

Gobernanza algorítmica

Revista Latinoamericana de Economía y Sociedad Digital

Issue Número Especial 1, Julio 2022

Autores: [Christian Katzenbach](#) , Lena Ulbricht

DOI: [10.53857/CWYI6959](https://doi.org/10.53857/CWYI6959)

Cita sugerida: Katzenbach, C., & Ulbricht, L. (2022). Gobernanza Algorítmica. Revista Latinoamericana de Economía Y Sociedad Digital. <https://doi.org/10.53857/cwyi6959>

Licencia: Creative Commons Atribución-NoComercial 4.0 Internacional ([CC BY-NC 4.0](#))

Tipo: [Ensayo](#), [Traducción](#)

Palabras clave: [automatización](#), [gobernanza](#), [gobernanza algorítmica](#), [moderación de contenidos](#), [politización](#), [Regulación](#), [transparencia](#)

Nota sobre la traducción

Este artículo fue originalmente publicado en la sección [Concepts of the Digital Society](#) de la revista Internet Policy Review, editada por Christian Katzenbach y Thomas Christian Bächle. Es publicado en la Revista Latinoamericana de Economía y Sociedad Digital según los lineamientos de acceso abierto del artículo original, con licencia Creative Commons Attribution 3.0 Germany.

Título Original: Algorithmic Governance

Fecha de publicación original: 29 de Noviembre de 2019

Link al artículo original: <https://policyreview.info/concepts/algorithmic-governance>

DOI original: <https://doi.org/10.14763/2019.4.1424>

Traducido del idioma inglés por [The Pillow Books](#) para la Revista Latinoamericana de Economía y Sociedad Digital en Mayo 2022.

Resumen

En cuanto concepto clave en controversias sobre la sociedad digital emergente, la gobernanza algorítmica pone de manifiesto la idea de que las tecnologías digitales producen cierto ordenamiento social de una forma específica. Este artículo comienza por los orígenes del concepto y presenta diferentes perspectivas y objetos de estudio en los que la

gobernanza algorítmica ha adquirido prominencia, y que van desde el sector público hasta la gestión laboral y el ordenamiento de la comunicación digital. En todos los sectores, las controversias recurrentes, como la datificación y la vigilancia, los sesgos, la agencia y la transparencia, indican que el concepto de gobernanza algorítmica nos permite incorporar a una conversación colectiva objetos de estudio y campos de investigación que no se habían relacionado hasta ahora. Los estudios de caso breves sobre políticas predictivas y moderación automatizada de contenidos muestran que la gobernanza algorítmica es múltiple, contingente y disputada. Adquiere distintas formas en distintos contextos y jurisdicciones, y está conformada por intereses, poder y resistencia.

Abstract

Algorithmic governance as a key concept in controversies around the emerging digital society highlights the idea that digital technologies produce social ordering in a specific way. Starting with the origins of the concept, this paper portrays different perspectives and objects of inquiry where algorithmic governance has gained prominence ranging from the public sector to labour management and ordering digital communication. Recurrent controversies across all sectors such as datafication and surveillance, bias, agency and transparency indicate that the concept of algorithmic governance allows to bring objects of inquiry and research fields that had not been related before into a joint conversation. Short case studies on predictive policy and automated content moderation show that algorithmic governance is multiple, contingent and contested. It takes different forms in different contexts and jurisdictions, and it is shaped by interests, power, and resistance.

1. INTRODUCCIÓN

El concepto de gobernanza algorítmica surgió durante la última década, pero retoma una idea que ha estado presente durante un tiempo mucho más prolongado: que las tecnologías digitales estructuran lo social de formas particulares. Enfrentarse al concepto de gobernanza algorítmica resulta complejo, ya que muchos campos de investigación están interesados en el fenómeno, utilizan distintos términos y tienen distintos enfoques. La investigación sobre lo que constituye la gobernanza algorítmica es una importante contribución a la teoría social contemporánea, dado que cuestiona el papel de los algoritmos y su efecto ordenador. Definimos *algoritmos* como procedimientos epistémicos computarizados que son particularmente complejos, aunque lo complejo depende del contexto. Los algoritmos conforman procedimientos a partir de sus lógicas matemáticas y prácticas estadísticas inherentes. Por ello, el discurso en torno a la gobernanza algorítmica a menudo se superpone y se cruza con debates sobre la datificación (cf. Mejias y Couldry, 2019 como parte de esta sección especial) y la inteligencia artificial (IA). No obstante, los algoritmos también operan a veces con conjuntos pequeños de datos o *small data* y utilizan procedimientos de cálculo que no aprenden ni son adaptativos.

Aun cuando *gobernanza* es un término controvertido, definimos su núcleo como la coordinación entre actores basada en reglas. Más allá de la regulación, la gobernanza no es necesariamente intencional ni dirigida a objetivos (Black, 2001), y además incluye una coordinación no intencionada (Hofmann, Katzenbach y Gollatz, 2016). Sin embargo, la gobernanza excluye todas las formas de ordenamiento social de carácter puramente ocasionales y que no se basan en algún tipo de regla; la gobernanza supone un grado mínimo de estabilidad que es necesario para que los actores desarrollen expectativas, las cuales son un prerequisite para la coordinación (Hobbes, 1909). Elegimos utilizar el término *gobernanza algorítmica* en lugar de *regulación algorítmica* porque la gobernanza nos permite dar cuenta de la multiplicidad del ordenamiento social en lo que toca a actores, mecanismos, estructuras, grados de institucionalización y distribución de la autoridad. Este término abarca de forma deliberada el ordenamiento social que está descentralizado desde el punto de vista analítico y estructural y que no se centra en el Estado. Así pues, *gobernanza algorítmica* refleja mejor la ambición de este artículo: escudriñar a fondo las formas en que los algoritmos generan orden social. En este sentido, nos enfocamos en la gobernanza *por* algoritmos en lugar de la gobernanza *de* algoritmos (Musiani, 2013; Just y Latzer, 2017). En suma, la gobernanza algorítmica es una forma de ordenamiento social que depende de la coordinación entre actores, se basa en reglas e incorpora procedimientos epistémicos computarizados particularmente complejos.

La relevancia de enfrentarse a la gobernanza algorítmica se hace evidente en lo que toca a las narrativas antagónicas sobre lo que *cambia* en la gobernanza cuando se hace uso de algoritmos: una narrativa es, por ejemplo, que la gobernanza se vuelve más poderosa, invasiva y dominante. Una narrativa distinta hace énfasis en que la gobernanza se vuelve más inclusiva, receptiva y permite una mayor diversidad social, tal como subrayamos en los siguientes capítulos.

Si consideramos el concepto de manera amplia, sus raíces se pueden rastrear en la historia y la sociología de la ciencia, la tecnología y la sociedad. La tecnología siempre ha reflejado y reorganizado lo social (Bijker y Law, 1992; Latour, 2005). Desde las inquietudes de Sócrates en relación con la escritura y la alfabetización (Ong, 1982), hasta la conexión radicalmente interdisciplinaria que hay en la cibernética entre los sistemas técnico, biológico y social, así como su control (Wiener, 1948), pasando por la distopía burocrática de una “sociedad tecnológica” según Jacques Ellul (1964) y hasta la muy citada, pero disputada, “política de los artefactos” de Langdon Winner (1980), la idea de que la tecnología y los artefactos gobiernan de alguna manera la sociedad y las interacciones sociales es un tema recurrente. El antecedente más directo de la gobernanza algorítmica es el famoso lema de Lawrence Lessig, “el código es ley”. Aquí, el código de *software* o, de manera más general, las arquitecturas técnicas, son vistas como uno de cuatro factores que regulan el comportamiento social (junto con la ley, el mercado y las normas sociales). Los estudiosos también han conceptualizado el carácter institucional del *software* y los algoritmos (Katzenbach, 2017, 2012; Napoli, 2013; Orwat *et al.*, 2010). Mientras que Rouvroy y Berns utilizaron el término “*gouvernance algorithmique*” en 2009, los primeros en conceptualizar

el término “*algorithmic governance*” fueron Müller-Birn, Dobusch y Herbsleb (2013), quienes lo presentaron como un mecanismo de coordinación opuesto a la “gobernanza social”^[1]. El concepto de “regulación algorítmica” fue introducido por el editor estadounidense Tim O’Reilly (2013), y subrayaba la eficiencia de los espacios gobernados de forma automática, pero pasaba por alto la despolitización de temas muy controversiales que tiene lugar cuando éstos se delegan a las soluciones tecnológicas (Morozov, 2014). En contraste con el determinismo tecnológico implícito en estas explicaciones, el campo interdisciplinario de los estudios críticos del *software* ha complejizado —en el mejor sentido— las intrincadas dependencias mutuas entre el *software* y los algoritmos, por una parte, y las interacciones y estructuras sociales, por la otra (MacKenzie, 2006; Fuller, 2008; Berry, 2011; Kitchin y Dodge, 2011). Este artículo se propone ofrecer un manual sobre el concepto de gobernanza algorítmica, incluido un resumen sobre las perspectivas y áreas de interés principales (sección 2); una presentación de controversias recurrentes en este espacio (sección 3); una descripción analítica de los distintos tipos de gobernanza algorítmica (sección 4); y un breve debate sobre las políticas predictivas y la moderación automatizada de contenidos como casos de estudio ilustrativos (sección 5). Buscamos eludir el ímpetu determinista de la trayectoria hacia una automatización cada vez mayor, al tiempo que nos tomamos en serio el giro hacia la gestión cada vez más común de espacios e interacciones sociales mediante sistemas algorítmicos.

