros deben establecer sanciones por incumplimientos, pero no especifica la naturaleza o el alcance de estas sanciones. Esto puede llevar a sanciones inconsistentes entre los diferentes Estados miembros, lo que podría permitir a algunas empresas escapar de sanciones significativas.

## Carencias

A pesar de los aciertos y la dirección generalmente positiva de la legislación europea, hay varias áreas en las que existen carencias significativas en cuanto a la regulación de la responsabilidad y la transparencia de las empresas de redes sociales. Algunas de estas carencias son:

- **Reglamento General de Protección de Datos (GDPR):**
  - **Algoritmos e Inteligencia Artificial (IA):** Aunque el Artículo 22 establece límites en la toma de decisiones automatizada y perfilado, el GDPR no profundiza lo suficiente en la regulación de los algoritmos de IA y cómo estos pueden impactar en la privacidad y los derechos de los individuos. Esta es una creciente preocupación en la era de las redes sociales y la gran cantidad de datos personales que estas plataformas procesan.
  - **Falta de protección contra la recopilación de datos no consensuada (Artículos generales):** Aunque el GDPR pone mucho énfasis en el consentimiento explícito para la recopilación de datos, no hay suficientes salvaguardias contra la recopilación de datos no consensuada, especialmente en el caso de las empresas de redes sociales que a menudo recopilan grandes cantidades de datos sin el conocimiento explícito del usuario.
- **Propuesta de Ley de Servicios Digitales (DSA):**
  - **Falta de protección contra el sesgo algorítmico (Artículos generales):** Aunque la DSA requiere transparencia en torno a los algoritmos utilizados por las grandes plataformas en línea, no hay suficiente protección contra el sesgo algorítmico, que puede tener efectos significativos en la forma en que se presenta y distribuye la información en las redes sociales.

- **Manipulación de la Opinión Pública:** La DSA no incluye disposiciones específicas sobre la manipulación de la opinión pública a través de las redes sociales. Esta es una preocupación importante, dada la influencia que pueden tener estas plataformas en la formación de la opinión pública.

Los aciertos de la legislación en este bloque son más evidentes, la promoción de la transparencia de los algoritmos, el énfasis en el consentimiento del usuario para la recopilación y procesamiento de datos, y las regulaciones específicas para las grandes plataformas en línea, son pasos sustanciales hacia la protección de los derechos de los usuarios y el mantenimiento de la salud democrática de nuestra sociedad digital. De hecho, acertar en este punto es de suma importancia también para el resto de bloques temáticos, porque será cuando la responsabilidad de lo que ocurre en estas plataformas caiga sobre las empresas, cuando todos los demás aspectos recibirán una mayor atención.

Sin embargo, quedan pasos por dar. La falta de protección contra la recopilación no consensuada o la tibieza a la hora de regular los algoritmos de IA son aspectos donde la legislación debe ser más firme y exhaustiva. Las problemáticas que pueden derivarse de estos dos aspectos donde la regulación no aprieta son peligrosas para los usuarios y muy lucrativas para las empresas.

Por otro lado, las sanciones. No coordinar unas sanciones entre los Estados miembros de la Unión Europea no solo provoca una discordancia en la ejecución de la ley, sino que también provoca demoras al estudiar cada caso. Unas directrices fijas y tajantes aplicables por todos los Estados harían más efectivo el cumplimiento de las sanciones.

## 4. Regulación de publicidad y contenido sensible

Este bloque se ocupa del uso de cierta información sensible para la recomendación de publicidad y contenido en redes sociales. Debe haber una privacidad mínima que incluya aspectos sensibles, como pueden ser historiales médicos, preferencias políticas, asuntos económicos. Es peligroso, por ejemplo, que a personas con problemas de alcoholismo continuamente se les bombardee con publicidad de bebidas alcohólicas o discotecas. La regulación de este tipo de contenidos también apunta a combatir las facilidades que las plataformas de redes sociales brindan para la manipulación de los usuarios. Sin embargo, hay que tener en cuenta que este tipo de contenido, a pesar de ser sensible o peligroso, no puede ser excluido por completo, porque atentaría contra otros derechos de empresas que se dediquen a la venta de alcohol, por ejemplo, por lo que una solución podría ser identificar ese contenido como sensible siempre que aparezca.

### Aciertos

Algunos aciertos de la legislación europea respecto a la regulación de la publicidad y el contenido sensible son los siguientes:

- **Reglamento General de Protección de Datos (GDPR):**
  - **Consentimiento explícito para el procesamiento de datos (Artículos 6 y 7):** El GDPR exige que los usuarios den su consentimiento explícito para el procesamiento de sus datos, lo que también incluye el uso de dichos datos para fines publicitarios. Esto garantiza que los usuarios tengan control sobre cómo y cuándo sus datos son utilizados para publicidad.
  - **Derecho a oponerse al procesamiento de datos para fines de marketing (Artículo 21):** Este artículo permite a los usuarios oponerse al procesamiento de sus datos para fines de marketing directo, incluyendo la creación de perfiles.
- **Ley de Servicios Digitales (DSA):**
  - **Transparencia de la publicidad y algoritmos (Artículos 24 y 30):** Los "porteros" tienen que explicar los parámetros principales que determinan la clasificación, junto con cualquier opción para influir en ellos. Los porteros también deben proporcionar información sobre la existencia de publicidad pagada y sobre la persona en cuyo nombre se entrega la publicidad.

- **Protección contra contenido ilegal y nocivo (Artículos 14-20):** Estos artículos requieren que las empresas de plataformas tomen medidas contra el contenido ilegal y proporcionen a los usuarios mecanismos para denunciarlo.

## Errores

Vistos los aciertos, pasamos a ver los principales errores que nosotros encontramos en la legislación, y donde se puede afinar para disminuir los problemas derivados de los puntos que dan nombre a este bloque:

- **Reglamento General de Protección de Datos (GDPR):**
  - **Definición y clasificación poco clara de los datos sensibles (Artículo 9):** El GDPR prohíbe el procesamiento de "datos sensibles" sin el consentimiento explícito del individuo. Sin embargo, puede haber ambigüedad en la interpretación de qué constituye datos sensibles en el contexto de la publicidad y el contenido.
  - **Falta de detalles sobre la implementación del consentimiento (Artículos 6 y 7):** Aunque el GDPR requiere el consentimiento explícito para el procesamiento de datos, no proporciona directrices claras sobre cómo se debe obtener y verificar este consentimiento, lo que puede dar lugar a abusos.
- **Ley de Servicios Digitales (DSA):**
  - **Falta de claridad sobre la supervisión de contenido (Artículos 14-20):** Si bien la DSA exige que las empresas de plataformas tomen medidas contra el contenido ilegal, no proporciona pautas claras sobre cómo estas empresas deben supervisar y moderar el contenido sin violar la privacidad de los usuarios.
  - **Falta de control sobre la influencia de los algoritmos (Artículo 24):** Aunque la DSA exige transparencia sobre el funcionamiento de los algoritmos que determinan la clasificación de la publicidad, no establece medidas de control sobre cómo estos algoritmos pueden ser manipulados para influir en el comportamiento de los usuarios.

