



Revista Iberoamericana de Ciencia, Tecnología y Sociedad
- CTS

ISSN: 1668-0030

ISSN: 1850-0013

secretaria@revistacts.net

Consejo Nacional de Investigaciones Científicas y Técnicas
Argentina

Gómez, Carlos E.

**Estudios críticos sobre algoritmos: ¿un punto de
encuentro entre la ingeniería y las ciencias sociales?**

Revista Iberoamericana de Ciencia, Tecnología y Sociedad
- CTS, vol. 14, núm. 41, 2019, Marzo-Junio, pp. 215-232
Consejo Nacional de Investigaciones Científicas y Técnicas
Argentina

Disponible en: <https://www.redalyc.org/articulo.oa?id=92460273013>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

UAEH  redalyc.org

Sistema de Información Científica Redalyc
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto

**Estudios críticos sobre algoritmos:
¿un punto de encuentro entre la ingeniería y las ciencias sociales?**

**Estudos críticos sobre algoritmos:
um ponto de encontro entre engenharia e ciências sociais?**

***Critical Studies on Algorithms:
A Meeting Point between Engineering and the Social Sciences?***

Carlos E. Gómez *

Estamos en la era del algoritmo o, como señalan algunos expertos, en una “algocracia” donde las matemáticas y las ciencias de la computación se están transformando en un poderoso mecanismo de influencia, conformando y guiando nuestro comportamiento y la gobernanza de la sociedad. La “algocracia”, más allá de nuestra visión o de nuestra voluntad de acción, condiciona cada vez más nuestra existencia, y su uso creciente, al mismo tiempo que nos dota de una poderosa herramienta de conocimiento, nos restringe, manipula, controla y provoca, algunas veces de manera más benigna y otras de forma más riesgosa y problemática. De momento percibimos de manera confusa su efectividad y legitimidad. Las ciencias sociales y la ingeniería tienen un campo de trabajo en común y de dialogo para hacer frente a los nuevos desafíos sociales que plantea la sociedad digital. Este trabajo describe e identifica algunos ejes centrales que se plantean en los estudios críticos sobre algoritmos, sus propiedades fundamentales y los problemas sociales, éticos, políticos y legales producidos o reforzados por su utilización, y las metodologías de las ciencias sociales más aptas para analizarlos.

215

Palabras clave: algoritmos; ciencias sociales; ingeniería; investigación

* Docente investigador del Departamento de Ingeniería Industrial de la Universidad Tecnológica Nacional-Regional San Nicolás (FRSN-UTN) y profesor titular de la Facultad de Trabajo Social de la Universidad Nacional de Entre Ríos (FTS-UNER), Argentina. Correo electrónico: cegomez@frsn.utn.edu.ar.

Estamos na era do algoritmo o, como alguns especialistas apontam, em uma “algocracia” onde a matemática e a ciência da computação estão se transformando em um poderoso mecanismo de influência, moldando e guiando o nosso comportamento e a governança da sociedade. A “algocracia”, além de nossa visão ou nossa vontade de ação, condiciona cada vez mais a nossa existência, e o seu uso crescente, ao mesmo tempo em que nos dota de uma poderosa ferramenta de conhecimento, nos restringe, manipula, controla e provoca, algumas vezes de forma mais benigna e outras de forma mais arriscada e problemática. Na hora percebemos de modo confuso sua efetividade e legitimidade. As ciências sociais e a engenharia têm um campo de trabalho em comum e de diálogo para enfrentar os novos desafios sociais colocados pela sociedade digital. Este trabalho descreve e identifica alguns eixos centrais que surgem nos estudos críticos sobre algoritmos, suas propriedades fundamentais e os problemas sociais, éticos, políticos e legais que são produzidos ou reforçados pelo seu uso, e as metodologias das ciências sociais mais adequadas para analisá-los.

Palavras-chave: algoritmos; ciências sociais; engenharia; pesquisa

We are in the Era of the Algorithm or, as is pointed out by some experts, in an “algocracy” where math and computer sciences are becoming a powerful mechanism that influences and guides our behavior and the governance of society. This “algocracy”, beyond our vision or will for action, more and more conditions our existence. In addition, its increasing use, while providing us with a powerful tool for knowledge, also restricts, manipulates, controls, and provokes us, at times in a benign way and at others in a riskier and more problematic one. At present, our perception confuses its effectiveness and legitimacy. Engineering and the social sciences have a common field of work and dialogue to face the new social challenges presented by the digital society. This paper describes and identifies some lynchpins that are set forth in critical studies on algorithms, their fundamental properties and social, ethical, political and legal issues, which are produced or reinforced by their use, as well as the most suitable social science methodologies to analyze them.

Keywords: algorithms; social sciences; engineering; research

Una aclaración

Este trabajo requiere de una breve presentación aclaratoria, ya que parte de un desconocimiento de las particularidades profundas del objeto que se intenta abordar, al menos en su vertiente de más pura materialidad. Cierta inevitable inseguridad científica al entrar en campos tradicionalmente ajenos a las ciencias sociales, no obstante, no supone desconocer los múltiples y fructíferos vínculos existentes entre tecnólogos y científicos sociales dentro del amplio campo de los estudios sociales de la ciencia y la tecnología. Hablar de estudios críticos sobre algoritmos, en cambio, carece de una adecuada tradición y nos remite sin más a la pura actualidad. Nos transforma en unos profanos que intentan inmiscuirse en disciplinas de una cierta ajenidad.