2. GOBERNANZA ALGORÍTMICA: PERSPECTIVAS Y OBJETOS DE ESTUDIO

La noción de gobernanza algorítmica se aborda y discute en distintos contextos y disciplinas. Todas ellas comparten una comprensión similar sobre la importancia de los algoritmos para el ordenamiento social, pero eligen distintos objetos de estudio. La selección de conjuntos relevantes de literatura aquí presente se centra en la investigación realizada desde los estudios sobre ciencia y tecnología (STS, por sus siglas en inglés), la sociología, la ciencia política, la comunicación y las ciencias de la comunicación, pero incluye también investigaciones de otras disciplinas relevantes interesadas en la gobernanza algorítmica, como las ciencias informáticas, los estudios jurídicos, la economía y la filosofía.

Varias áreas de investigación estrechamente relacionadas y superpuestas se interesan por la forma en que los algoritmos contribuyen a reorganizar y cambiar las interacciones y las estructuras sociales. Sin embargo, en contraste con lo que sucede en el debate público, en estas investigaciones se rechaza la noción de algoritmo como fuerza independiente, externa, que gobierna por sí sola nuestro mundo. Complejizan este panorama tecnodeterminista al plantear la gran relevancia que tienen los algoritmos (Gillespie, 2014) y, al mismo tiempo, poner de manifiesto los contextos económicos, culturales y políticos que conforman el diseño de los algoritmos y permiten su funcionamiento. De esta manera, los estudios empíricos en este campo suelen enfocarse en las interacciones sociales estudiadas e

interrogan el papel de los algoritmos y su efecto ordenador en estos contextos específicos (Kitchin, 2016; Seaver, 2017; Ziewitz, 2016). También comparten un interés por la forma en que los conjuntos de datos, los modelos matemáticos y los procedimientos de cálculo allanan el camino hacia una nueva cualidad de la cuantificación y la clasificación social. Las nociones de “regulación algorítmica” (Yeung, 2018) y de “gobernanza algorítmica” (Just y Latzer, 2016; König, 2019) emanan del campo de investigación sobre regulación y gobernanza, compuesto sobre todo por especialistas en la ciencia jurídica, la ciencia política, la economía y la sociología. El efecto de los estudios pertinentes ha sido organizar y estimular la investigación sobre la gobernanza algorítmica con una comprensión compartida sobre la regulación entendida como los “esfuerzos intencionales para gestionar el riesgo o alterar el comportamiento a fin de lograr cierto objetivo previamente especificado” (Yeung, 2018). Este enfoque sobre las intervenciones intencionales, dirigidas a objetivos, prepara el escenario para investigaciones con un interés explícito en los algoritmos como una forma de gobierno empleada a propósito para regular los contextos sociales y alterar el comportamiento de las personas, por ejemplo, en el trato a la ciudadanía o la gestión de la clase trabajadora. Otros enfoques también estudian las formas no deliberadas de ordenamiento social a través de y con algoritmos.

Un enfoque ligeramente distinto coloca *los sistemas técnicos en el centro*, y no las estructuras y relaciones sociales. Los estudios pertinentes, en particular los de las ciencias informáticas, apuntan a construir y optimizar sistemas algorítmicos para resolver problemas sociales específicos: detectar contenido controvertido, comportamientos anómalos y preferencias u opiniones. En pocas palabras, construyen los instrumentos mismos que suelen emplearse en la gobernanza algorítmica. La meta común de este enfoque suele ser detectar de forma eficaz patrones en los datos, por ejemplo, traducir el contexto social a procesos computables (es decir, optimizar la detección). Esta corriente de investigación busca formas eficientes, sólidas, justas y responsables para clasificar tanto sujetos como objetos en categorías generales (como especies) así como en dimensiones específicas, como tipos psicométricos, estados emocionales, capacidad crediticia o preferencias políticas (Schmidt y Wiegand, 2017; Binns *et al.*, 2017). Los productores y los proveedores de servicios impulsados por algoritmos no solo optimizan la detección de patrones en conjuntos de datos existentes, sino que, a su vez, a menudo también buscan optimizar sus sistemas para dar un pequeño empujón al comportamiento de la persona usuaria de la manera más efectiva, con miras a maximizar los beneficios para la organización (optimización de comportamientos). Al poner a prueba de forma sistemática distintas versiones de pantallas de usuario u otros aspectos (prueba A/B) y al realizar análisis tanto de las personas usuarias como de sus comportamientos, las empresas trabajan de forma continua para dirigir las interacciones de éstas de manera más eficaz hacia un mayor involucramiento y una menor fricción (Guerses *et al.*, 2018). Sin embargo, es importante señalar que no existe una división clara entre la investigación que desarrolla y optimiza la gobernanza algorítmica y la investigación que analiza sus implicaciones para la sociedad; ambas se superponen y hay muchos estudios que apuntan hacia ambos objetivos. Un ejemplo son los estudios sobre

sesgos algorítmicos, imparcialidad y rendición de cuentas que conceptualizan y al mismo tiempo ponen a prueba sus parámetros (por ejemplo, Waseem y Hovy, 2016). Otra área importante de investigación, que es a la vez aplicada y crítica, son los estudios sobre “sesgo de automatización”, “sesgo de máquina” o “dependencia excesiva” que analizan bajo qué condiciones los agentes humanos pueden tomar una decisión verdaderamente autónoma (Lee y See, 2004; Parasuraman y Manzey, 2010).

Un ámbito importante de investigación, relevante sobre todo para los STS, la comunicación y las ciencias de la comunicación, es la *comunicación digital y las redes sociales*. Desde hace más de una década, las personas que se dedican a su estudio han mostrado interés por la forma en que los buscadores y las plataformas de redes sociales organizan y estructuran la información que está disponible en la red y en cómo esto afecta a la subjetivación (Couldry y Langer, 2005). Las plataformas dan prioridad a ciertos tipos de contenido (por lo general, con base en parámetros de “*engagement*” o “*involucramiento*”), dando forma a un nuevo modo dominante de atribuir relevancia en la sociedad, que complementa las rutinas periodísticas tradicionales. Las plataformas también despliegan algoritmos para regular el contenido por medio de bloqueos o filtros de lenguaje, videos y fotografías que se consideran inaceptables o ilegales (Gillespie, 2018; Gorwa, 2019). En una escala cada vez mayor y con una creciente presión política, las plataformas recurren sin reparos a soluciones técnicas para atender los difíciles dilemas que plantea la gobernanza de plataformas como el discurso de odio, la desinformación y los derechos de autor (Gorwa, Binns y Katzenbach, 2019). Otras áreas estudiadas que hacen uso de la detección automatizada de contenidos son las verificaciones de plagio en la escritura escolar y académica (Introna, 2016) y el análisis de sentimiento (o minería de opinión) para la mercadotecnia comercial y política (Tactical Tech, 2019).

Las estipulaciones de servicio en el sector público, la gestión ciudadana y la vigilancia constituyen otra área de interés clave para quienes se dedican al estudio de la gobernanza algorítmica. En particular, especialistas de las áreas de las ciencias políticas y el ámbito jurídico investigan sobre procedimientos automatizados para la provisión de servicios estatales y la toma de decisiones administrativas. Aquí, la meta es que los algoritmos aumenten potencialmente la eficiencia y eficacia de los servicios estatales, por ejemplo, al racionalizar la toma de decisiones burocráticas mediante el direccionamiento de la información y las intervenciones hacia perfiles precisos o la elección de las mejores opciones en materia de política disponibles (OECD, 2015). Sin embargo, estas promesas han sido muy cuestionadas. Las personas que se dedican a su estudio han mostrado que el despliegue de sistemas algorítmicos en el sector público ha dado lugar a muchas consecuencias no planeadas ni reveladas (Veale y Brass, 2019; Dencik, Hintz, Redden y Warne, 2018). La aplicación de herramientas algorítmicas en el gobierno a menudo se basa en nuevas formas de vigilancia y clasificación de la población llevadas a cabo por el Estado y los actores corporativos (Neyland y Möllers, 2017; Lupton, 2016; Bennet, 2017). Los fundamentos de muchos proyectos de prestación de servicios digitales y elección de políticas con base en algoritmos son sistemas para clasificar, calificar y predecir el

comportamiento, las preferencias y la opinión de la ciudadanía. Estos sistemas se utilizan para la asignación de beneficios sociales, para combatir la evasión fiscal y el fraude, para dar forma a la jurisdicción, para la vigilancia policial y la prevención del terrorismo, así como para el control de fronteras y la gestión de la migración.