## Carencias

Por último, revisamos aquellos aspectos donde aun queda trabajo por hacer:

- Reglamento General de Protección de Datos (GDPR):
  - **Enfoque limitado en la publicidad dirigida (Artículos 6 y 7):** Aunque el GDPR habla del consentimiento para el procesamiento de datos, no aborda explícitamente los problemas específicos relacionados con la publicidad dirigida y cómo los datos pueden ser utilizados (y potencialmente abusados) en este contexto.
  - **Falta de medidas específicas para proteger a los niños (Artículo 8):** Aunque el GDPR establece que el consentimiento debe obtenerse de un adulto cuando se procesan los datos de los niños, no proporciona orientación o medidas adicionales para proteger a los niños de la publicidad y el contenido sensibles.
- Ley de Servicios Digitales (DSA):
  - **Falta de regulaciones claras sobre la transparencia de la publicidad (Artículos 24-30):** Aunque la DSA exige cierto nivel de transparencia en la publicidad, no proporciona una orientación suficientemente detallada sobre cómo debe lograrse esta transparencia.
  - **No aborda suficientemente la desinformación y las noticias falsas:** La DSA proporciona medidas para eliminar el contenido ilegal, pero no aborda completamente el problema de la desinformación y las noticias falsas, que pueden ser legales pero dañinas.

El propósito de este cuarto bloque y de los dos siguientes son de los que más interesan para este estudio, como se ha podido apreciar en la segunda parte de este. En este bloque, podemos reconocer grandes aciertos de la legislación al salvaguardar el derecho de los usuarios a oponerse al procesamiento de datos con fines de marketing y asegurarse de que tengan mecanismos para denunciar el contenido nocivo o ilegal.

Pero, de nuevo, la falta de claridad puede llevar a situaciones no deseadas en estas plataformas. Hablar de contenido sensible sin definirlo lleva a la interpretación propia de las empresas y, por tanto, es como si no reguláramos nada. De hecho, esta definición tampoco parece demasiado conflictiva: es contenido sensible el que presenta violencia, contenido sexual explícito, lenguaje ofensivo, manifestaciones de odio o discriminación... También la publicidad que aprovecha características estrictamente personales como la orientación sexual o las ideas políticas pueden considerarse sensible. No queremos decir que todo este contenido deba suprimirse por ley, con un mecanismo para que cada individuo pueda denunciarlo si quiere dejar de verlo es suficiente,

pero también es importante que todo esto esté clasificado como sensible previamente, y para eso tendría que haber sido definido por la legislación.

También es importante comentar la poca atención que se presta a las generaciones más jóvenes en el GDPR, que son las más vulnerables y pueden tener acceso a todo lo que un adulto ve. Es evidente que son los tutores los que tienen que preocuparse de lo que hacen los niños en internet, pero hay que pedir responsabilidad a estas empresas que saben que hay niños y adolescentes presentes en sus plataformas.

## 5. Control de la desinformación

De nuevo, tratamos en este bloque de contenido sensible, pero esta vez no por aprovechar información de riesgo, sino por ser deliberadamente falso. El control de la desinformación es una labor imprescindible al regular la ciencia de datos, porque la capacidad de manipulación de este tipo de contenidos excede en riesgo a cualquier otro aspecto. Hablamos en este punto del concepto de posverdad, que tanto juego ha dado en los últimos años, que es empleado, más inocentemente para generar visitas y debate en cosas intrascendentes, y más seriamente para la manipulación y el control de la opinión pública de la mano de los poderes políticos y económicos. Aquí, de nuevo, hacemos una llamada a la responsabilidad de las empresas de redes sociales para identificar este tipo de contenido y eliminarlo de la plataforma.

### Aciertos

Comenzando con los aciertos en la regulación del contenido que promueve la desinformación en los usuarios encontramos:

- Ley de Servicios Digitales (DSA):
  - **Mayor transparencia y responsabilidad para las plataformas (Artículos 5 y 6):** La DSA exige a las grandes plataformas que sean transparentes acerca de las medidas que toman para mitigar la propagación de contenido ilegal, incluyendo desinformación. Esto es un acierto en cuanto a la exigencia de responsabilidad y transparencia a las plataformas.
  - **Cooperación con las autoridades nacionales (Artículo 7):** Las plataformas deben designar un representante legal en cada estado miembro en el que operen para garantizar la cooperación con las autoridades nacionales. Esto podría facilitar las investigaciones sobre casos de desinformación y aumentar la responsabilidad de las plataformas.
  - **Enfoque en la moderación de contenido (Artículos 14 y 15):** La DSA establece que las plataformas de contenido deben tener un sistema en lugar para la moderación de contenido, lo que puede ayudar a combatir la desinformación.
  - **Procedimientos de gestión de riesgos (Artículos 26 y 27):** Las plataformas están obligadas a realizar una evaluación anual de los riesgos de sus sistemas y a tomar medidas para mitigar cualquier problema identificado, incluyendo la

desinformación. Esto es crucial para la prevención y control de la propagación de información falsa.

- **Exigencia de auditorías independientes (Artículo 28):** Las grandes plataformas deben someterse a auditorías independientes de sus sistemas de moderación de contenido y medidas contra la desinformación. Esto añade una capa adicional de responsabilidad y supervisión.
- **Código de Prácticas contra la Desinformación:** Aunque no es una ley per se, este código voluntario de la Unión Europea al que muchas de las principales plataformas se han adherido representa un importante paso adelante en el control de la desinformación. Incluye compromisos para mejorar la transparencia de la publicidad política, cerrar cuentas falsas y utilizar verificadores de hechos independientes.

## Errores

Pasando a los errores, vemos los siguientes:

- **Ley de Servicios Digitales (DSA):**
  - **Evaluación de riesgos (Artículos 26 y 27):** Aunque se exige a las plataformas que evalúen los riesgos asociados a su uso y que adopten medidas para mitigarlos, la DSA no establece criterios claros ni suficientemente detallados sobre cómo debería realizarse esta evaluación y qué tipo de medidas de mitigación deberían adoptarse. Esto puede dar lugar a interpretaciones demasiado amplias o laxas por parte de las plataformas.
  - **Transparencia (Artículos 24, 30):** Aunque se hace hincapié en la necesidad de transparencia, se proporciona poca orientación sobre cómo las plataformas deben divulgar información acerca de sus políticas, operaciones y algoritmos. Esto puede dificultar la capacidad de los reguladores y del público para entender cómo funcionan exactamente estas plataformas y cómo manejan la desinformación.
  - **Auditorías externas (Artículo 28):** Aunque se requiere que las plataformas muy grandes realicen auditorías externas, la DSA no especifica quién debería realizar estas auditorías o cómo deberían realizarse. Esto podría dar lugar a situaciones en las que las auditorías no sean imparciales o exhaustivas.
- **Código de Prácticas contra la Desinformación:**

- **Carácter voluntario:** Aunque muchas plataformas se han adherido al Código de Prácticas contra la Desinformación de la UE, su naturaleza voluntaria significa que su implementación puede variar considerablemente entre diferentes plataformas, y no hay sanciones formales para aquellas plataformas que no lo sigan adecuadamente.
- **Evaluación de la efectividad:** El Código de Prácticas carece de un sistema robusto para evaluar su efectividad en la lucha contra la desinformación. Sin tal sistema, es difícil medir si el Código está logrando sus objetivos.
- **Falta de detalles sobre implementación:** El Código de Prácticas da amplias directrices, pero carece de detalles sobre cómo las plataformas deben implementar estas directrices en la práctica. Esto podría permitir a las plataformas interpretar las directrices de manera que no sean efectivas para combatir la desinformación.

## Carencias

En este bloque, las carencias son más significativas como podemos ver:

- **Definición clara de desinformación:** Ni el Reglamento General de Protección de Datos (RGPD), la Ley de Servicios Digitales (DSA) ni el Código de Prácticas contra la Desinformación definen claramente qué constituye desinformación. Esta falta de claridad puede dificultar la identificación y moderación efectiva de la desinformación.
- **Voluntariedad del Código de Prácticas contra la Desinformación:** Este código es voluntario, lo que significa que las empresas de redes sociales pueden optar por no participar o cumplir solo parcialmente con sus directrices. Una carencia es la falta de un mecanismo obligatorio para garantizar el cumplimiento.
- **Algoritmos de recomendación:** Los algoritmos que impulsan las recomendaciones de contenido en las plataformas de redes sociales pueden amplificar la desinformación. La regulación actual no aborda de manera efectiva este aspecto del problema.
- **Identificación y seguimiento de la desinformación en tiempo real:** Las plataformas digitales tienen la capacidad de identificar y rastrear la propagación de la desinformación en tiempo real. Sin embargo, las políticas actuales carecen de requisitos claros para la detección y el seguimiento proactivos de la desinformación.