Acercarse al análisis de un particular campo de conocimiento, sin muchos juicios previos y con gran curiosidad intelectual, constituye un ejercicio estimulante para elaborar esta aproximación a un objeto de estudio con el bagaje de un “inexperto”, al menos sobre las coordenadas profundas que constituyen y conforman el campo de conocimiento de las ciencias matemáticas y de la computación. Este trabajo intenta ser una reflexión fundamentada sobre un área de conocimiento a la que las ciencias sociales, fundamentalmente la sociología y la antropología, comenzaron a acercarse en los últimos años, de la misma forma en que los antropólogos clásicos se acercaban a territorios desconocidos, sobre los cuales tenían escasos conocimientos previos y, por lo tanto, escasas referencias.

Como sugiere un investigador, al señalar las carencias de ciertos enfoques tradicionales en la sociología, es quizá necesario acercarse a determinadas prácticas sociales vinculadas a la tecnología, alejándose un tanto de las prescripciones y convenciones más habituales de las ciencias sociales:

“... la sociología clásica trabaja con criterios que le permiten clasificar a la sociedad en grupos del tipo ‘familia’, ‘pueblo’, o ‘clase’. La sociología futura elaborara criterios que le permitan también clasificar a la sociedad según los tipos de imágenes, por ejemplo en ‘telespectadores’, ‘jugadores de computadoras’ o ‘público de cine’...” (Vilem Fluxe, 2017: 79).

Constituye, entonces, un empeño orientado a conocer algunos interrogantes centrales sobre los cada vez más crecientes debates sobre los algoritmos y su influencia sobre nuestra vida cotidiana, intentando aportar algunas ideas al interrogante central sobre el necesario abordaje de estos temas por tecnólogos y científicos sociales.

Este trabajo no busca desentrañar todas las facetas de naturaleza técnica que inciden en el funcionamiento de los algoritmos, ni cómo funcionan en detalles las matrices de redes neuronales que hace que los algoritmos “aprendan” (Burrell, 2016: 6). Se propone un primer acercamiento orientado a conocer algunos aspectos centrales de su funcionamiento, para entender que operaciones supuestamente

técnicas pueden no serlo demasiado o, en el mejor de los casos, reconocer que estamos sujetos a la lógica decisoria de sistemas aplicados a una variedad de situaciones y problemas complejos que son resueltos, en no pocas oportunidades, de una manera poco afortunada.

Constituye también este trabajo un intento de dotar de sentido al concepto de interdisciplina, tan mentado habitualmente como necesario, pero que no encuentra muchas veces en nuestras tareas de investigación oportunidades claras donde esta pretensión, el ejercicio fructífero del intercambio entre campos diversos, pueda ser transformada en acción sin recurrir a ejercicios meramente retóricos.

Una de las dificultades más habituales para avanzar en el estudio de los algoritmos es la falta de perspectivas compartidas entre los tecnólogos y otras disciplinas, lo que resulta en una baja capacidad aún de teorización sobre estos aspectos, que abarcan de manera creciente diferentes puntos de vista y disciplinas, algunas de manera evidente, como el derecho, la economía, y las ciencias sociales en general. Estos problemas se dan también de manera habitual entre los expertos que componen el ámbito efectivo del campo de la computación y la programación, producto de la separación entre los niveles de dominio de los expertos y los desarrolladores, para responder a los crecientes requerimientos legales y éticos.

Nada desdeñable como problema es la ausencia de un lenguaje común entre académicos, decisores políticos y tecnólogos para encontrar una manera adecuada de reconocer las categorías que utilizan diferentes disciplinas, para explicar un mismo fenómeno y el modo en que diferentes prácticas educativas perpetúan las dificultades de la comunicación interdisciplinaria (Danaher *et al.*, 2017: 14).

Una de las dificultades no menores para crear un puente de colaboración entre disciplinas es, entonces, el conocimiento del lenguaje y de la jerga computacional, y el acuerdo sobre la naturaleza socio-técnica del marco de análisis de un constructo que puede aparecer a simple vista como de irreducible naturaleza técnica. Por lo tanto, este trabajo es un esfuerzo de aprendizaje necesario para comprender la lógica sobre la que se sustenta una tecnología que constituye un sistema conceptual, lógica sin la cual un científico social no podría acercarse para comprender que posibles aéreas de intervención son factibles de ser compatibles, entre las diferentes orientaciones profesionales que conforman la ingeniería, las ciencias de la computación y las ciencias sociales.

Sabemos que las comunidades de expertos y sus límites se sustentan en la capacidad de mantener el control sobre la circulación y el uso apropiado de los *argots* profesionales, demarcando con claridad quiénes ostentan el derecho de hablar y quiénes no, y los algoritmos encuentran un terreno abonado para esta disputa por sus propiedades específicas. Si existe una disciplina en las ciencias sociales con una fuerte tradición para hacer referencia a un encuentro posible entre distintas tradiciones de conocimiento es la antropología, fundamentalmente por la creciente importancia que se da a los enfoques etnográficos para estudiar comunidades profesionales y comportamientos vinculados a las tecnologías digitales.