La clasificación y la calificación no solo se aplican a la ciudadanía, sino también a las personas consumidoras, como han señalado los estudios sobre valoración y cuantificación referentes a la capacidad crediticia (Avery, Brevoort y Canner, 2012; Brevoort, Grimm y Kambara, 2015; Fourcade y Healy, 2016, 2017). Estos estudios señalan la forma en que las prácticas de valoración basadas en algoritmos configuran los mercados y crean mecanismos de estratificación que pueden superponer la clase social y reconfigurar las relaciones de poder, a menudo en detrimento de las personas en situación de pobreza y aquellas “mal apreciadas” (Fourcade y Healy, 2017; Zarsky, 2014).

La gobernanza a través de algoritmos también constituye también un tema importante para las personas que se dedican al estudio de la transformación digital del trabajo, como la sociología laboral o la economía laboral. Los objetos de estudio aquí son la gobernanza automatizada en plataformas de trabajo y la *gestión laboral* dentro de las empresas, por ejemplo, a través de la gestión de la productividad y los sistemas de evaluación (Lee, Poltrock, Barkhuus, Borges y Kellogg, 2017; Rosenblat, 2018). Este campo de investigación se caracteriza por los estudios de caso empíricos que examinan las implicaciones que tienen la gestión algorítmica y la vigilancia en los centros de trabajo sobre los salarios, la autonomía, el bienestar, los derechos y la seguridad social de la clase trabajadora; así como sobre la desigualdad social y los Estados de bienestar (Wood, Graham, Lehdonvirta y Hjorth, 2019). Algunos objetos de estudio relacionados son los sistemas algorítmicos de realidad aumentada, reconocimiento de voz y sistemas de asistencia para ejecución de tareas, capacitación y control de calidad (Gerber y Krzywdzinski, 2019). Algunos sectores económicos importantes bajo estudio son la logística, la producción industrial, la paquetería y los servicios. Otras áreas de interés relevantes se enfocan en la gestión algorítmica del transporte y el tráfico, la energía, los residuos y el agua, por ejemplo, en proyectos de “ciudades inteligentes”.

Entre las personas que se dedican a su estudio, hay quienes abordan la gobernanza algorítmica en un metanivel, como una forma descentralizada de *coordinación y participación*. Subrayan su capacidad para procesar un gran número de entradas o *inputs* y así lidiar con un alto grado de complejidad. En consecuencia, ven la gobernanza algorítmica como un modo de coordinación que ofrece nuevas oportunidades para la participación, la inclusión social, la diversidad y la capacidad de respuesta democrática (König, 2019; Schrape, 2019). Hay numerosas investigaciones sobre las posibilidades que el *software* puede ofrecer para mejorar la participación política a través de las herramientas de participación en línea (Boulianne, 2015; Boulianne y Theocharis, 2018), como las elecciones y las peticiones electrónicas, la comunicación en redes sociales y la participación masiva o *crowdsourcing* legislativo. Además, se están desarrollando infinitas herramientas

algorítmicas con el objetivo explícito de “escuchar a más voces” y de mejorar la relación entre personas usuarias y plataformas o entre la ciudadanía y las élites políticas. No obstante, la gobernanza algorítmica a través de herramientas de participación es aún a menudo jerárquica y se caracteriza por una distribución desigual del poder (Kelty, 2017).

3. CONTROVERSIAS E INQUIETUDES

En todas estas perspectivas y sectores diversos existen controversias e inquietudes recurrentes que por lo general se traen a colación cada vez que se discute el fenómeno de la gobernanza algorítmica. Al observar estas controversias más de cerca a menudo podemos detectar un movimiento dialéctico entre las connotaciones positivas y las negativas.

DATIFICACIÓN Y VIGILANCIA

La literatura sobre gobernanza algorítmica muestra un amplio consenso en lo que respecta a la forma en que el *big data*, los algoritmos y la inteligencia artificial cambian las perspectivas de la sociedad sobre ciertas poblaciones e individuos. Esto se debe al “diluvio de datos”, a un crecimiento y una diversificación de los datos recolectados por los aparatos digitales, a los rastreadores en línea y a la vigilancia de espacios (Beer, 2019). La “datificación” (cf. Mejias y Couldry, 2019 como parte de esta sección especial) también se beneficia de infraestructuras cada vez más poderosas que permiten un análisis de datos mayor y más rápido, y de normas sociales que benefician la cuantificación, la clasificación y la vigilancia (Rieder y Simon, 2016). La investigación sobre gobernanza algorítmica siempre se ha preocupado, no obstante, por los muchos riesgos de la datificación y la vigilancia. Vigilar a poblaciones enteras y crear perfiles detallados de las personas con base en sus “dobles de datos” genera amplias oportunidades para la clasificación social, la discriminación, la opresión estatal y la manipulación de la ciudadanía y de las personas consumidoras (Lyon, 2014; Gandy, 2010). La vigilancia sin restricciones representa un riesgo para muchos derechos civiles y humanos, como la libertad de expresión, la libertad de reunión y la privacidad, por mencionar unos cuantos.

AGENCIA Y AUTONOMÍA

La ubicuidad de los algoritmos como herramientas de gobernanza ha generado inquietudes en torno a los efectos que tienen sobre la agencia y la autonomía humanas (Hildebrandt, 2016), este último un concepto central de la Ilustración y una característica fundamental del individuo moderno. Mientras que algunos enfoques anteriores concebían los algoritmos como potenciadores o reductores de la agencia humana, se ha vuelto claro que la interacción entre agentes humanos y máquinas es compleja y requiere de una diferenciación mayor. Mientras que las tipologías y los debates suelen construir una distinción binaria entre humanos que están al tanto (de las opciones y las decisiones) y humanos que no lo están, esta dicotomía no se sostiene en análisis profundos de las realidades múltiples que caracterizan la interacción humano-computadora (Gray y Suri, 2018). Además, la agencia

humana no puede evaluarse únicamente en lo tocante a las máquinas, sino que también debe considerar las restricciones planteadas por organizaciones y normas sociales (Caplan y boyd, 2018).

TRANSPARENCIA Y OPACIDAD

La supuesta opacidad de los algoritmos y la gobernanza algorítmica es un tema fuerte y persistente en el debate, que suele ir de la mano con el llamado a una mayor transparencia (Kitchin, 2016; Pasquale, 2015). Sin embargo, argumentos más recientes señalan que el acceso al código informático no debería volverse un fetiche: la transparencia absoluta a menudo no es posible ni deseable, ni es la solución a la mayoría de los problemas relacionados con la gobernanza algorítmica, como la imparcialidad, la manipulación, la civilidad, etcétera (Ananny y Crawford, 2017; Mittelstadt, Allo, Taddeo, Wachter y Floridi, 2016). Además, la implementación de normas sociales en el código no solo genera opacidad, sino que también revela normas y procesos que antes estaban ocultos. Un buen ejemplo son las controversias en torno a los sistemas de puntaje sobre riesgo de desempleo, tal como se desataron en Austria y Polonia (AlgorithmWatch y Bertelsmann Stiftung, 2019), y sobre capacidad crediticia (AlgorithmWatch, 2019). El interés público en la gobernanza algorítmica ha motivado a los actores de la sociedad civil y a quienes se dedican a su estudio a investigar sobre la composición y la racionalidad del puntaje algorítmico y a cuestionar los valores sociales que le subyacen. Dado lo anterior, el giro actual hacia la gobernanza algorítmica podría incluso favorecer una mayor transparencia, ya que el código de *software*, una vez revelado, requiere de la articulación de supuestos subyacentes en modelos explícitos.

DESPOLITIZACIÓN Y REPOLITIZACIÓN

En una lógica similar, existe un animado debate público sobre los efectos despolitizantes y repolitizantes de los algoritmos. Los algoritmos a menudo han sido criticados como despolitizantes debido a su “aura de objetividad y verdad” (boyd y Crawford, 2012) y su promesa de solucionar problemas de complejidad social solamente por el tamaño de los datos y el mayor poder de cómputo (Kitchin, 2013; Morozov, 2013). No obstante, y como consecuencia, muchos estudios han cuestionado la idea de que los algoritmos pueden ser objetivos y neutrales. La desigualdad social, la injusticia y la discriminación se traducen en conjuntos y prácticas de datos sesgados. Esta nueva sospecha pública sobre las implicaciones sociales de los algoritmos ha motivado críticas que buscan cuestionar de manera parecida las motivaciones de las campañas políticas, la desigualdad social en la provisión de servicios públicos y las implicaciones de la vigilancia corporativa sobre los derechos civiles. De esta manera, la gobernanza algorítmica ha contribuido a repolitizar la gobernanza y la toma de decisiones en algunas áreas. No obstante, éste podría ser un logro de corto aliento, ya que la instalación de la gobernanza algorítmica en cuanto infraestructura social conducirá sin duda a su integración profunda en nuestras rutinas a lo largo del tiempo, y al final se dará por sentada, como casi todas las infraestructuras una vez

que están colocadas (Plantin *et al.*, 2018; Gorwa, Binns y Katzenbach, 2019).