En este bloque, los comentarios deben comenzar con las carencias. Que la legislación no incluya una definición clara sobre un fenómeno sobre el que han corrido ríos de tinta como es la desinformación y el concepto de posverdad es algo inadmisibile. A partir de esto, cualquier regulación sobre este tipo de contenido es en vano. La definición del objeto de estudio es la primera parte fundamental de cualquier análisis, como hemos visto en este estudio sobre la CD.

Es cierto, que aspectos que sí contempla la legislación, como la cooperación con las autoridades nacionales, es importante para hacer frente a este fenómeno, pero el problema es que si no tenemos claro a que nos referimos no podemos luchar contra ello.

Por otro lado, el hecho de que el Código de Prácticas contra la Desinformación sea de carácter voluntario implica que cualquier empresa que quiera saltárselo puede hacerlo siempre que quiera. De hecho, supone más peligro que seguridad, porque puede dar la impresión de que una red social lucha contra la desinformación, pero luego no es así, y solo elimina el contenido que considera como tal, tachándolo quizás sin criterio, y promueve otro que sí lo es y favorece sus propios intereses.

Todo el estudio que hemos realizado hasta aquí nos lleva a afirmar que lo contenido en este bloque temático materializada las implicaciones más peligrosas de la CD y, por tanto, su regulación debe realizarse sin demora.

## 6. Control de contenido sesgado

El contenido del bloque seis se centra en la regulación del funcionamiento de los algoritmos de redes sociales para evitar que el contenido sesgado pueda influir en las opiniones de los usuarios. Es de suma importancia que el contenido que los usuarios visualizan no está sesgado para evitar las cámaras de eco de las que hemos hablado. La pluralidad de las opiniones debe ser un derecho fundamental para cualquier usuario de redes sociales, que debe primar sobre la maximización insaciable de las interacciones. Si un usuario está interesado en ciertos temas, las plataformas deben actuar responsablemente, dejando a un lado los objetivos puramente económicos, y brindando todos los puntos de vista existentes sobre los que los usuarios puedan construir su visión de conjunto. Este también es un bloque clave para evitar la manipulación y el control de la sociedad, y para la disolución de ideologías irracionales y tribus digitales.

### Aciertos

Los aciertos que encontramos en relación a este bloque son los siguientes:

- **Reglamento General de Protección de Datos (GDPR)**
  - **Transparencia y consentimiento (Artículos 5, 6 y 7):** El GDPR proporciona un marco fuerte para obtener el consentimiento de los usuarios antes de procesar sus datos. Esto puede actuar como una salvaguarda contra el sesgo, ya que los usuarios tienen la opción de no participar en sistemas que consideren parciales.
  - **Derecho a la explicación (Artículo 15):** Este derecho puede ser útil para combatir el sesgo, ya que permite a los usuarios entender cómo se toman las decisiones automatizadas que les afectan.
  - **Protección contra decisiones automatizadas (Artículo 22):** Los usuarios tienen el derecho de no estar sujetos a decisiones basadas únicamente en el procesamiento automatizado, incluyendo la elaboración de perfiles. Este artículo es especialmente relevante en el contexto del control de contenido sesgado.
- **Ley de Servicios Digitales (DSA):**
  - **Implementación de sistemas de quejas (Artículo 15):** El Artículo 15 de la DSA requiere que las plataformas en línea establezcan un sistema de gestión de quejas interno. Esto puede permitir a los usuarios expresar sus preocupaciones

sobre contenido sesgado y puede proporcionar a las plataformas una retroalimentación valiosa sobre la forma en que sus sistemas de recomendación y moderación de contenido están funcionando.

- **Transparencia de los sistemas de recomendación (Artículo 29):** Los usuarios tienen derecho a recibir información sobre el funcionamiento de los sistemas de recomendación y a modificar sus parámetros. Esto puede ayudar a prevenir el sesgo al permitir a los usuarios tener más control sobre el contenido que ven.
- **Transparencia en la publicidad (Artículo 30):** Este artículo exige a las plataformas que sean transparentes acerca de por qué se les ha mostrado una determinada publicidad. Esto puede ayudar a combatir el sesgo al permitir a los usuarios comprender mejor la selección y orientación de los anuncios.

## Errores

Los aspectos donde la regulación no termina de atacar el problema correctamente son los siguientes:

- Reglamento General de Protección de Datos (GDPR):
  - **Ambigüedad en la interpretación de sesgo (Artículos 5 y 22):** Aunque el GDPR intenta regular el procesamiento de datos y garantizar la equidad, existe cierta ambigüedad en la interpretación de lo que constituye un 'sesgo'. Esto podría resultar en una aplicación inconsistente de estas disposiciones y permitir la existencia de prácticas sesgadas.
  - **Falta de claridad en el "Derecho a la explicación":** El Artículo 15 del GDPR, que se refiere al "derecho a la explicación", ha sido objeto de debate. Algunos expertos argumentan que no proporciona un derecho claro y efectivo a una explicación de las decisiones automatizadas. Esta falta de claridad puede ser un obstáculo para combatir el sesgo en el contenido.
- Propuesta de Ley de Servicios Digitales (DSA):
  - **Falta de normas concretas sobre la transparencia de los algoritmos:** Aunque la DSA exige una mayor transparencia en la moderación de contenidos y los sistemas de recomendación (Artículos 24 y 29), estas disposiciones podrían ser más específicas. Por ejemplo, las empresas podrían estar obligadas a realizar y publicar auditorías de sus algoritmos para detectar y corregir cualquier sesgo.

- **Falta de mecanismos de supervisión efectivos:** La DSA establece varias obligaciones para las plataformas en términos de transparencia y responsabilidad. Sin embargo, aún se podría mejorar en términos de establecer mecanismos de supervisión efectivos para garantizar que las plataformas cumplan con estas obligaciones.
- **Falta de regulación en el diseño de algoritmos (Artículo 30):** La DSA no se mete en detalles sobre cómo deberían diseñarse los algoritmos para prevenir el sesgo. Más directrices o regulaciones en esta área podrían ayudar a prevenir el sesgo en el contenido.
- **Falta de sanciones específicas para el sesgo en el contenido:** La DSA no proporciona sanciones específicas para las empresas que permiten el sesgo en su contenido. La introducción de tales sanciones podría incentivar a las empresas a hacer más para prevenir el sesgo.

## Carencias

Y, por último, algunas carencias que la legislación presenta a la hora de regular el contenido sesgado pueden resumirse con los siguientes puntos:

- Reglamento General de Protección de Datos (GDPR):
  - **Falta de claridad sobre el sesgo de los algoritmos:** El GDPR no menciona específicamente los algoritmos o cómo pueden contribuir al sesgo en el contenido que ven los usuarios. Aunque la interpretación general del GDPR puede aplicarse a los algoritmos, una orientación más específica sobre cómo se deben diseñar y usar los algoritmos para evitar el sesgo podría ser útil.
  - **Falta de herramientas prácticas para el control del usuario sobre sus datos:** Aunque el GDPR proporciona una serie de derechos a los usuarios en cuanto a sus datos, puede haber una falta de herramientas prácticas que permitan a los usuarios ejercer efectivamente estos derechos en términos de controlar cómo sus datos pueden ser utilizados para personalizar y sesgar el contenido que ven.
- Propuesta de Ley de Servicios Digitales (DSA):
  - **Falta de orientación detallada sobre la auditoría de algoritmos (Artículo 28):** Si bien la DSA requiere que ciertas plataformas muy grandes realicen auditorías de sus sistemas de clasificación y recomendación, no proporciona orientación

detallada sobre cómo deberían realizarse estas auditorías y cómo se deberían abordar los posibles sesgos.