Los científicos de la computación se preocupan, a diferencia de los científicos sociales, por las cuestiones de eficiencia y por cómo los algoritmos interactúan con estructuras de información y datos. Un antropólogo o un sociólogo, en tanto, se preocupan por develar cómo un algoritmo materializa valores y significados. Ello no significa que sean dos preocupaciones alejadas, sino que están asociadas a diferentes proyectos con distintos objetivos: mientras un científico social está mal situado para entender la eficiencia de un algoritmo, un experto en computación carece de la posibilidad de entender, quizá, el contexto cultural en el que los algoritmos son construidos e implementados.

Estamos en una era del algoritmo o, como señalan algunos expertos, en una “algocracia” donde las matemáticas y la computación se están transformando en un poderoso mecanismo de influencia, conformando y guiando nuestro comportamiento y la gobernanza de la sociedad (Danaher *et al.*, 2017: 1). Desde los tiempos de Max Weber se advertía sobre los peligros de la mecanización y el control, tanto en el Estado como en la sociedad industrial, como consecuencia de las modernizaciones organizativas mediante las estructuras burocráticas y las máquinas. Sin embargo, esto no significa que no estemos ante un desafío muy diferente y de mayor envergadura que el señalado por Weber. La algocracia, más allá de nuestra visión o voluntad de acción, condiciona cada vez más nuestra existencia, y su uso creciente nos restringe, manipula, controla y provoca, algunas veces de manera más benigna y otras de manera más riesgosa.

El reconocimiento de esta realidad no presupone que debamos escribir necesariamente los tópicos clásicos críticos sobre racionalidad, cuantificaciones y procedimientos de estos objetos, y publicar todo lo publicable en formas de manifiestos críticos. Habitualmente suelen coexistir en estos casos puntos de vista de más radicalidad, en cuanto a su crítica y planteamientos sobre las intervenciones posibles sobre la tecnología, que nos ayudan a tener una actitud vigilante y crítica. Ubicamos, no obstante, este trabajo entre los intentos de lograr algún tipo de gobernanza sobre las consecuencias de la utilización de los sistemas de algoritmos, que por el momento percibimos de manera confusa en términos de efectividad y legitimidad.

219

Características de los algoritmos

Algunos definen ya al algoritmo como un nuevo mito moderno al cual se le atribuyen gran significancia y poder, pero cuyas propiedades no están muy bien definidas (Barocas *et al.*, 2013: 1). Para los profanos en el tema los algoritmos son fórmulas matemáticas que nadie entiende. La denominada *big data* hace referencia a macro-datos, datos masivos o datos a gran escala; son *big* en términos de volumen, velocidad y variedad, lo que habitualmente se denomina las “tres v”. Kitchin y McArdle (2016) agregan siete dimensiones particulares para los macro-datos que complejizan el sistema de algoritmos: exhaustividad, resolución, indexicalidad, relacionalidad, extensionalidad y escalabilidad.

En los años recientes, el diseño de algoritmos ha evolucionado de un modelo de “arriba hacia abajo”, en el cual un programador o un equipo de programadores definen exhaustivamente el conjunto de reglas, a uno de “abajo hacia arriba”, lo que supone un sistema en el cual los algoritmos tienen incorporadas reglas de aprendizaje que permiten instalar un orden de reglas propio en las bases de datos a las que entran.

Su propia definición ha supuesto un cambio en su presentación pública: de ser un oscuro término utilizado entre los científicos de la computación pasó a ser un concepto que polariza discusiones y se hace habitual en los medios de comunicación, donde se hace referencia a algunas respuestas estandarizadas y erróneas en diferentes actividades con pretensiones de estandarización, o por el contrario a la idea de sus múltiples beneficios por su supuesto carácter objetivo sobre los sesgados sistemas de decisión humanos. La connotación del término también tiene sus respuestas críticas en los contra-discursos sobre la automatización, el control corporativo y los monopolios mediáticos. En términos generales, en un plano más elemental se asocia a la idea de que en las organizaciones públicas y privadas existen procedimientos internos que no son entendidos por parte de quienes son sujetos de ellos, y estos procedimientos son usualmente descriptos como algoritmos.

Parece claro, entonces, que existen nuevos significados adicionales para definir el concepto de algoritmo en función de sus nuevas aplicaciones en la vida cotidiana. Adicionalmente la polarización de los discursos públicos sobre la temática, muchos de los cuales son nuevos en estos dominios, se vincula fundamentalmente al carácter invasivo de las tecnologías y las técnicas de manipulación de información (*link clicks*), a los movimientos geo-espaciales y a la universalización de la telefonía móvil, los servicios y las aplicaciones de una constante conectividad. Muchas veces, sin embargo, la composición de la información y las preocupaciones nuevas sobre el concepto de privacidad pueden ser más importantes que los propios algoritmos (Burrell, 2016: 2).

Se han desarrollado investigaciones centradas en los aspectos legales y éticos, y en identificar cuatro etapas que conforman lo que se entiende como el “sistema de gobernanza”: recolección, procesamiento, utilización, aprendizaje y *feedback* (Pasquale, 2015). El sistema de gobernanza algorítmica adquiere información, la procesa, la utiliza y aprende de lo que ha realizado. Por lo tanto, un sistema de gobernanza algorítmica funciona como un sistema cuasi-inteligente y adaptativo. No es preciso abundar demasiado sobre los problemas en términos de opacidad y transparencia que estos sistemas generan.