SESGO E IMPARCIALIDAD

Otra inquietud clave es la del *sesgo algorítmico*. La toma de decisiones automatizada por medio de sistemas algorítmicos suele favorecer a las personas y los colectivos que ya gozan de privilegios, mientras que discriminan contra las personas en situación de marginación (Noble, 2018). Si bien es cierto que esto constituye una inquietud importante que se debe atender en la creciente automatización de lo social, no se trata de un fenómeno nuevo, y el algoritmo no es el (único) culpable. Los conjuntos de datos y las reglas de decisión sesgados también generan discriminación. Esto más bien pone de relieve la observación general de que cualquier procedimiento tecnológico y burocrático materializa clasificaciones como el género, la clase social, el espacio geográfico y la raza. Éstos no se originan en los sistemas, sino que son un mero reflejo de los sesgos y prejuicios prevalentes, las desigualdades y estructuras de poder y, una vez en funcionamiento, amplifican una y otra vez las desigualdades inscritas. La actual politización de estos temas puede considerarse como una oportunidad para pensar cómo puede darse una mayor imparcialidad en sociedades que ya cuentan con sistemas automatizados (Barocas y Selbst, 2016; boyd y Barocas, 2017; Hacker, 2018).

4. DE “FUERA DE CONTROL” A “FAVORABLE A LA AUTONOMÍA”: EVALUACIÓN DE TIPOS DE GOBERNANZA ALGORÍTMICA

Conforme los sistemas algorítmicos se expandan hacia diversos sectores sociales, otros campos de investigación se desarrollarán, se integrarán y crearán subcampos. Al mismo tiempo, las controversias cambiarán y darán forma a futuros desarrollos. Esto hace difícil e incluso imposible sintetizar la diversidad de perspectivas sobre la gobernanza algorítmica y sus numerosas áreas de interés en una tipología sistemática. En cualquier caso, las tipologías siempre están supeditadas a las prioridades y los motivos de quienes las desarrollan y su percepción del fenómeno. Aun así, existe una creciente demanda por parte de las personas responsables de la elaboración de políticas en todo el mundo y por parte del público en general para evaluar el despliegue de sistemas de gobernanza algorítmica y guiar su desarrollo en el futuro. Los motivos para ello son buenos; la gobernanza algorítmica, como otros sistemas sociotécnicos, depende de fuerzas sociales, políticas y económicas y puede adoptar distintas formas.

Por todas estas razones, presentamos una tipificación que atiende al diseño y la funcionalidad de los sistemas algorítmicos y los evalúa según criterios normativos clave. A pesar del carácter dinámico del campo, hemos elegido el grado de automatización y transparencia porque se destacan en lo que toca a sus implicaciones normativas en materia de rendición de cuentas y democracia y, por lo tanto, lo más probable es que se mantengan como elementos clave en las evaluaciones de distintos tipos de gobernanza algorítmica que

se realicen en el futuro^[2].

La *transparencia* es importante en cuanto que constituye una piedra angular de la democracia y la autodeterminación (Passig, 2017), pero se ve particularmente desafiada de cara a la complejidad inherente de los sistemas algorítmicos. Por ende, la transparencia no solo es uno de los temas académicos de mayor importancia cuando de regulación algorítmica se trata, sino también una cuestión general importante en la controversia pública (Hansen y Flyverbom, 2015). Solo la transparencia (en un cierto grado) abre los sistemas de toma de decisiones y sus normas sociales inscritas al escrutinio, la deliberación y el cambio. La transparencia, por ende, es un elemento importante de la legitimidad democrática. También cabe notar, empero, que la evaluación de un caso dado de gobernanza algorítmica diferirá entre los desarrolladores de software, el público y los organismos supervisores. Como ya se ha mencionado (cf. sección 3), en comparación con constelaciones de gobernanza anteriores, los sistemas de gobernanza algorítmica desafían las fronteras informativas: exigen formalización, por lo cual las normas sociales y los intereses organizacionales deben ser explicados y traducidos al código, lo cual incrementa *en potencia* la porción de información socialmente observable. Sin embargo, *en la práctica*, la gobernanza algorítmica a menudo viene acompañada por un decremento real en la información socialmente inteligible y accesible, debido a los límites cognitivos (la inteligibilidad del aprendizaje automático) y las barreras sistémicas (la falta de acceso a los algoritmos debida al secreto comercial, cuestiones de seguridad y la protección de la privacidad) (Ananny y Crawford, 2017; Pasquale, 2015; Wachter, Mittelstadt y Floridi, 2017).

El *grado de automatización* importa mucho porque la legitimidad de los regímenes de gobernanza depende de la responsabilidad y la rendición de cuentas de una persona que toma decisiones en su calidad de profesional (un juez, doctora, periodista) y sujeto ético. Enfocarse en el grado de automatización marca también la opción de problematizar la compleja interacción dentro de sistemas sociotécnicos: los sistemas algorítmicos pueden conceder más o menos autonomía a las personas que toman decisiones. Aquí reducimos la escala gradual de compromiso a la distinción binaria entre sistemas totalmente automatizados, donde las decisiones no son supervisadas por un operador humano, y sistemas de recomendación donde las personas que los operan ejecutan o aprueban las decisiones (“humanos que están al tanto”) (Christin, 2017; Kroes y Verbeek, 2014; Yeung, 2018).

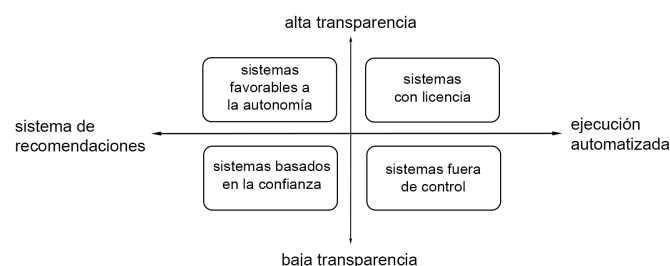


FIGURA 1: Tipos de sistemas de gobernanza algorítmica.

La combinación de ambas dimensiones arroja cuatro tipos ideales [en el sentido weberiano] de sistemas de gobernanza algorítmica con distintas características: “sistemas favorables a la autonomía”, que proporcionan una alta transparencia y dejan las decisiones a los humanos; “sistemas basados en la confianza”, que operan con una baja transparencia y con humanos que toman decisiones; “sistemas con licencia”, que combinan alta transparencia con una ejecución automatizada; y finalmente “sistemas fuera de control”, que demuestran una baja transparencia y ejecutan decisiones de manera totalmente automatizada.

5. GOBERNANZA ALGORÍTMICA EN FUNCIONAMIENTO: POLÍTICA PREDICTIVA Y MODERACIÓN AUTOMATIZADA DE CONTENIDOS

Los cuatro tipos ideales pueden encontrarse en toda la gama de sectores y ámbitos que emplean sistemas de gobernanza algorítmica en la actualidad (véase AlgorithmWatch y Bertelsmann Stiftung, 2019; Arora, 2019; Dencik, Hintz, Redden y Warne, 2018; Tactical Tech, 2019 para revisiones recientes). Para ilustrar la gobernanza algorítmica tanto en el sector privado como en el público y en las plataformas, presentamos brevemente dos casos prominentes y controvertidos: la evaluación de riesgo automatizada para la vigilancia policial (“vigilancia policial predictiva”) está entre las formas más difundidas de gobernanza algorítmica pública en los países industrializados; y la moderación automatizada de contenidos en las plataformas de redes sociales se cuenta entre las diversas maneras en que las plataformas privadas utilizan la gobernanza algorítmica en una escala global. Los casos muestran que la gobernanza algorítmica no es una sola cosa, sino que toma distintas formas en distintas jurisdicciones y contextos, y que está conformada por intereses, poder y resistencia. La gobernanza algorítmica es múltiple, contingente y disputada.

VIGILANCIA POLICIAL PREDICTIVA

Las autoridades policiales emplean la gobernanza algorítmica al combinar y analizar diversas fuentes de datos con la finalidad de evaluar el riesgo de crímenes y de prevenirlos (esto es, el hurto, el robo de automóviles, la agresión violenta, etcétera). Este análisis de riesgo se dirige tanto a personas como a zonas geográficas; algunos sistemas se enfocan en quienes cometen los delitos, otros en las víctimas potenciales. Los resultados son predicciones de riesgo con las que se manobra para orientar la vigilancia policial. La gobernanza algorítmica puede dirigirse hacia el comportamiento de la ciudadanía, o bien, de las fuerzas policiales. Algunas acciones típicas consisten en la asignación de una mayor presencia policial a ciertas zonas geográficas, la vigilancia de criminales potenciales o la advertencia a las víctimas potenciales.