- **Falta de medidas de transparencia para los usuarios:** Si bien la DSA exige cierta transparencia de las empresas, puede que no vaya lo suficientemente lejos en términos de exigir que las empresas expliquen cómo sus algoritmos funcionan a los usuarios. Esto puede dejar a los usuarios en la oscuridad sobre cómo y por qué ven cierto contenido.

Los contenidos que dan forma a este sexto bloque temático corresponden, como indicamos en el anterior bloque, a uno de los fenómenos que mayor preocupación ha generado a lo largo de este estudio. La creación de cámaras de eco es un problema que se extiende incluso a la vida real, siendo uno de los instrumentos de manipulación pública más sofisticados. Por ello, una legislación férrea sobre los contenidos de este bloque es muy necesaria, pero vemos que el trabajo realizado y las directrices desplegadas no terminan de penetrar en el problema.

En primer lugar, podemos ver cómo entre los aciertos no hemos sido capaces de encontrar mención al sesgo. En lugar de esto, tenemos sistemas de quejas o derechos a la explicación, que, si bien son necesarios, son solo trámites que se obliga a las empresas a pasar para lavar su imagen, porque por mucho que te expliquen o nos quejemos los algoritmos que hay detrás de la recomendación de contenido siguen funcionando de la misma manera. La solución tiene que ir por atacar directamente la constitución y el funcionamiento de estos algoritmos.

Luego, resaltamos el artículo 22 del GDPR, sobre la protección contra las decisiones automatizadas, pero de nuevo caemos en la ambigüedad con este artículo. La indefinición clara de lo que es el sesgo nos deja indefensos ante estas prácticas.

De nuevo, la legislación en este bloque se muestra tibia en la ejecución, estableciendo una serie de derechos que no garantizan realmente nada y exigiendo una transparencia que no es capaz de garantizar por completo. Aunque los esfuerzos para establecer un marco regulatorio son notables y necesarios, existen brechas significativas que dificultan su efectividad real.

## 7. Alfabetización digital y mediática

Este último bloque centra su atención en la labor desde los aparatos educativos de proporcionar a los estudiantes un programa que los prepare y enseñe cómo desenvolverse en un medio digital mediado por unas tecnologías que potencian la adicción, la polarización y la manipulación. La labor de la educación debe ser, siempre, preparar a las nuevas generaciones para enfrentarse a un entorno hostil y competitivo, donde somos nosotros los que debemos adaptarnos. En la era del capitalismo digital, las redes sociales e internet en su conjunto, son partes imprescindibles del entorno y, por tanto, aprender a desenvolverse en ellos es tremendamente necesario.

Como ya hemos señalado, para este último bloque la legislación a revisar será la ley educativa española. La razón para la elección de esta ley radica en la pertenencia al marco regulatorio de este país y a la ausencia de una ley educativa comunitaria en la Unión Europea (por suerte). Sobra decir que desde los otros reglamentos estudiados en este análisis la alfabetización digital a la que nos referimos no es un tema contemplado, el GDPR o la DSA buscan mediar los mercados digitales para evitar abusos, pero la forma en que las personas que participan en esos mercados interpretan las informaciones no es de su competencia.

De esta manera, pasamos a ver cómo desde los aparatos educativos de nuestro país se hace frente a los desafíos que la era digital nos presenta.

### Aciertos

En primer lugar, nos centramos en los aparentes aciertos de una ley educativa que parece tener intención de formar tanto a profesorado como a alumnado en materia digital. Algunos de los aciertos más notables que podemos encontrar son los siguientes:

- **Integración de las competencias digitales en el currículo (LOMLOE, Artículo 1, Apartado 1, Punto m):** Este artículo establece que uno de los principios y objetivos del sistema educativo español es el desarrollo de la competencia digital y el pensamiento computacional en todos los estudiantes. Esto muestra un claro compromiso por parte del legislador de preparar a los estudiantes para un entorno digital.
- **Promoción de la formación del profesorado en competencias digitales (LOMLOE, Artículo 102):** Este artículo destaca la importancia de que el profesorado tenga la

formación adecuada en competencias digitales, para que puedan impartir una educación digital de calidad a los estudiantes.

- **Inclusión de la educación en medios en el currículo (LOMLOE, Artículo 1, Apartado 1, Punto e):** Este punto establece que uno de los objetivos del sistema educativo es la promoción de la competencia mediática, incluyendo el análisis crítico de los medios de comunicación.
- **Promoción de la igualdad de acceso a la tecnología (LOMLOE, Artículo 84, Apartado 3):** Este artículo establece la necesidad de garantizar que todos los estudiantes tengan acceso a las tecnologías de la información y la comunicación, independientemente de su situación socioeconómica.
- **Promoción de la educación en seguridad y privacidad digital (LOMLOE, Artículo 1, Apartado 1, Punto l):** Este artículo establece que uno de los objetivos del sistema educativo es educar a los estudiantes en el uso seguro y responsable de la información y las comunicaciones digitales, incluyendo la protección de la privacidad y los datos personales.

## Errores

Revisados los aciertos, pasamos a ver una falta evidente de detalles que generan ciertas dificultades en el cumplimiento de los programas:

- **Falta de detalle en los contenidos específicos (LOMLOE, Artículo 1, Apartado 1, Punto m, e, l):** Aunque la ley menciona la importancia de la competencia digital, el pensamiento computacional, la competencia mediática, y la seguridad y privacidad digital, no especifica qué contenidos concretos se deben enseñar o cómo se deben enseñar.
- **Falta de un plan de implementación detallado (LOMLOE, Artículo 102):** A pesar de que se hace énfasis en la formación del profesorado en competencias digitales, no se ofrece un plan de implementación detallado que permita lograr este objetivo. Esto puede hacer que la formación sea inconsistente o insuficiente.
- **Falta de evaluación de las competencias digitales y mediáticas (LOMLOE, Artículos Varios):** Aunque la ley promueve la alfabetización digital y mediática, no se establece un método claro de evaluación de estas competencias. Esto puede hacer que sea difícil medir el progreso y la eficacia de la educación digital y mediática.

- **Falta de mención a las Fake News (LOMLOE, Artículo 1, Apartado 1, Punto e):** A pesar de que la competencia mediática es mencionada, no se hace una mención explícita a la problemática actual de las Fake News y cómo educar a los estudiantes para que puedan detectarlas y combatirlas.
- **Ausencia de inclusión de la ética digital (LOMLOE, Artículo 1, Apartado 1):** En un mundo cada vez más mediado por la tecnología, es importante también educar en la ética digital. No se hace mención a cómo enseñar a los estudiantes a comportarse de forma ética en el entorno digital.
- **Poca atención a la formación en algoritmos y ciencia de datos (LOMLOE, Artículos Varios):** A medida que la sociedad se convierte en una sociedad de datos, es importante que los estudiantes comprendan cómo funcionan los algoritmos y la ciencia de datos, así como su impacto en nuestras vidas. La LOMLOE no aborda explícitamente esta necesidad.
- **Ausencia de una aproximación multidisciplinar a la alfabetización digital (LOMLOE, Artículos Varios):** La alfabetización digital no debe ser solo una materia aislada, sino que debe integrarse en todas las áreas del currículo. No se hace suficiente énfasis en esta necesidad en la LOMLOE.
- **Ausencia de medidas para reducir la brecha digital (LOMLOE, Artículos Varios):** La brecha digital puede ser un gran obstáculo para la alfabetización digital y mediática. La LOMLOE no aborda suficientemente cómo reducir esta brecha.