La comprensión y la participación humana en el sistema de gobernanza son un factor determinante de su legitimidad política y social. Paralelamente, el impacto del sistema sobre el comportamiento humano es clave para su evaluación ética y normativa. Citron y Pasquale (2014) han puesto atención en las tipologías de estos sistemas, sobre la base del tipo y grado de involucramiento humano que implican, así han distinguido entre sistemas en los que el ser humano está: a) en el circuito; b) sobre el circuito; o c) fuera del circuito.

Existe una creciente preocupación por los problemas sociales, éticos, políticos y legales que pueden ser producidos y reforzados por estos sistemas. Es cada vez más extendida la producción de estudios sobre los problemas de privacidad y protección de datos. Hay muchos especialistas hablando de la opacidad inherente a la gobernanza algorítmica, particularmente cuando es determinada por el denominado mecanismo con aprendizaje de máquina, o aprendizaje automático, y cuando los artefactos son amparados por una red de leyes de secreto, orientadas a proteger a los algoritmos y a sus propietarios.

Las preocupaciones se extienden a las consecuencias de las inexactitudes, ineficiencias y consecuencias no deseadas del sistema. Zarsky (2016) provee una taxonomía para categorizar y ordenar todos los problemas que se discuten en la literatura especializada. Propone que los sistemas de decisión basados en algoritmos tienen dos propiedades centrales: son potencialmente opacos y pueden ser autómatas. Esta taxonomía define dos ramas o espacios de problemas: espacio de la eficiencia y espacio de la justicia. En el espacio de la eficiencia podemos ubicar aquellos problemas predictivos que obedecen a las dificultades para predecir el comportamiento humano y a la baja calidad de los datos por sus inexactitudes. En el espacio de las objeciones basadas en la justicia tendríamos las injustas transferencias de riquezas: de los consumidores a las empresas, entre los consumidores, lejos de los grupos protegidos; problemas de arbitrariedad, es decir: personas tratadas de manera diferente por razones irrelevantes y daños basados en la autonomía, la falta de consentimiento o la falta de comprensión.

221

La opacidad y la transparencia son, por lo tanto, factores que generan preocupaciones crecientes, tanto desde el campo específico de los usuarios y las organizaciones, como así también desde las ciencias sociales, ya que como disciplina pueden aportar cierta perspectiva crítica dada la naturaleza ambigua de los conceptos puestos en juego y lo cambiante de su significado en diferentes periodos históricos.

La opacidad es una preocupación central en la que suelen detenerse de manera recurrente los estudios críticos sobre los algoritmos, que están sujetos a diversas posibilidades de interpretación (Burrell, 1916). Generalmente, se suele identificar la opacidad como un problema relacionado con las consecuencias sociales de los mecanismos de clasificación y ranking: filtros de *spam*, detección de fraudes con tarjetas de crédito, ingenios de búsqueda, nuevas tendencias, segmentación de mercados y publicidad, criterios para el otorgamiento de créditos, seguros y calificación para préstamos.

Apoyados en algoritmos computacionales, y más específicamente en los mecanismos de aprendizaje de máquina que realizan estas tareas, habituales en el capitalismo avanzado y en las redes de trabajo, estos sistemas de “encasillamiento” son grandes productores de opacidad. Los algoritmos son el mecanismo central de todas las clasificaciones, ya que operan sobre la recepción de información produciendo un resultado, es decir: una clasificación en general vinculada a la posibilidad de acceder a un crédito o a decidir que alguna información debe ser desechada. Las decisiones de clasificación son opacas, ya que no tenemos idea de cómo ese resultado ha sido consecuencia de la información que han manipulado.

Desconocemos también la información que obtienen de nosotros. ¿Cuál es la razón para que desconozcamos esa información? Solemos atribuirlo a que los algoritmos tienen propietarios o son técnicamente ambiguos. No obstante, los fundamentos de la opacidad son más complejos. Hay al menos tres formas de opacidad que podríamos identificar: 1) una de naturaleza intencional que obedece a causas corporativas o de protección institucional; 2) la proveniente de las prácticas de un conocimiento experto donde la escritura y lectura de códigos es una habilidad especializada; y 3) el desfase entre la optimización matemática y las características de los sistemas de aprendizaje automático en relación a las demandas del razonamiento humano y a los estilos de interpretación semántica.

La opacidad existe en un primer nivel por cuestiones de propiedad; los algoritmos son cerrados para mantener las ventajas competitivas y estar unos pasos adelante de la competencia. No obstante, sobre la lógica de estas estrategias hay al menos dos perspectivas. Una sugiere que la opacidad es parte de una estrategia de protección de las empresas y los usuarios, es decir: tienen que ser opacos en aras de ser efectivos. La otra (Pasquale, 2015) señala que la opacidad no es ventajosa en nombre de la manipulación de los consumidores y los usuarios, esconde patrones de discriminación y por lo tanto es una opacidad remediable que puede ser abordada desde las regulaciones, a lo que se propone la figura de un auditor confiable que pueda mantener el necesario secreto y servir a la vez al interés público. Este punto de vista refleja la idea de auditar los códigos y la manera en que se crean, para evaluar las posibles discriminaciones de la clasificación; para ello se hace necesario “entrar” en el algoritmo desafiando su propia impenetrabilidad.