Es necesario evaluar los grados de *transparencia* desde dos perspectivas: en lo que toca al público y en lo referente a la organización que utiliza la vigilancia policial predictiva. En muchos casos, la recolección de datos, el análisis de datos y las medidas de gobernanza son responsabilidad de las agencias policiales y de las empresas privadas, a menudo en

constelaciones complejas (Egbert, 2019). Algunos proyectos se basan en datos de carácter estrictamente delictivo, mientras otros utilizan datos adicionales, como información sobre el clima, el tráfico, las redes, el consumo y el comportamiento en línea. En la mayoría de los casos, el *software* y sus racionalidades básicas no son públicos. Lo mismo se puede afirmar de los resultados del análisis y su interpretación^[3]. No existe un sistema de vigilancia policial predictivo que permita que los datos y el código sean accesibles al público, de manera que la mayoría de las aplicaciones en dicho espacio son *sistemas basados en la confianza*. En algunos casos, como el estado alemán de Renania del Norte-Westfalia, el *software* ha sido desarrollado por la policía. No es público, pero es un *sistema favorable a la autonomía* desde la perspectiva de la policía. Ésta justifica el elevado grado de opacidad con el argumento de que la transparencia le permitiría a la delincuencia “engañar al sistema” y reduciría la eficacia de la gobernanza algorítmica. La opacidad, sin embargo, dificulta las evaluaciones de los efectos sociales que tiene la gobernanza algorítmica sobre la vigilancia policial. Las principales temas de interés público giran en torno a si la vigilancia policial predictiva refuerza formas ilegítimas de discriminación, si amenaza los valores sociales y si es efectiva y eficiente (Ulbricht, 2018).

En lo que respecta al grado de *automatización*, vale la pena destacar que, en la mayoría de los casos de gobernanza algorítmica para la vigilancia policial, el *software* todavía está diseñado como un sistema de recomendaciones: las personas que lo operan reciben información o recomendaciones generadas por computadora. Su responsabilidad es tomar la decisión final de actuar o no y cómo hacerlo. No obstante, los oficiales de la policía se han quejado de la falta de prudencia al decidir dónde se patrulla (Ratcliffe, Taylor y Fisher, 2019). Otra inquietud es que el personal policial puede no tener la capacidad para tomar una decisión autónoma y rechazar la recomendación generada por el algoritmo (Brayne, 2017), con lo que la vigilancia policial predictiva se convertiría, efectivamente, en un sistema de gobernanza algorítmica *fuera de control o con licencia*. La enorme cantidad de investigaciones y proyectos piloto en este ámbito indican que, en un futuro cercano, el grado de automatización en la gobernanza de la vigilancia policial predictiva y el control de fronteras aumentará de forma considerable.

MODERACIÓN AUTOMATIZADA DE CONTENIDOS EN PLATAFORMAS DE REDES SOCIALES

Otro campo muy importante y controvertido de gobernanza algorítmica en operación es la moderación y regulación (parcialmente) automatizada de contenidos en las plataformas de redes sociales. Dos desarrollos están guiando el giro hacia la IA y los algoritmos en este campo (Gollatz, Beer y Katzenbach, 2018): (a) la cantidad de comunicación y contenido que circula en estas plataformas es tan gigantesca que es difícil imaginar a unos personas moderadoras capaces de lidiar manualmente con todas las publicaciones y otros materiales, filtrándolos según el cumplimiento de la ley pública y las reglas de la plataforma. Al tiempo que las plataformas progresan en la búsqueda de soluciones que guarden la escala de su alcance global, están muy interesadas, desde el punto de vista económico, en encontrar

soluciones técnicas. Esto (b) se ve reforzado por la creciente presión política sobre las plataformas para afrontar cuestiones relacionadas con expresiones que incitan al odio, desinformación y violación de derechos de autor en sus sitios, lo cual ha llevado a una regulación que se dirige en parte hacia la responsabilidad inmediata de las plataformas por la publicación de contenidos ilegales (Helberger, Pierson y Poell, 2019). De esta manera, las plataformas desarrollan, prueban y cada vez con mayor frecuencia ponen en funcionamiento sistemas automatizados que apuntan a identificar el discurso de odio, a cotejar contenido subido con obras protegidas por derechos de autor y a etiquetar las campañas de desinformación (Gillespie 2018; Duarte, Llanso y Loup, 2018).

En lo que respecta a la *transparencia*, plataformas como Facebook, YouTube y Twitter han sido sumamente reservadas en relación con estos procesos, los criterios de decisión y las tecnologías y datos específicos que utilizan. La creciente politización de la moderación de contenidos, empero, ha presionado a las empresas para que aumenten la transparencia en este ámbito, aunque con beneficios limitados. Hoy en día, Facebook, por ejemplo, revela el diseño del proceso general de moderación, así como los criterios de decisión que le subyacen, pero aún es reservada en lo que respecta a los elementos específicos del proceso y los datos detallados acerca de la remoción^[4]. El sistema de YouTube para bloquear o monetizar el contenido protegido por derechos de autor, llamado ContentID, proporciona una base de datos de las obras registradas accesible al público. Los criterios de alto nivel para bloquear contenidos han sido comunicados, pero la crítica argumenta que el sistema bloquea de forma exagerada y masiva contenido legítimo y que YouTube se muestra aún demasiado reservado y poco receptivo sobre los procesos de apelación, lo cual incluye los criterios para definir el uso legítimo e ilegítimo de los contenidos protegidos por derechos de autor (Erickson y Kretschmer, 2019; Klonick, 2018). El Foro Global de Internet Contra el Terrorismo (GIFCT, por sus siglas en inglés), un esfuerzo conjunto de Facebook, Google, Twitter y Microsoft para combatir la difusión de contenido terrorista en línea, alberga una base de datos compartida, pero reservada, de imágenes, video, audio y texto conocidos con contenido terrorista.

En lo que toca a la *automatización*, la mayoría de los sistemas de moderación de contenidos no funcionan de manera completamente automatizada, sino que suelen señalar los contenidos controvertidos para que sean revisados por una persona, esto pese a las afirmaciones de la industria sobre la eficacia de los sistemas de IA. Por ejemplo, Facebook tiene en funcionamiento clasificadores técnicos de discurso de odio que al parecer evalúan toda publicación subida y marcan los elementos considerados ilegítimos para su posterior revisión por personas^[5]. En contraste, ContentID opera por lo general de manera totalmente automatizada, lo que significa que las decisiones son ejecutadas sin intervención humana rutinaria: el contenido subido que concuerda con contenido registrado se bloquea, se monetiza para quien posee los derechos o se tolera según las provisiones de quien se supone posee tales derechos. En el caso del GIFCT, los primeros comunicados de prensa hacían énfasis en que “el contenido que concuerda no será removido de forma automática”

(Facebook Newsroom, 2016). Sin embargo, la respuesta de las plataformas a incidentes importantes como el tiroteo en Christchurch, Nueva Zelandia, y a la propaganda de grandes organizaciones terroristas como ISIS y Al-Qaeda parece indicar ahora que ciertas concordancias del GIFCT se ejecutan y, por tanto, se bloquean de forma automática, sin la intervención de personas que estén al tanto que actúen como moderadoras (sistemas fuera de control) (Gorwa, Binns y Katzenbach, 2019).

Como lo muestran estos ejemplos, la clasificación binaria de la transparencia y la automatización de un sistema dado no siempre resulta fácil de delinear. Pero, hasta hace poco, la mayoría de estas implementaciones de gobernanza algorítmica podían considerarse con todo derecho sistemas fuera de control. La reciente presión política y discursiva sin duda ha orillado a las empresas hacia una mayor transparencia, aunque en nuestra evaluación esto todavía no las califica como sistemas favorable a la autonomía ni sistemas con licencia, dado que todavía carecen de una transparencia significativa.

6. CONCLUSIÓN

El concepto de gobernanza algorítmica engloba un amplio espectro de prácticas sociotécnicas que ordenan y regulan lo social de formas específicas, que van desde la vigilancia policial predictiva hasta la gestión laboral y la moderación de contenidos. Una de las ventajas del concepto es que reúne estos conjuntos diversos de fenómenos, discursos y campos de investigación, y de esta manera contribuye a identificar controversias y desafíos clave de la sociedad digital emergente. Los sesgos y la imparcialidad, la transparencia y la agencia humana son temas importantes que necesitan ser abordados siempre que los sistemas algorítmicos se hallen profundamente integrados en procesos organizacionales, sin importar el sector o su aplicación específica. La gobernanza algorítmica tiene muchos rostros: se ve como ordenamiento, como regulación y modificación de comportamientos, como una forma de gestión, de optimización y de participación. Según el área de estudio, se caracteriza por su inescrutabilidad, por la inscripción de valores e intereses, por la eficiencia y la efectividad, por la asimetría de poder, por la inclusión social, por nuevas exclusiones, por competencia, capacidad de respuesta, participación, cocreación y sobrecarga. Para la mayoría de los observadores, la gobernanza se convierte en algo más poderoso, invasivo y generalizado con la algoritmización y la datificación. Una narrativa distinta destaca que la gobernanza se convierte en algo más inclusivo, receptivo y que permite una mayor diversidad social.