## Carencias

Por último, revisamos las carencias más significativas:

- **Inclusión de la enseñanza de inteligencia artificial y machine learning (LOMLOE, Artículos Varios):** En un mundo donde la IA y el machine learning están teniendo cada vez más importancia, no se incluye explícitamente la enseñanza de estos temas en el currículo.
- **Enseñanza de habilidades de programación y codificación (LOMLOE, Artículos Varios):** Aunque se menciona la competencia digital, no se menciona específicamente la enseñanza de habilidades de programación y codificación, que son cada vez más importantes en el mundo laboral.

- **Formación específica para profesores en ciberseguridad (LOMLOE, Artículo 102):** Aunque se menciona la formación del profesorado en competencias digitales, no se hace mención a la ciberseguridad, que es una parte vital de la competencia digital.
- **Fomento del pensamiento crítico frente a la desinformación (LOMLOE, Artículos Varios):** En el contexto de la desinformación y las fake news, es esencial educar a los estudiantes en el pensamiento crítico para que puedan discernir la información verídica de la falsa. La LOMLOE no aborda explícitamente esta necesidad.

A simple vista puede verse la insuficiencia de una ley educativa que se limita a nombrar una serie de conocimientos sobre los que luego no profundiza ni en programas ni en temarios. Se citan una serie de objetivos como la formación del profesorado en competencias digitales o la preparación mediática de los alumnos, pero no se establecen directrices claras sobre cómo llevarlo a la práctica, por lo que a efecto tangible queda en un juego retórico. Da la impresión de que los responsables saben que estos contenidos deben enseñarse en los centros educativos, por su evidente importancia, pero que carecen de medios e ideas sobre cómo hacerlo. La ausencia de contenido específico, de sistemas de implementación o de formas de evaluar la consecución de estas competencias nos dejan bastante claro que con esta ley educativa han querido marcar una casilla en la lista de requisitos de la educación contemporánea, pero se ha quedado corta en proporcionar un marco sólido y efectivo para la implementación de la alfabetización digital y mediática en los centros educativos.

Además, la ley se queda corta en la mayoría de ocasiones: no hay referencias a la inteligencia artificial, ni a la CD, las fake news y la desinformación tampoco aparecen, la ética digital brilla por su ausencia. Todos estos son temas de rabiosa actualidad y su ausencia en el currículo incita a pensar en dos posibilidades, o que su conocimiento no es de interés de los que han desarrollado la ley educativa o que su incompetencia les ha impedido verlos. No se cuál de las dos opciones es peor.

Por último, destacar la ausencia de una aproximación desde la asignatura de filosofía y, de paso, romper una lanza a favor de esta disciplina, que, si bien los desafíos a los que se enfrenta la sociedad deben ser examinados desde un enfoque multidisciplinar, es el saber filosófico el que nos presta las herramientas más adecuadas para enfrentar estos conocimientos. Un acercamiento desde una filosofía crítica a los acontecimientos que dan forma a nuestro presente nos permite desnudar, como aquí en este estudio hemos hecho, los fenómenos que parecen inherentes a nosotros. Damos por hecho ciertas ideas, como la idea de "libertad", de

"democracia" o de "verdad", pero hay que tener en cuenta que estas sufren modificaciones continuamente porque son ideas tremendamente poderosas que se utilizan desde los poderes fácticos para moldear la opinión pública y generar narrativas que determinan nuestro presente. La labor de la filosofía en la educación debe alejarse del conocimiento filológico de obras y autores para acercarse a la realidad que nos rodea, que se da por hecho acríticamente para nuestro perjuicio, mediante el examen de los conceptos, de los sistemas y de los temas que están a la orden del día.

El progreso de la CD no va a detenerse en este punto en el que nos encontramos, los últimos avances en IA son prueba de ello, y las problemáticas que ya hemos visto en nuestra investigación podrían quedarse cortas de aquí a unos años. Esto sumado a la insuficiencia manifiesta de la legislación europea para hacer frente a todos estos problemas, nos hace poner todavía más la mirada en la labor de la filosofía en la educación como único medio para dar a la sociedad los instrumentos para discernir entre lo real y lo falso, entre lo bueno y lo malo, lo justo y lo injusto. El objetivo final, por tanto, de la filosofía en particular, y de la educación en general, debe ser el desengaño de una sociedad que se encuentra dormida, movida de un lado para otro por unos poderes que todavía no comprende.

## VI. Conclusiones

Llegados a este punto, queda por ver de qué manera hemos sido capaces de alcanzar los distintos objetivos marcados al inicio del estudio. Recordando lo propuesto al inicio, nos instábamos, a grandes rasgos, a analizar los comportamientos de las sociedades contemporáneas perpetrados por la CD, atendiendo, específicamente, a los tres tipos de relaciones que las definen, a saber, las relaciones económicas, las políticas (entre poder y pueblo) y las sociales (entre ciudadanos), y a la actuación desempeñada por las herramientas legislativas, comunitarias y nacionales, en la regulación de estas relaciones. Para ello, como paso previo, era imprescindible construir una interpretación clara de los conceptos e ideas a los que nos íbamos a referir, a través de una genealogía y una ontología de la CD. Vayamos, ahora, viendo los resultados de cada uno de estos pasos.

Partiendo de la primera fase de este estudio, en *2.¿Qué es la ciencia de datos?* hemos fijado la atención en las distintas disciplinas donde podemos encontrar los primeros antecedentes de la CD, hemos estudiado cómo se desarrolla el concepto de CD en la década de los 60, de la mano de John W. Tukey, y cómo el desarrollo tecnológico y científico de la computación ha dado lugar a una serie de transformaciones en el entorno que han propiciado la constitución y el avance del cuerpo de la CD hasta lo que conocemos hoy en día, formado por un conjunto amplísimo de herramientas, tecnologías y conceptos pertenecientes a las ciencias de las que se nutre: la computación, la estadística, la ciencia de la información. Con la imagen clara de los distintos componentes que dan forma a la disciplina de la CD, nos lanzamos a la determinación de esta como conjunto de saberes tecnológicos, afín a la idea de ciencia en su acepción de saber-hacer, y no como ciencia positiva, cuyo concepto y definición hemos explorado a través de la Teoría del Cierre Categorical del filósofo Gustavo Bueno, pero cuyas características no concuerdan con el cuerpo que la CD presenta.

Las ideas principales que obtenemos de esta primera fase del estudio son las siguientes:

- Podemos definir la CD como el conjunto de tecnologías y teorías, que, procedentes de la estadística, la computación y la ciencia de la información, se despliegan sobre distintos conjuntos de datos generados por la actividad en otros campos del saber o del hacer, que también son incorporados puntualmente al cuerpo de la CD, con el objetivo de alcanzar unos propósitos prácticos y funcionales para los sujetos que la practican. Es, por tanto, la CD una disciplina multidisciplinar y transdisciplinar de carácter eminentemente práctico.

- Las ciencias positivas, según la Teoría del Cierre Categorical de Gustavo Bueno, presentan dos características principalmente: crean el mundo a través de verdades científicas a medida que se ejercen y son capaces de cerrar un campo categorial propio donde los resultados de las operaciones pertenecen siempre al mismo. La CD no cumple con ninguna de estas condiciones y, por tanto, no puede ser considerada una ciencia.