222

En un sentido diferente podemos tratar a la opacidad como propia de una competencia técnica compleja. Mientras los códigos requieren programación en lenguajes específicos y la sintaxis de estos lenguajes debe ser aprendida, ellos son en cierta manera diferentes al lenguaje humano; deben estar adheridos a estrictas reglas lógicas, tener precisión y ser gramaticalmente precisos. Una regla básica de cualquier código es cumplir simultáneamente con la necesidad de ser interpretable por los humanos, el programador original o alguien encargado del mantenimiento del código, y por el dispositivo computacional; por lo tanto, requiere especial exactitud, formalidad y complementariedad, aspectos no necesarios en la comunicación mediante lenguajes humanos.

Finalmente, para agregar más complejidad al concepto de opacidad, podríamos identificar aquella que se origina en la manera en que los sistemas operan: la denominada escala de aplicabilidad. Los expertos hacen notar que los algoritmos son a menudo sistemas de multi-componentes desarrollados por equipos que producen una ambigüedad que los programadores deben contener o limitar.

Hay ciertos desafíos producto de la escala y complejidad distintivas en los algoritmos de aprendizaje de máquina. Están relacionados no simplemente con el número de líneas o páginas de códigos, sino fundamentalmente con el número de miembros de los equipos integrados por ingenieros, y sobre todo a la variedad de inter-vínculos entre módulos o sub-rutinas. Estos no son desafíos sólo de comprensión y lectura de códigos, sino que permiten comprender a los algoritmos en

acción, operando sobre los datos. Los aprendizajes de máquina, en muchos casos, son dependientes de la “dimensión” de la información con que trabajan; billones o trillones de informaciones o datos poseen muchas veces una inabordable complejidad, propia de las heterogéneas propiedades de los datos.

Los vínculos entre los mecanismos de aprendizaje de máquina se hacen complejos y opacos, pese a que los datos tabulados (*dataset*) y los códigos pueden ser escritos con claridad en su interrelación. Los algoritmos de aprendizaje automático son poderosos generalizadores y predictores que funcionan con una cantidad creciente de información, incluyendo dos operaciones paralelas o algoritmos distintivos: un “clasificador” y un “aprendiz”. El clasificador recoge una información de una serie de elementos y produce un resultado, una categoría. Para remitirnos a un ejemplo familiar, los filtros de spam crean categorías de una información como “no deseada” o “deseada”. De manera más preocupante, quizá un sistema de soporte para decisiones de diagnóstico clínico toma una información como síntoma, análisis de sangre, dolor de cabeza, y produce un diagnóstico de hipertensión, cáncer o enfermedad del corazón (Burrell, 2016: 7).

Pero las clasificaciones no son neutras: contienen y sostienen consecuencias sociales. Tomemos el ejemplo sencillo de los correos no deseados; cuando un mensaje legítimo es un denominado “falso positivo” puede ser censurado por el sistema por contener palabras como Nigeria o Ghana y es enviado a la carpeta de no deseados. El sistema hizo sus ponderaciones para establecerlo y determinar su destino. Igualmente ocurre cuando contiene envíos de unas determinadas promesas (garantías), insinúa una autoridad colectiva (nosotros) o especifica una cuantificación de ganancias o beneficios especialmente monetarios (Burrell, 2016: 8).

223

Los ejemplos dados nos proveen de algunos elementos para entender y considerar la lógica interna de los algoritmos y las preocupaciones sobre la falta de justicia y sus efectos discriminatorios. Los científicos de la computación achacan estos efectos a un simple problema de “interpretación”. Pero esto supone imponer un proceso de razonamiento interpretativo humano sobre un proceso de optimización estadística. Un enfoque posible sería construir clasificadores más entendibles para el usuario final, que no impliquen sólo las consecuencias del producto, sino también conocer algo de la lógica de su clasificación.

La continua expansión del “poder algorítmico” ha producido ciertas estrategias de optimización, que exageran este problema particular de opacidad como complejidad de escala (Burrell, 2017: 9). Otros autores (Domingos, 2012) señalan que la intuición falla en depurar o mejorar el algoritmo, se vuelve más difícil, con más cualidades o características provistas como insumo; cada sutil e imperceptible modificación cambia la resultante clasificación. Un alternativa posible podría ser evitar los aprendizajes de máquina en ciertos dominios críticos de aplicación.

Una posible agenda de las ciencias sociales podría orientarse también a observar más profundamente los aprendizajes de máquina y la naturaleza de su opacidad, sumándose al interés sociológico natural en la clasificación y la discriminación relacionadas con la “desigualdad digital”. Se trataría de una concepción orientada a

la distribución desigual de los recursos computacionales y las habilidades, pero que hasta ahora no ha avanzado demasiado en el análisis de cómo las personas pueden estar sujetas a las clasificaciones computacionales, a la invasión de la privacidad y a otras servidumbres posibles, en la manera que se pueden violar las protecciones regulatorias.

En concreto, hay un desfase entre la representación que construyen los mecanismos; su optimización no repara en las formas humanas de construcción semántica, ni siquiera en la de los propios expertos. Últimamente parece haber algunas asociaciones y algunos vínculos entre los académicos de áreas legales, ciencias sociales, expertos en dominios y científicos computacionales que buscan desarrollar estrategias que palien los problemas creados por las clasificaciones.