Y, en efecto, la gobernanza algorítmica es múltiple. No sigue un camino puramente funcional, teleológico, que progrese hacia una optimización perenne. Más bien depende de su contexto social, económico y político. Los estudios de caso ilustrativos sobre la vigilancia policial predictiva y la moderación de contenidos muestran que la gobernanza algorítmica puede adoptar múltiples formas, y que cambia sin cesar: a veces se optimiza según los intereses de las empresas, a veces se ve presionada por la regulación y las controversias públicas. Los tipos ideales de gobernanza algorítmica propuestos para fungir como medios

de evaluación constituyen una forma de evaluar estos sistemas según estándares normativos. Elegimos la transparencia y el grado de automatización como criterios clave, lo cual resultó en un espectro de implementación que va desde los sistemas fuera de control hasta los sistemas favorables a la autonomía. Otros criterios de evaluación podrían ser los tipos de datos que se aportan o los modelos de decisión. En cualquier caso, estas formas estructuradas e integradas de pensar sobre la gobernanza algorítmica podrían ayudarnos en el futuro a evaluar sobre una base más sólida qué formas de gobernanza algorítmica son legítimas y apropiadas para qué propósitos y bajo qué condiciones... y dónde podríamos no querer ninguna forma en absoluto de gobernanza algorítmica.

REFERENCIAS

Algorithm Watch. (2019). OpenSCHUFA: The campaign is over, the problems remain - what we expect from SCHUFA and Minister Barley. Consultado en <https://openschufa.de/english/>

Algorithm Watch y Bertelsmann Stiftung. (2019). *Automating Society. Taking Stock of Automated Decision-Making in the EU*. https://algorithmwatch.org/wp-content/uploads/2019/02/Automating_Society_Report_2019.pdf.

Ananny, M. y Crawford, K. (2017). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 33(4), 973-989. doi:[10.1177/1461444816676645](https://doi.org/10.1177/1461444816676645)

Arora, P. (2019). Benign dataveillance? Examining novel data-driven governance systems in India and China. *First Monday*, 24(4). doi:[10.5210/fm.v24i4.9840](https://doi.org/10.5210/fm.v24i4.9840)

Avery, R. B., Brevoort, K. P. y Canner, G. (2012). Does Credit Scoring Produce a Disparate Impact? *Real Estate Economics*, 40(3), S65-S114. doi:[10.1111/j.1540-6229.2012.00348.x](https://doi.org/10.1111/j.1540-6229.2012.00348.x)

Barocas, S. y Selbst, A. D. (2016). Big Data's Disparate Impact. *California Law Review*, 104(3), 671-732. doi:[10.15779/Z38BG31](https://doi.org/10.15779/Z38BG31)

Beer, D. (2019). *The Data Gaze: Capitalism, Power and Perception*. SAGE Publications.

Bennett, C. J. (2017). Voter databases, micro-targeting, and data protection law: Can political parties campaign in Europe as they do in North America? *International Data Privacy Law*, 6(4), 261-275. doi:[10.1093/idpl/ipw021](https://doi.org/10.1093/idpl/ipw021)

Berry, D. M. (2011). *The philosophy of software: code and mediation in the digital age*. Basingstoke, Hampshire; Nueva York: Palgrave Macmillan. doi:[10.1057/9780230306479](https://doi.org/10.1057/9780230306479)

Binns, R. Veale, M., Van Kleek y Shadbolt, M. (2017) Like trainer, like bot? Inheritance of bias in algorithmic content moderation. En G. L. Ciampaglia, A. Mashhadi, y T. Yasseri (eds.), *Social Informatics*, (pp. 405-415). doi: [10.1007/978-3-319-67256-4_32](https://doi.org/10.1007/978-3-319-67256-4_32)

Bijker, W. E. y Law, J. (eds.). (1992). *Shaping Technology/Building Society: Studies in*

Sociotechnical Change. Cambridge, MA: The MIT Press.

Black, J. (2001). Decentring Regulation: Understanding the Role of Regulation and Self-Regulation in a 'Post-Regulatory' World. *Current Legal Problems*, 54(1), 103-146. doi:[10.1093/clp/54.1.103](https://doi.org/10.1093/clp/54.1.103)

Boulianne, S. (2015). Social media use and participation: A meta-analysis of current research. *Information, Communication & Society*, 18(5), 524-538. doi:[10.1080/1369118X.2015.1008542](https://doi.org/10.1080/1369118X.2015.1008542)

Boulianne, S. y Theocharis, Y. (2018). Young People, Digital Media, and Engagement: A Meta-Analysis of Research. *Social Science Computer Review*, 19(1). doi:[10.1177/0894439318814190](https://doi.org/10.1177/0894439318814190)

boyd, d. y Crawford, K. (2012). Critical questions for big data. *Information, Communication y Society*, 15(5), 662-667. doi:[10.1080/1369118X.2012.678878](https://doi.org/10.1080/1369118X.2012.678878)

boyd, d. y Barocas, S. (2017). Engaging the Ethics of Data Science in Practice. *Communications of the ACM*, 60(11), 23-25. doi:[10.1145/3144172](https://doi.org/10.1145/3144172)

Brayne, S. (2017). Big Data Surveillance: The Case of Policing. *American Sociological Review*, 82(5), 977-1008. doi:[10.1177/0003122417725865](https://doi.org/10.1177/0003122417725865)

Brevoort, K. P., Grimm, P. y Kambara, M. (2015). *Data Point: Credit Invisibles* [informe de investigación]. Washington, D.C.: Consumer Financial Protection Bureau. Consultado en <https://www.consumerfinance.gov/data-research/research-reports/data-point-credit-invisibles/>

Caplan, R., y boyd, d. (2018). Isomorphism through algorithms: Institutional dependencies in the case of Facebook. *Big Data & Society*, 5(1). doi:[10.1177/2053951718757253](https://doi.org/10.1177/2053951718757253)

Christin, A. (2017). Algorithms in practice: Comparing web journalism and criminal justice. *Big Data & Society*, 4(2). doi:[10.1177/2053951717718855](https://doi.org/10.1177/2053951717718855)

Couldry, N. y Langer, A. I. (2005). Media Consumption and Public Connection: Toward a Typology of the Dispersed Citizen. *The Communication Review*, 8(2). 237-257. doi:[10.1080/10714420590953325](https://doi.org/10.1080/10714420590953325)

Dencik, L., Hintz, A., Redden, J. y Warne, H. (2018). *Data scores as Governance: Investigating uses of citizen scoring in public services project report* [informe de proyecto]. Cardiff University. Consultado en el sitio web de Open Society Foundations: <http://orca.cf.ac.uk/117517/>

DeVito, M. A. (2017). From Editors to Algorithms. *Digital Journalism*, 5(6), 753-773. doi:[10.1080/21670811.2016.1178592](https://doi.org/10.1080/21670811.2016.1178592)

Duarte, N., Llanso, E. y Loup, A. (2018). *Mixed Messages? The Limits of Automated Social*