De esta manera dejamos claro el objeto al que nos hemos estado refiriendo en todo el estudio y podemos continuar con el análisis de las problemáticas que se derivan de su implementación en los procesos y servicios que actúan e intervienen en los tres tipos de relaciones que hemos contemplado.

Siguiendo el orden lógico economía-política-sociedad, hemos desarrollado cómo, en una economía donde el valor de los productos se define en base a percepciones subjetivas de los compradores, la CD se ha extendido a todos los sectores económicos por la capacidad de sus herramientas para la predicción y manipulación de estos pensamientos subjetivos, generando ciertas problemáticas sobre el uso desmedido de unas tecnologías que aprovechan las debilidades cognitivas y emocionales de unos individuos, que son los mismos generadores de la información que las empresas emplean para su manipulación.

Las mismas posibilidades que tienen y aprovechan las empresas, son puestas en práctica desde los poderes políticos tanto occidentales como orientales. Hemos visto las consecuencias de la infraestructura de tecnologías de CD desplegada en ciudades chinas y la capacidad de los partidos políticos occidentales para influir en el voto de los ciudadanos en cuestiones tan importantes como unas elecciones democráticas. También el fenómeno de la posverdad ha sido analizado, viendo cómo los algoritmos de CD pueden favorecer a este tipo de contenido por encima del procedente de fuentes fiables, en base a criterios de repercusión e interacciones, dos conceptos también clave para comprender el funcionamiento de los entornos sociales digitales, donde el tercer tipo de relación analizada se desarrolla continuamente, en una especie de campo de batalla por la repercusión mediática, que provoca, en última instancia, la polarización y la creación de cámaras de eco, que modifican la visión del mundo de los individuos, impidiendo que se adapten correctamente al medio.

Por último, en lo que consideramos el final necesario a una investigación sobre una tecnología y sus problemáticas generadas, atendemos a la legislación como medio por el cual las sociedades políticamente organizadas son capaces de crear un marco de libertad que garantice la seguridad de sus ciudadanos a la par que mantiene su capacidad de operar. En este punto hemos dividido en siete bloques temáticos, elaborados a partir de los resultados del punto anterior, los distintos

aspectos que queremos analizar, y no podemos evitar señalar la insuficiencia de la legislación para regular las situaciones analizadas a lo largo de nuestro estudio. Si bien es cierto que se ha desarrollado un trabajo bastante considerable e informado (en algunos aspectos más que en otros), la velocidad con la que el campo de la CD evoluciona, como suele pasar con todas las nuevas tecnologías, es muchísimo mayor que la velocidad de los aparatos legislativos para regular estos fenómenos, lo que ha dejado en muchos aspectos desfasado el marco regulatorio. Este aspecto del derecho es algo que debemos tener siempre en cuenta, siempre va por detrás de las causas materiales que se necesitan regular, y en temas de tanto dinamismo e innovación puede resultar realmente complicado desarrollar una normativa que sea duradera en el tiempo.

Así, llegamos al último apartado de la fase correspondiente al estudio crítico, donde analizamos la ley educativa española y su aproximación a la CD y sus implicaciones, que muestra una incapacidad bastante mayor que las anteriores para dar solución al motivo que da nombre al bloque temático, la alfabetización digital y mediática. Aquí no podemos hacer más que repetir las palabras de esa última conclusión como resultado y conclusión final de nuestra investigación. Ante una legislación deficiente, no en su constitución necesariamente, sino en su aplicabilidad en el tiempo, la única forma de actuar en entornos digitales mediados por herramientas de CD, donde existe la posibilidad constante de ser manipulados, influidos o, simplemente, monitoreados, las posibilidades que tenemos como sociedad parten de una correcta educación que nos prepare para desenvolvemos en los entornos del presente. Los centros educativos no pueden preservar a los alumnos del entorno en el que van a desarrollarse y a operar con fines de protección y seguridad. Esta es una actitud que se está tomando, pero que es completamente errónea, pues en el momento en el que el alumnado termine la educación obligatoria tendrá que enfrentarse a un mundo que no mira por su seguridad, sino que es una suerte de biocenosis en la que cada uno debe buscarse sus formas de prosperar. La educación debe retomar el camino de la exigencia y la intensidad, en lugar de esquivar los desafíos que el presente plantea, las instituciones educativas deben abordarlos de frente, proporcionando a los estudiantes las habilidades que necesitarán para navegar y prosperar en estos nuevos paisajes.

## VII. Bibliografía

- Adorno, T. W., & Horkheimer, M. (2007). *Dialéctica de la Ilustración*. Ediciones AKAL.
- Althusser, L. (1994). *Ideología y aparatos ideológicos del estado: Louis Althusser*.
- Andersen, R. (2020). China's Artificial Intelligence Surveillance State Goes Global. *The Atlantic*. <https://www.theatlantic.com/magazine/archive/2020/09/china-ai-surveillance/614197/>
- Aristóteles. (2014). *Metafísica*.
- Asale, R.-. (s. f.). *posverdad | Diccionario de la lengua española*. «Diccionario de la lengua española» - Edición del Tricentenario. <https://dle.rae.es/posverdad>
- Babbage, C. (1826). *On a Method of Expressing by Signs the Action of Machinery*.
- Ball, J. (2017). *Post-Truth: How Bullshit Conquered the World*. Biteback Publishing.
- Bartlett, J. (2018). *The People Vs Tech: How the internet is killing democracy (and how we save it)*. Random House.
- Baudrillard, J. (1981). *Simulacres et simulation*. Editions Galilée.
- Bauman, Z. (2015). *Modernidad líquida*. Fondo de Cultura Económica.
- Bazerman, C. (1988). *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*.
- Bentham, J. (1791). *Panopticon Or the Inspection House*.
- Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. N. (2017). Julia: A Fresh Approach to Numerical Computing. *Siam Review*, 59(1), 65-98. <https://doi.org/10.1137/141000671>
- Borko, H. (1968). Information science: What is it? *American Documentation*, 19(1), 3-5. <https://doi.org/10.1002/asi.5090190103>
- Breiman, L. (2001a). *Statistical modelling: the two cultures*.
- Breiman, L. (2001b). Random Forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/a:1010933404324>
- Bueno, G., & Martínez, G. B. (1995). *¿Qué es la filosofía?: el lugar de la filosofía en la educación : el papel de la filosofía en el conjunto del saber constituido por el saber político, el saber científico y el saber religioso de nuestra época*.

- Bühler, K. (1934). *Sprachtheorie: Die Darstellungsfunktion der Sprache von Karl Bühler*.
- Campbell-Kelly, M., Aspray, W., Ensmenger, N., & Yost, J. R. (2013). *Computer: A History of the Information Machine*. Hachette UK.
- Carr, N. (2020). *The Shallows: What the Internet Is Doing to Our Brains*. W. W. Norton & Company.
- Castells, M. (2001). *La galaxia Internet*. Arete.
- Cavoukian, A. (2011). Privacy by design: The 7 foundational principles. Information and Privacy Commissioner of Ontario, Canada.
- Chamberlin, D. D., & Boyce, R. F. (1974). *SEQUEL*. <https://doi.org/10.1145/800296.811515>
- Ciencia de datos | PTI Ciencia Digital CSIC*. (s. f.). <https://pti-cienciadigital.csic.es/ciencia-de-datos/>
- Codd, E. F. (1970). A relational model of data for large shared data banks. *Communications of The ACM*, 13(6), 377-387. <https://doi.org/10.1145/362384.362685>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297. <https://doi.org/10.1007/bf00994018>
- Data Mining for Direct Marketing, Problems and Solutions*. (1999).
- Davenport, T. H., & Prusak, L. (2000). *Working Knowledge: How Organizations Manage What They Know*. Harvard Business Press.
- David, F. N., & Tukey, J. W. (1977). Exploratory Data Analysis. *Biometrics*, 33(4), 768. <https://doi.org/10.2307/2529486>
- De Cervantes Saavedra, M. (1605). *Don Quijote de la Mancha*.
- Dean, J., & Ghemawat, S. (2008). MapReduce. *Communications of The ACM*, 51(1), 107-113. <https://doi.org/10.1145/1327452.1327492>
- Deleuze, G. (2017). Postscript on the Societies of Control. En *Routledge eBooks* (pp. 35-39). <https://doi.org/10.4324/9781315242002-3>
- Directiva de Derechos de Autor en el Mercado Único Digital. (2019). Directiva (UE) 2019/790 del Parlamento Europeo y del Consejo, de 17 de abril de 2019, sobre los derechos de autor y los derechos afines en el mercado único digital y por la que se modifican las Directivas 96/9/CE y 2001/29/CE. Diario Oficial de la Unión Europea, L 130/92.