Los algoritmos como cultura

Un acercamiento posible y necesario también desde las ciencias sociales, especialmente desde una perspectiva antropológica, es analizar los algoritmos como una cultura (Seaver, 1917: 2). Podríamos abordarlos más apropiadamente como artefactos incrustados “en” la cultura, o formando parte de ella. En un sentido diferente, también podríamos observarlos “como” cultura.

Es necesario destacar que diferentes personas en distintos periodos históricos, y en distintas situaciones sociales, han definido a los algoritmos y a sus cualidades sobresalientes de diferentes maneras. Un científico que pertenece a Facebook, un matemático que realiza pruebas en la universidad en 1940 o un médico que establece procedimientos para tratamientos en 1995 podrían asegurarnos, de manera acertada, que trabajan con algoritmos, aunque lógicamente no están hablando de lo mismo. Su definición, sentido y significado están de alguna manera configurados por condicionantes de naturaleza cultural, en el sentido antropológico del término. Es preciso, por lo tanto, tomar cierta distancia de la creencia, a veces extendida, de que existe un sentido unívoco para su definición. En parte podríamos consentir, desde una perspectiva más antropológica, que muchas veces su definición forma parte de un término “émico”; constituye la arquitectura lingüística a través de la cual los ingenieros de sistemas dan sentido a su mundo y a su trabajo.

No obstante, no estamos tratando simplemente con “materiales técnicos”. Los algoritmos están imbricados en la cultura como objetos discretos, que pueden estar localizados dentro de contextos culturales y entran en “conversación” con elementos de esa cultura, alterando los flujos materiales de nuestra cultura y conformándola, y a su vez son simultáneamente determinados por ella. Por lo tanto, los algoritmos afectan a la cultura y son afectados por ella.

En un sentido diferente, analizar a los algoritmos “como” cultura implicaría que, como otros aspectos de la cultura, los algoritmos son transformados en los intercambios entre los ingenieros que los desarrollan y los usuarios; no son objetos técnicos inmutables y estables, sino que son modificados en contacto con actores no técnicos que cambian sus funciones en respuesta tanto a los aprendizajes de

máquina como a los ingenieros, acomodando sus códigos en respuesta a los usuarios.

En síntesis, los algoritmos son una fuerza externa o exógena que transforma la cultura. Un “recomendador” de películas es cultural por su capacidad de influir sobre los flujos de cultura material, que pasaría a ser parte de una cultura del algoritmo. Los algoritmos son cultura en este caso no sólo porque se vinculen con prácticas de consumo popular y colectivo, sino fundamentalmente porque están compuestos de prácticas humanas y colectivas múltiples. Producto de la intersección de diversas prácticas sociales y materiales, culturalmente, históricamente e institucionalmente situadas. Tomarlas como la expresión de un esfuerzo mental, o de manera más abstracta, como procesos de abstracción, es desprenderse de las proximidades y relaciones que los algoritmos articulan (Mackenzie, 2007: 93).

En otras palabras, un algoritmo sucede en un contexto de procesos de ensayo y error, juego, colaboración, discusión y negociación entre diversos actores intervinientes (Kitchin, 2015: 18). Tales prácticas adicionalmente son complementadas por otras: investigaciones sobre el concepto a desarrollar, seleccionar y manipular datos; formación de equipos de codificación; ajustes de parámetros; la venta de una idea y un producto; y la búsqueda de apoyo financiero, entre otros. Todas estas acciones están enmarcadas en formas de pensamiento y conocimiento, culturas organizacionales e institucionales, restricciones legales y subjetividades.

Problemas de una gobernanza algorítmica

225

Es habitual encontrarse en la literatura de ciencias de la computación con el concepto de “gobernanza algorítmica”, término que intenta describir las condiciones del sistema de algoritmos y las posibilidades de establecer alguna forma de arbitraje y control sobre él, fundamentalmente orientado a controlar las debilidades o posibles distorsiones en la constitución de estos modelos que pueden determinar un resultado sesgado, discriminatorio, injusto, o inescrutable. El concepto de gobernanza supone además tener en cuenta a una multiplicidad de actores estatales, sociales y privados que deben involucrarse y articularse para una acción colectiva orientada al diálogo entre los actores. Si bien existe una importante cantidad de expertos que confían y conciben posible esta gobernanza, otros expertos dudan de sus posibilidades y expresan también interrogantes importantes sobre su alcance y posibilidades (Barocas *et al.*, 2013).

Los interrogantes planteados alrededor de los principios de gobernanza y su marco de referencia están vinculados al concepto de legitimidad y efectividad y ciertamente tienen un marcado anclaje en el *mainstream* político liberal, distante de los puntos de vista más radicales que desafían esta perspectiva proponiendo, por el contrario, una anti-gobernanza. No habría que descartar ni dar por sentados que sean los únicos abordajes posibles, críticos con el ideario de la gobernanza algorítmica, ni descartar tampoco que la rápida y cambiante naturaleza de las tecnologías obliguen a replantear los presupuestos de esta gobernanza.