- Media Content Analysis*. Informe. Washington, D.C.: Center for Democracy & Technology.
- Egbert, S. (2019). Predictive Policing and the Platformization of Police Work. *Surveillance & Society*, 17(1/2), 83-88. doi:[10.24908/ss.v17i1/2.12920](https://doi.org/10.24908/ss.v17i1/2.12920)
- Ellul, J. (1964). *The technological society*. Nueva York: Alfred A. Knopf.
- Erickson, K. y Kretschmer, M. (2018). "This Video is Unavailable": Analyzing Copyright Takedown of User-Generated Content on YouTube. *JIPITEC*, 9(1). Consultado en <http://www.jipitec.eu/issues/jipitec-9-1-2018/4680>
- Eyert, F., Irgmaier, F. y Ulbricht, L. (2018). Algorithmic social ordering: Towards a conceptual framework. En G. Getzinger (ed.), *Critical Issues in Science, Technology and Society Studies* (pp. 48-57). Consultado en <https://conference.aau.at/event/137/page/6>
- Facebook. (2016). Partnering to Help Curb Spread of Online Terrorist Content [publicación de Blog]. Consultado en el sitio web de la Sala de Prensa de Facebook <https://newsroom.fb.com/news/2016/12/partnering-to-help-curb-spread-of-online-terrorist-content>
- Fourcade, M. y Healy, K. (2016). Seeing like a market. *Socio-Economic Review*, 15(1), 9-29. doi:[10.1093/ser/mww033](https://doi.org/10.1093/ser/mww033)
- Fourcade, M. y Healy, K. (2017). Categories All the Way Down. *Historical Social Research*, 42(1), 286-296. doi:[10.12759/hsr.42.2017.1.286-296](https://doi.org/10.12759/hsr.42.2017.1.286-296)
- Fuller, M. (2008). *Software Studies: A Lexicon*. Cambridge, MA: MIT Press. doi:[10.7551/mitpress/9780262062749.001.0001](https://doi.org/10.7551/mitpress/9780262062749.001.0001)
- Gandy, O. H. (2010). Engaging rational discrimination: exploring reasons for placing regulatory constraints on decision support systems. *Ethics and Information Technology*, 12(1), 29-42. doi:[10.1007/s10676-009-9198-6](https://doi.org/10.1007/s10676-009-9198-6)
- Gerber, C. y Krzywdzinski, M. (2019): Brave New Digital Work? New Forms of Performance Control in Crowdwork. En: Steve P. Vallas y A. Kovalainen (eds.), *Work and Labor in the Digital Age* (pp. 48-57). Bingley: Emerald Publishing. doi: [10.1108/S0277-283320190000033008](https://doi.org/10.1108/S0277-283320190000033008)
- Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. New Haven: Yale University Press.
- Gillespie, T. (2014). The relevance of algorithms. En T. Gillespie, P. J. Boczkowski, y K. A. Foot (eds.), *Media Technologies. Essays on Communication, Materiality, and Society* (pp. 167-193). Cambridge, MA: MIT Press.
- Gollatz, K., Beer, F. y Katzenbach, C. (2018). *The Turn to Artificial Intelligence in Governing Communication Online* [Informe de taller] Berlín: Instituto Alexander von Humboldt para

Internet y la Sociedad. <https://nbn-resolving.org/urn:nbn:de:0168-ssolar-59528-6>

Gorwa, R. (2019). What is platform governance? *Information, Communication & Society*, 22(6). doi:[10.1080/1369118X.2019.1573914](https://doi.org/10.1080/1369118X.2019.1573914)

Gorwa, R., Binns, R. y Katzenbach, C. (2019). Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance. *Big Data & Society*, de próxima publicación.

Gray, M. L. y Suri, S. (2019). *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Boston: Houghton Mifflin Harcourt.

Hacker, P. (2018). Teaching fairness to artificial intelligence: Existing and novel strategies against algorithmic discrimination under EU law. *Common Market Law Review*, 55(4), 1143-1185.

Hansen, H. K. y Flyverbom, M. (2014). The politics of transparency and the calibration of knowledge in the digital age. *Organization*, 22(6), 872-889. doi:[10.1177/1350508414522315](https://doi.org/10.1177/1350508414522315)

Helberger, N., Pierson, J. y Poell, T. (2018). Governing online platforms: From contested to cooperative responsibility. *The Information Society*, 34(1), 1-14. doi:[10.1080/01972243.2017.1391913](https://doi.org/10.1080/01972243.2017.1391913)

Hildebrandt, M. (2015). *Smart Technologies and the End(s) of Law. Novel Entanglements of Law and Technology*. Cheltenham: Edward Elgar.

Hobbes, T. (1909). *Hobbes's Leviathan: reprinted from the edition of 1651*. Oxford: Clarendon Press. Consultado en <https://archive.org/details/hobbessleviathan00hobbuoft>

Hofmann, J., Katzenbach, C. y Gollatz, K. (2016). Between coordination and regulation: Finding the governance in Internet governance. *New Media & Society*, 19(9). doi:[10.1177/1461444816639975](https://doi.org/10.1177/1461444816639975)

Introna, L. D. (2016). Algorithms, Governance, and Governmentality: On Governing Academic Writing. *Science, Technology, & Human Values*, 41(1), 17-49. doi:[10.1177/0162243915587360](https://doi.org/10.1177/0162243915587360)

Jarke, J. y Gerhard, U. (2018). Using Probes for Sharing (Tacit) Knowing in Participatory Design: Facilitating Perspective Making and Perspective Taking. *i-com*, 17(2), 137-152. doi:[10.1515/icom-2018-0014](https://doi.org/10.1515/icom-2018-0014)

Danaher, J. Hogan, M. J., Noone, C., Kennedy, R., Behan, B., de Paor, A. ... Shankar, K. (2017). Algorithmic governance: Developing a research agenda through the power of collective intelligence. *Big Data & Society*, 4(2). doi: [10.1177/2053951717726554](https://doi.org/10.1177/2053951717726554)

Just, N. y Latzer, M. (2016). Governance by algorithms: Reality construction by algorithmic selection on the Internet. *Media, Culture & Society*, 39(2), 238-258.

doi:[10.1177/0163443716643157](https://doi.org/10.1177/0163443716643157)

Katzenbach, C. (2017). *Die Regeln digitaler Kommunikation. Governance zwischen Norm, Diskurs und Technik* [Las reglas de la comunicación digital. La gobernanza entre la norma, el discurso y la tecnología]. Wiesbaden: Springer VS. doi:[10.1007/978-3-658-19337-9](https://doi.org/10.1007/978-3-658-19337-9)

Katzenbach, C. (2012). Technologies as Institutions: Rethinking the Role of Technology in Media Governance Constellations. En N. Just y M. Puppis (eds.), *Trends in Communication Policy Research: New Theories, New Methods, New Subjects* (pp. 117-138). Bristol: Intellect.

Kelty, C. M. (2017). Too Much Democracy in All the Wrong Places: Toward a Grammar of Participation. *Current Anthropology*, 58(S15), S77-S90. doi:[10.1086/688705](https://doi.org/10.1086/688705)

Kitchin, R. y Dodge, M. (2011). *Code/Space: Software in Everyday Life*. Cambridge, MA: MIT Press.

Kitchin, R. (2013). Big data and human geography: Opportunities, challenges and risks. *Dialogues in Human Geography*, 3(3), 262-267. doi:[10.1177/2043820613513388](https://doi.org/10.1177/2043820613513388)

Kitchin, R. (2016). Thinking critically about and researching algorithms. *Information, Communication & Society*, 20(1), 14-29. doi:[10.1080/1369118X.2016.1154087](https://doi.org/10.1080/1369118X.2016.1154087)

Klonick, K. (2018). The New Governors: The People, Rules, and Processes Governing Online Speech. *Harvard Law Review*, 131, 1598-1670. Consultado en <https://harvardlawreview.org/2018/04/the-new-governors-the-people-rules-and-processes-governing-online-speech/>

König, P. D. (2019). Dissecting the Algorithmic Leviathan. On the Socio-Political Anatomy of Algorithmic Governance. *Philosophy & Technology*. doi:[10.1007/s13347-019-00363-w](https://doi.org/10.1007/s13347-019-00363-w)

Kroes, P. y Verbeek, P.-P. (2014). Introduction: The Moral Status of Technical Artefacts. En P. Kroes y P.-P. Verbeek (eds.), *Philosophy of Engineering and Technology. The Moral Status of Technical Artefacts* (pp. 1-9). Dordrecht: Springer. doi:[10.1007/978-94-007-7914-3_1](https://doi.org/10.1007/978-94-007-7914-3_1)

Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford; Nueva York: Oxford University Press.

Lee, J. D. y See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, 46(1), 50-80. doi:[10.1518/hfes.46.1.50_30392](https://doi.org/10.1518/hfes.46.1.50_30392)

Lee, C. P., Poltrock, S., Barkhuus, L., Borges, M. y Kellogg, W. (eds.) 2017. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17*. Nueva York: ACM Press.

Lupton, D. (2016). Personal Data Practices in the Age of Lively Data. En J. Daniels, K. Gregory y T. M. Cottom (eds.). *Digital sociologies*. Bristol; Chicago: Policy Press

(pp. 339-354).

Lyon, D. (2014). Surveillance, Snowden, and Big Data: Capacities, consequences, critique. *Big Data & Society*, 1(2). doi:[10.1177/2053951714541861](https://doi.org/10.1177/2053951714541861)

MacKenzie, D. A. (2006). *An engine, not a camera: How financial models shape markets*. Cambridge, MA: MIT Press.

Mejias, U. y Couldry, N. (2019) Datafication. *Internet Policy Review*, 8(4). doi:[10.14763/2019.4.1428](https://doi.org/10.14763/2019.4.1428)

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S. y Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2). doi:[10.1177/2053951716679679](https://doi.org/10.1177/2053951716679679)

Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism*. Nueva York: Public Affairs.

Musiani, F. (2013). Governance by algorithms. *Internet Policy Review*, 2(3). doi:[10.14763/2013.3.188](https://doi.org/10.14763/2013.3.188).