Ester, M., Kriegel, H., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial Databases with Noise. En *Knowledge Discovery and Data Mining* (pp. 226-231). <https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf>

Feced, C. G. (2022, 26 noviembre). Qué es la ley de crédito social de China, comparada con 'Black Mirror' Business Insider España. <https://www.businessinsider.es/ley-credito-social-china-que-es-1159317>

Foucault, M. (2012). *Vigilar y castigar: nacimiento de la prisión*.

Freedman, D. A. (2006). Statistical Models: Theory and Practice. *Technometrics*, 48(2), 315. <https://doi.org/10.1198/tech.2006.s403>

Frege, G. (1879). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5). <https://doi.org/10.1214/aos/1013203451>

Gruber, T. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199-220. <https://doi.org/10.1006/knac.1993.1008>

Hacking, I., & Hacking, E. U. P. I. (1990). *The Taming of Chance*. Cambridge University Press.

Haidt, J. (2012). *The Righteous Mind: Why Good People are Divided by Politics and Religion*. Penguin UK.

Han, J., Pei, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques*. Elsevier.

Hayek, F. A. (1945). *The Use of Knowledge in Society*.

Hearst, M. A. (1999). *Untangling text data mining*. <https://doi.org/10.3115/1034678.1034679>

Hey, A. J. G., Tansley, S., & Tolle, K. M. (2009). *The Fourth Paradigm: Data-intensive Scientific Discovery*.

Hinton, G. E., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. S., Jaitly, N., W, A., Senior, Vanhoucke, V., Nguyen, P. L., Sainath, T. N., & Kingsbury, B. (2012). Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Processing Magazine*, 29(6), 82-97. <https://doi.org/10.1109/msp.2012.2205597>

Kahneman, D., & Tversky, A. (1979). *Prospect Theory: An Analysis of Decision Under Risk*.

Kline, M. (1990). *Mathematical Thought from Ancient to Modern Times: Volume 1*. Oxford University Press.

Kluyver, T., Ragan-Kelley, B., Pérez, F. L., Granger, B. E., Bussonnier, M., Frederic, J., Kelley, K. P., Hamrick, J. B., Grout, J., Corlay, S., Ivanov, P., Avila, D. E. G., Abdalla, S., & Willing, C. (2016). Jupyter Notebooks – a publishing format for reproducible computational workflows. En *International Conference on Electronic Publishing* (pp. 87-90). <https://doi.org/10.3233/978-1-61499-649-1-87>

Kraft, V. (1966). *El círculo de Viena*.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. En *Neural Information Processing Systems* (Vol. 25, pp. 1097-1105). [http://books.nips.cc/papers/files/nips25/NIPS2012\\_0534.pdf](http://books.nips.cc/papers/files/nips25/NIPS2012_0534.pdf)

Lacan, J. (1997). *La ética del psicoanálisis*.

Lanier, J. (2013). *Who Owns the Future*. <http://ci.nii.ac.jp/ncid/BB12897197>

Lanier, J. (2018). *Ten Arguments For Deleting Your Social Media Accounts Right Now*. Random House.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, 521, 436-444.

Ley de Servicios Digitales (DSA). (2022). Reglamento del Parlamento Europeo y del Consejo relativo a los servicios digitales y por el que se modifica la Directiva 2000/31/CE (Reglamento de servicios digitales). Diario Oficial de la Unión Europea, C 32/1.

Ley Española de Educación (LOMLOE). (2020). Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación. Boletín Oficial del Estado, núm. 340, de 30 de diciembre de 2020, pp. 105910-105945.

Luhn, H. P. (1957). A Statistical Approach to Mechanized Encoding and Searching of Literary Information. *IBM journal of research and development*, 1(4), 309-317. <https://doi.org/10.1147/rd.14.0309>

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Berkeley*, 1, 281-297. [http://digitalassets.lib.berkeley.edu/math/ucb/text/math\\_s5\\_v1\\_article-17.pdf](http://digitalassets.lib.berkeley.edu/math/ucb/text/math_s5_v1_article-17.pdf)

Maestro, J. G. (2017). *Crítica de la Razón Literaria: el Materialismo Filosófico como Teoría, Crítica y Dialéctica de la Literatura*.

- Manyika, J. (2011). *Big Data: The Next Frontier for Innovation, Competition, and Productivity*.
- Martínez, G. B., & Bueno, G. (1972). *Ensayos materialistas*.
- Martínez, G. B., Bueno, G., & Sánchez, G. B. (1995). *Qué es la ciencia?: la respuesta de la teoría del cierre categorial*.
- Martínez, G. B., & Sánchez, G. B. (1993). *Teoría del cierre categorial (Vol. 1-5)*
- Marx, K. (1976). *EL CAPITAL. CRÍTICA DE LA ECONOMÍA POLÍTICA. LIBRO PRIMERO: EL PROCESO DE PRODUCCIÓN DEL CAPITAL*. Ediciones AKAL.
- McCulloch, W. S., & Pitts, W. (1990). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of mathematical biophysics*, 5(4), 115-133. <https://doi.org/10.1007/bf02478259>
- McIntyre, L. (2018). *Post-Truth*. MIT Press.
- Menger, C. (1997). *Principios de economía política*.
- Mitchell, T. M. (1997). *Machine Learning*.
- Morris, C. W. (1938). *Foundations of the Theory of Signs*.
- Nichols, T. M. (2017). *The Death of Expertise: The Campaign Against Established Knowledge and why it Matters*. Oxford University Press.
- Noveck, B. S. (2015). *Smart Citizens, Smarter State: The Technologies of Expertise and the Future of Governing*. Harvard University Press.
- Orwell, G. (1978). *Mi guerra civil española*.
- Orwell, G. (2013). *1984*. HarperCollins.
- Oxford Word of the Year 2016 | Oxford Languages*. (2020, 16 junio). <https://languages.oup.com/word-of-the-year/2016/>
- Page, L. M., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank Citation Ranking : Bringing Order to the Web. En *The Web Conference* (Vol. 98, pp. 161-172). <http://dbpubs.stanford.edu:8090/pub/1999-66/>
- Pariser, E. (2011). *The Filter Bubble: What The Internet Is Hiding From You*. Penguin UK.
- Petzold, C. (1999). *Code: The Hidden Language of Computer Hardware and Software*. <https://ptgmedia.pearsoncmg.com/images/9780735611313/samplepages/9780735611313.pdf>

Platón. (2014). *La República o El Estado*. Grupo Planeta Spain.

Popper, K. (1991). Conjeturas y refutaciones: El desarrollo del conocimiento científico. En *Paidós Ibérica eBooks*. <http://atlas.umss.edu.bo:8080/jspui/bitstream/123456789/553/1/LD-300-145.pdf>

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81-106. <https://doi.org/10.1007/bf00116251>

Rainie, L., & Wellman, B. (2012). *Networked: The New Social Operating System*. MIT Press.