Nos centraremos aquí en el marco de sentido de este trabajo. Seguiremos, entonces, con las posibilidades de desarrollar un marco de investigación o una agenda de investigación que contenga una cierta complementariedad entre la ingeniería y las ciencias sociales. El objetivo principal es lograr una mejor comprensión de los procesos por el cual los sistemas de gobernanza son diseñados e implementados. Existen diversos puntos de vista acerca de los aspectos centrales que deben ser analizados. Para Kitchin (2017), el foco de atención debe ser puesto en los problemas de traslado o “traducción” que se producen cuando los objetivos de las decisiones necesitan convertirse en códigos computacionales. Identifica particularmente tres desafíos de investigación fundamentales para quien desee comprender el proceso y seis posibles enfoques metodológicos y métodos que se detallan a continuación.

Cuadro 1. Desafíos de investigación y enfoques metodológicos

Desafíos de investigación

- *Cajas negras*. Los algoritmos son a menudo privados y están bajo el control de compañías y gobiernos, y sus mecanismos precisos suelen estar ocultos a la vista.
- *Heterogeneidad e integración en el contexto*. Los algoritmos son habitualmente creados por grandes equipos, ensamblados desde paquetes de códigos preexistentes e integrados en complejas redes de otros algoritmos.
- *Característica ontogenética y performativa*. Los algoritmos no son estáticos e invariables. Son habitualmente modificados y adaptados en respuesta a interacciones con el usuario; se desarrollan y cambian de manera incontrolable e impredecible.

Enfoques metodológicos y métodos

- *Examen de códigos*. Deconstruir códigos examinando cuidadosamente la documentación, trazar genealogías de algoritmos, examinar cómo la misma tarea es traducida en diferentes lenguajes de código y ejecutados en diferentes plataformas.
- *Producir códigos reflexivamente*. El método auto-etnográfico, es decir: reflexionar sobre cómo se podría convertir un problema en un conjunto de reglas y códigos de manera asociada.
- *Ingeniería inversa*. Seleccionar datos ficticios y analizar las consecuencias bajo varios escenarios (Ejemplo: utilizando los algoritmos *pagerank* de Google o *edgerank* de Facebook), seguir debates entre usuarios, entrevistar a aquellos que intentan “ganarle” al sistema de algoritmos y otros.
- *Entrevistas y etnografías de programadores*. Observar cuidadosamente y entrevistar a miembros de equipos de programación mientras construyen un algoritmo.
- *Desarmar el ensamble socio-técnico*. Análisis discursivo de documentos empresariales, licitaciones de compras, estándares y marcos de referencia legales.
- *Estudiar efectos en el mundo real*. Llevar a cabo experimentos, entrevistas y etnografías de usuarios, y explorar de cualquier otra manera los efectos sociales de los algoritmos.

Fuente: adaptación propia del marco de investigación de Kitchin (2017) publicado por Danaher, Hogan, Noone y otros (2017: 1-21).

En una perspectiva más elaborada colectivamente, una serie de especialistas (Danaher *et al.*, 2017: 7) desarrollan una agenda posible de investigación para derribar las barreras que impidan legitimar una efectiva gobernanza algorítmica. Proponen otros desafíos adicionales aunque concordantes con Kitchin (2017), para identificar lo que denominan doce barreras que impedirían una efectiva y legítima gobernanza algorítmica. Consideramos trascendentes las conclusiones de estos especialistas básicamente del campo de las ciencias de la computación y otras disciplinas sociales, por la importancia que asignan a la investigación interdisciplinaria para lograr una adecuada legitimidad que haga posible una gobernanza efectiva.

Las barreras identificadas serían las siguientes: opacidad de los algoritmos, el papel de los denominados tecno-utópicos, como contraste la influencia de los tecno-pesimistas, la incertidumbre tecnológica, la capacidad y el conocimiento entre tecnólogos, la capacidad entre los servidores públicos y sus representados, la capacidad de los abogados y el sistema legal, la complejidad legal e institucional, el interés comercial y el público, la efectiva gobernanza vs los derechos individuales, la privacidad y el consentimiento informado.

Describiremos brevemente el sentido de cada una de las barreras identificadas, y las posibles metodologías de investigación social que sería pertinente utilizar, destacando el importante papel que podrían tener los enfoques etnográficos para abordarlos:

- *Opacidad de los algoritmos.* Ya hemos señalado anteriormente que una de las preocupaciones más señaladas por los expertos es la potencial opacidad de estos sistemas. La opacidad está vinculada al funcionamiento del sistema algorítmico, a cómo los codificadores programan y desarrollan algoritmos, y cómo podrían ser decodificados, sugiriendo la posibilidad de investigarlos en tiempo real, mediante filmaciones y estudios etnográficos. Colocando un foco particular en el desarrollo de los sistemas de aprendizaje de máquina y su capacidad de crear reglas de acceso (*ruleset*), e intentando superar la distancia entre el conocimiento del público en general y el de los expertos, en cómo funcionan los sistemas. Se sugieren métodos combinados de obtención de información, como encuestas, estudios de casos, entrevistas, métodos visuales cualitativos y etnografías destinadas a evaluar el impacto que estos sistemas tienen en comunidades y grupo diversos, complementariamente a la posibilidad de crear un índice de transparencia algorítmica, para incidir en los sistemas de gobernanza algorítmica.