Müller-Birn, C., Herbsleb, J. y Dobusch, L. (2013). Work-to-Rule: The Emergence of Algorithmic Governance in Wikipedia. *Proceedings of the 6th International Conference on Communities and Technologies*, 80-89. doi:[10.1145/2482991.2482999](https://doi.org/10.1145/2482991.2482999)

Napoli, P. M. (2013). *The Algorithm as Institution: Toward a Theoretical Framework for Automated Media Production and Consumption* [documento de trabajo núm. 26]. Nueva York: Centro McGannon, Universidad de Fordham. Consultado en https://fordham.bepress.com/mcgannon_working_papers/26

Neyland, D. y Möllers, N. (2017). Algorithmic IF ... THEN rules and the conditions and consequences of power. *Information, Communication & Society*, 20(1), 45-62. doi:[10.1080/1369118X.2016.1156141](https://doi.org/10.1080/1369118X.2016.1156141)

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. Nueva York: NYU Press.

OCDE. (2015). *Data-Driven Innovation: Big Data for Growth and Well-Being*. París: Publicaciones de la OCDE. doi:[10.1787/9789264229358-en](https://doi.org/10.1787/9789264229358-en) .

Ong, W. J. (1982). *Orality and literacy: the technologizing of the word*. Londres; Nueva York: Methuen.

O'Reilly, T. (2013). Open Data and Algorithmic Regulation. En B. Goldstein y L. Dyson (eds.), *Beyond transparency: Open data and the future of civic innovation* (pp. 289-300). San Francisco: Code for America Press.

Orwat, C., Raabe, O., Buchmann, E., Anandasivam, A., Freytag, J.-C., Helberger, N., ... Werle, R. (2010). Software als Institution und ihre Gestaltbarkeit [El software como

institución y su capacidad de diseño]. *Informatik Spektrum*, 33(6), 626-633.
doi:[10.1007/s00287-009-0404-z](https://doi.org/10.1007/s00287-009-0404-z)

Parasuraman, R., Manzey, D. H. (2010). Complacency and Bias in Human Use of Automation: An Attentional Integration. *Human Factors*, 52(3), 381-410.
doi:[10.1177/0018720810376055](https://doi.org/10.1177/0018720810376055)

Pasquale, F. (2015). *The black box society: the secret algorithms that control money and information*. Cambridge, MA: Harvard University Press.

Passig, K. (23 de noviembre de 2017). Fünfzig Jahre Black Box [Cincuenta años de la caja negra]. *Merkur*. Consultado en
<https://www.merkur-zeitschrift.de/2017/11/23/fuenfzig-jahre-black-box/>

Ratcliffe, J. H., Taylor, R. B. y Fisher, R. (2019). Conflicts and congruencies between predictive policing and the patrol officer's craft. *Policing and Society*.
doi:[10.1080/10439463.2019.1577844](https://doi.org/10.1080/10439463.2019.1577844)

Rieder, G. y Simon, J. (2016), Datatrust: Or, The Political Quest for Numerical Evidence and the Epistemologies of Big Data. *Big Data & Society*, 3(1). doi:[10.1177/2053951716649398](https://doi.org/10.1177/2053951716649398)

Rosenblat, A. (2018). *Uberland: How algorithms are rewriting the rules of work*. Oakland: University of California Press.

Schmidt, A. y Wiegand, M. (2017). A survey on hate speech detection using natural language processing. *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media*. Valencia: Association for Computational Linguistics.
doi:[10.18653/v1/W17-1101](https://doi.org/10.18653/v1/W17-1101)

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*, 4(2). doi:[10.1177/2053951717738104](https://doi.org/10.1177/2053951717738104).

Schrape, J.-F. (2019). The Promise of Technological Decentralization. A Brief Reconstruction. *Society*, 56(1), 31-37. doi:[10.1007/s12115-018-00321-w](https://doi.org/10.1007/s12115-018-00321-w)

Schuepp, W. (14 de septiembre de 2015). Achtung, bei Ihnen droht ein Einbruch [Atención, un robo en su casa es inminente]. *Tagesanzeiger Zürich*.

Suzor, N. P., West, S. M., Quodling, A. y York, J. (2019). What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation. *International Journal of Communication*, 13, 1526-1543. Consultado en
<https://ijoc.org/index.php/ijoc/article/view/9736>

Tactical Tech. (2019). Personal Data: Political Persuasion. Inside the Influence Industry. How it works. Consultado en
https://ourdataourselves.tacticaltech.org/media/Personal-Data-Political-Persuasion-How-it-works_print-friendly.pdf

Ulbricht, L. (2018). When big data meet securitization. Algorithmic regulation with passenger name records. *European Journal for Security Research*, 3(2), 139-161. doi:[10.1007/s41125-018-0030-3](https://doi.org/10.1007/s41125-018-0030-3)

Veale, M. y Brass, I. (2019). Administration by Algorithm? Public Management meets Public Sector Machine Learning. En K. Yeung y M. Lodge (eds.), *Algorithmic Regulation* (pp. 121-149). Oxford: Oxford University Press.

Wachter, S., Mittelstadt, B. y Floridi, L. (2017). Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*, 7(2), 76-99. doi: [10.1093/idpl/ix005](https://doi.org/10.1093/idpl/ix005)

Waseem, Z. y Hovy, D. (2016). Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter. *Proceedings of the NAACL Student Research Workshop*, 88-93. doi:[10.18653/v1/N16-2013](https://doi.org/10.18653/v1/N16-2013).

Wiener, N. (1948). *Cybernetics: or control and communication in the animal and the machine*. Cambridge, MA: MIT Press.

Winner, L. (1980). Do Artifacts Have Politics? *Daedalus*, 109(1), 121-136. Consultado en <http://www.jstor.org/stable/20024652>

Wood, A. J., Graham, M., Lehdonvirta, V. y Hjorth, I. (2019). Good Gig, Bad Gig: Autonomy and Algorithmic Control in the Global Gig Economy. *Work, Employment & Society: a Journal of the British Sociological Association*, 33(1), 56-75. doi:[10.1177/0950017018785616](https://doi.org/10.1177/0950017018785616)

Yeung, K. (2017). „Hypernudge“: Big Data as a mode of regulation by design. *Information Communication & Society*, 20(1), 118-136. doi:[10.1080/1369118X.2016.1186713](https://doi.org/10.1080/1369118X.2016.1186713)

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, 12(4), 505-523. doi:[10.1111/rego.12158](https://doi.org/10.1111/rego.12158)

Ziewitz, M. (2016). Governing Algorithms: Myth, Mess, and Methods. *Science, Technology, & Human Values*, 41(1), 3-16. doi:[10.1177/0162243915608948](https://doi.org/10.1177/0162243915608948)

Zarsky, T. Z. (2014). Understanding Discrimination in the Scored Society. *Washington Law Review*, 89(4). Consultado en <https://digitalcommons.law.uw.edu/wlr/vol89/iss4/10/>

Acerca de los autores

Christian Katzenbach (ZeMKI, University of Bremen): es Profesor de Medios y Comunicación en la Universidad de Bremen y Jefe del Laboratorio “Gobernanza de Plataformas, Medios y Tecnología”. Su trabajo se enfoca en la investigación y docencia sobre plataformas, gobernanza, discursos e inteligencia artificial.

Lena Ulbricht: es doctora en Ciencias Políticas y encabeza un grupo de investigación en el

“Weizenbaum Institute for the Networked Society” y el WZB Berlin Social Science Center. Es coanfitriona del podcast feminista “PurpleCode”. Sus temas son regulación, gobernanza y democracia en sociedades digitales.

Notas

El contexto del estudio son los mecanismos de gobernanza en la producción de contenido de Wikipedia. Los autores definen la gobernanza social como la
^{†1} coordinación que se basa en la comunicación interpersonal y la gobernanza algorítmica como la coordinación basada en reglas que ejecutan los algoritmos (en su mayoría, bots) (Müller-Born et al., 2013, p. 3).

Otras tipologías son demasiado granulares para el objetivo generalizador de este artículo y/o se enfocan en subcampos de gobernanza algorítmica (Danaher et al.,
^{†2} 2017), como la selección algorítmica (Just y Latzer, 2016), la moderación de contenidos (Gorwa, Binns y Katzenbach, 2019) y los modos de regulación (Eyert, Irgmaier y Ulbricht, 2018; Yeung, 2018).

Una excepción es el cantón de Argovia, en Suiza, que publica su propio mapa de
^{†3} riesgos (Schuepp, 2015).

Para un ejemplo, cf. El Informe de Transparencia de Facebook,
^{†4} <https://transparency.facebook.com>, y para una crítica, Suzor et al., 2019.

Cf. Facebook Newsroom, “Using Technology to Remove the Bad Stuff Before It’s
^{†5} Even Reported”, <https://perma.cc/VN5P-7VNU>.