Ramsay, J., & Silverman, B. W. (2006). *Functional Data Analysis*. Springer Science & Business Media.

Reglamento General de Protección de Datos (GDPR). (2016). Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos. Diario Oficial de la Unión Europea, L 119/1.

Ricardo, D. (2007). *Principios de Economía Política y Tributacion*. Claridad.

Ruder, S. (2016). An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747.

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Salton, G. (1968). *Automatic information organization and retrieval*. <http://ci.nii.ac.jp/ncid/BA07535881>

Samuel, A. L. (1969). *Some Studies in Machine Learning Using the Game of Checkers: Recent Progress*.

Schiller, D. (1999). *Digital Capitalism: Networking the Global Market System*. MIT Press.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3), 379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>

Simon, H. A. (1997). *Administrative Behavior, 4th Edition*. Simon and Schuster.

Simon, H. A. (2019). *The Sciences of the Artificial, reissue of the third edition with a new introduction by John Laird*. MIT Press.

Singer, P. W., & Brooking, E. T. (2018). *Likewar: The Weaponization of Social Media*. Eamon Dolan Books.

Smith, A. (1794). *La Riqueza de las Naciones*. Oficina de Viuda e Hijos de Santander. Valladolid 1794.

Stigler, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty Before 1900*. Harvard University Press.

Stillwell, J. (2020). *Mathematics and Its History: A Concise Edition*. Springer Nature.

Strang, G. (1988). *Linear Algebra and Its Applications*. Brooks/Cole Publishing Company.

Susskind, J. (2018). *Future Politics: Living Together in a World Transformed by Tech*. Oxford University Press.

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to Sequence Learning with Neural Networks. En *Neural Information Processing Systems* (Vol. 27, pp. 3104-3112). <http://cs224d.stanford.edu/papers/seq2seq.pdf>

Sutton, R. S., & Barto, A. G. (2005). Reinforcement Learning: An Introduction. *IEEE Transactions on Neural Networks*, 16(1), 285-286. <https://doi.org/10.1109/tnn.2004.842673>

Sweeney, L. (2002). k-ANONYMITY: A MODEL FOR PROTECTING PRIVACY. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05), 557-570. <https://doi.org/10.1142/s0218488502001648>

Tesich, S. (1992). A Government of Lies. *The Nation*.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press.

Tsichritzis, D., & Lochovsky, F. (1982). *Data Models*. Prentice-Hall, Inc.

Tukey, J. W. (1962). The Future of Data Analysis. *Annals of Mathematical Statistics*, 33(1), 1-67. <https://doi.org/10.1214/aoms/1177704711>

Turing, A. M. (1937). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of The London Mathematical Society*, s2-42(1), 230-265. <https://doi.org/10.1112/plms/s2-42.1.230>

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing*, 17(4), 395-416. <https://doi.org/10.1007/s11222-007-9033-z>

Von Mises, L. (2016). *Human Action*. Lulu Press, Inc.

Wu, C. F. J. (1997). Statistics = Data Science? Inaugural lecture for the chair in statistics at the University of Michigan.

Wylie, C. (2020). *Mindf\*ck: Cambridge Analytica and the Plot to Break America*. Random House Trade Paperbacks.

Yang, Z. (2022, 1 diciembre). Ni vigilancia total ni problema inmediato: radiografía del crédito social chino (A. Milutinovic, Trad.). MIT Technology Review. <https://www.technologyreview.es//s/14861/ni-vigilancia-total-ni-problema-inmediato-radiografia-del-credito-social-chino>

Young, D. (2008). *Distraction: A Philosopher's Guide To Being Free*. Melbourne Univ. Publishing.

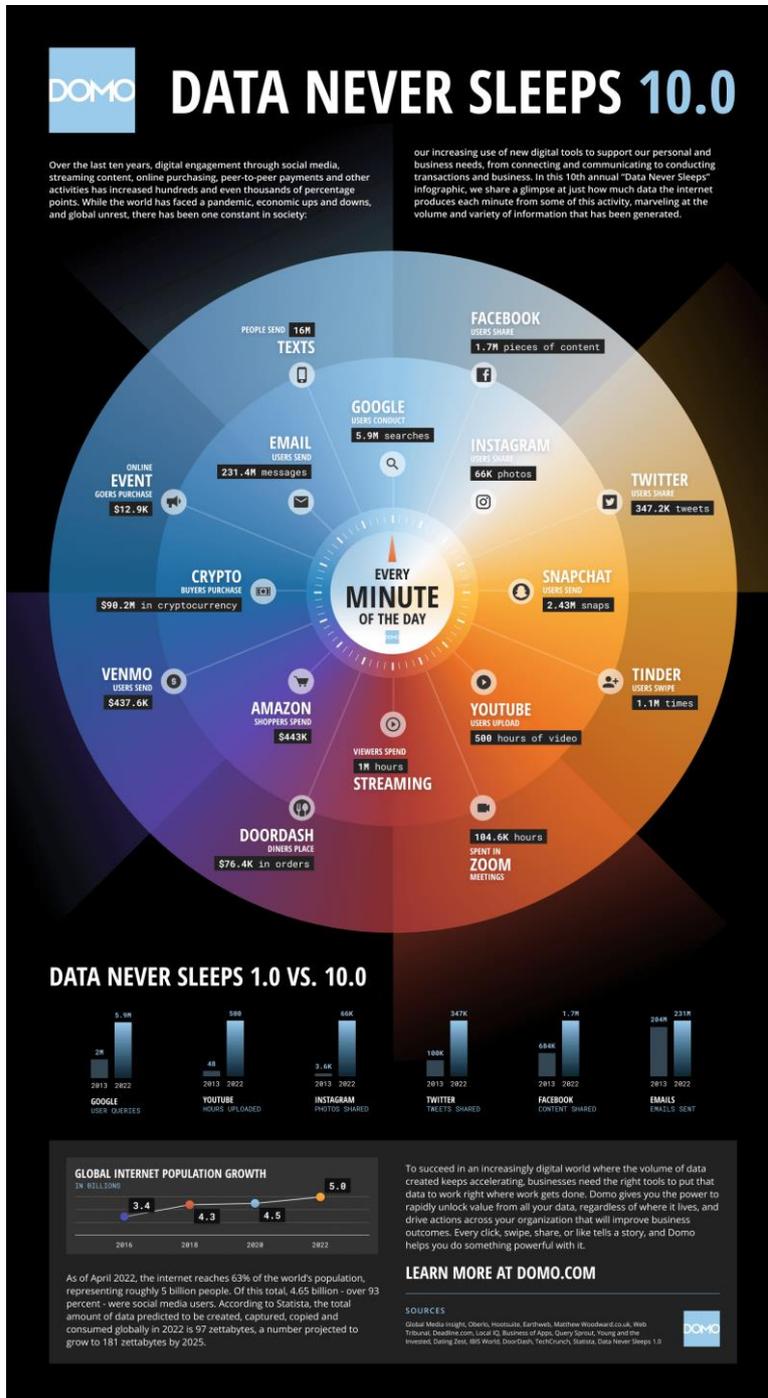
Zhang, L., & Gombeaud, A. (2020). *Tiananmen 1989: Our Shattered Hopes*. National Geographic Books.

Zikopoulos, P., & Eaton, C. (2011). *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. <https://dl.acm.org/citation.cfm?id=2132803>

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Books.

## VIII. Anexos

- Anexo 1:



Domo Inc. (2022). Data Never Sleeps 9.0. [Infografía].

<https://www.domo.com/learn/data-never-sleeps-9>