- *Tecno-utopismo.* También denominado tecno-optimismo, implica una mirada ingenua y fascinada por parte de las partes interesadas de la comunidad, los políticos y los tecnólogos. La ausencia de reflexión acerca de la expansión de los sistemas de algoritmos, sin considerar sus potenciales efectos negativos. Es decir: la creencia en el carácter virtuoso sin más de la tecnología, en cómo este sistema de creencias debilita un adecuado debate social sobre las consecuencias de su utilización, e impide ver más allá del impacto positivo de la automatización. Sería necesario crear también un sistema de medida y una efectiva definición de tecno-utopismo, analizando la visión sobre la tecnología de parte de los codificadores, desarrolladores y actores gubernamentales relevantes. Las metodologías más adecuadas de relevamiento serían estudios de caso de equipos de desarrollo y la utilización de estrategias observacionales.

• *Tecno-pesimismo*. Por contraste con la limitación anterior, una mirada distante o pesimista de la tecnología constituye también una barrera. A lo largo de la historia ha habido reacciones de resistencia a la incorporación y a los avances de nuevas tecnologías; no obstante, es preciso delimitar el alcance de este posible tecno-pesimismo y lograr un entendimiento común sobre lo que constituye este tecno-pesimismo o desafección a la tecnología, si sólo está orientado a una tecnología específica como la utilización de algoritmos o se extiende a la tecnología de manera genérica, discriminando adecuadamente sus causas, entre otras posibles: exageraciones en la protección de datos, la edad de las personas, la velocidad de su aparición, de sus consecuencias como las desventajas que podrían afectar a ciertos sectores de la población por la extensión de este tecno-pesimismo.

• *Incertidumbre tecnológica*. La aparente incertidumbre sobre las consecuencias del desarrollo tecnológico, básicamente la ausencia de certezas sobre su naturaleza, y la falta de perspectiva histórica sobre sus consecuencias y los cambios que se producen entre su desarrollo y la consecuente implementación. Desde un punto de vista subjetivo, la incertidumbre tecnológica parece sujeta a diferentes percepciones, atribuible a una falta de comprensión de la tecnología o a una falta de interés entre el público general y los decisores políticos. Plantea entonces un problema que se acentúa en la utilización de algoritmos en el sector público; es fundamentalmente preocupante, ya que el propio Estado contribuye a propagar la incertidumbre. Sería valorable, entonces, investigar los sistemas de gobiernos digitales (e-gobierno) mediante metodologías etnográficas para tener información de primera mano sobre los vínculos entre los decisores políticos y los expertos, y sobre cómo son desarrollados los sistemas.

• *Capacidad/conocimiento de los tecnólogos*. Los expertos tecnológicos ejercen un gran poder en el diseño efectivo y en la legitimidad de los sistemas, pero existen dudas sobre su grado de conciencia acerca de las consecuencias legales y de distinta naturaleza de otros sistemas con los cuales interactúan. Los algoritmos no pueden ser formulados ni diseñados aisladamente bajo el predominio de una lógica computacional; están condicionados por otras lógicas de conocimiento, legales, institucionales, de mercado o financieras, y fundamentalmente con personas bajo muy variadas circunstancias. Constituyen un ensamblaje socio-técnico; es necesario examinar de forma cuidadosa su funcionamiento en el mundo real y sus vinculaciones con otros actores clave en su implementación. No tener en cuenta su dimensión socio-técnica puede generar hostilidades en aquellos actores que no poseen el grado de conocimiento de los tecnólogos. Cabe entonces preguntarse si los expertos no desarrollan su propia cultura profesional y un pensamiento grupal, que origina distancias y resistencias críticas con los profanos externos a sus modelos mentales. Desarrollar un acercamiento sobre sus modos de pensar y actuar constituye un desafío que podría ser implementado mediante estudios de caso, encuestas y entrevistas. Particularmente importante podría ser conocer el grado de conocimientos legales, la predisposición a considerarlos y a incorporarlos en sus procesos a medida que los cambios legales se produzcan, fundamentalmente entre los programadores y codificadores, y su posible tendencia a defenderse de posibles litigios. También es preciso considerar las posibilidades

de un modelo de aprendizaje continuo y permanente para tecnólogos e incorporarlos a sus contenidos curriculares educativos.

- *Ausencia de conocimiento de los servidores públicos.* La falta de conocimiento y competencia técnica de nuestros decisores públicos, sobre todo la ausencia de conocimiento sobre habilidades matemáticas, que podría ser extensiva también a competencias de los sistemas públicos de información y sus organizaciones. Sería conveniente definir qué implica tener una competencia y una capacidad sobre estos problemas, para influir adecuadamente en las decisiones. Los estudios de caso y las narrativas de los actores públicos clave podrían constituir una información importante sobre el tema, derivando en un esquema competencial básico sobre el razonamiento matemático de los sistemas de información. Analizar las diferentes orientaciones de las organizaciones públicas en diferentes países y utilizar los métodos de la teoría del actor-red podría ser una estrategia para dar cuenta de las relaciones entre los individuos y las organizaciones sociales.

- *Capacidad de los juristas y el sistema legal.* Si la gobernanza es posible recientemente mediante las leyes, es muy importante entender la capacidad del sistema legal y sus actores para poder intervenir. Son limitaciones obvias la falta de entrenamiento de los abogados, el conservadurismo natural de los sistemas legales y los vacíos regulatorios que disuaden a las personas a desafiar los posibles resultados negativos de una gobernanza global. Esta tendencia puede modificarse identificando las figuras legales, entrevistando a sus ejecutores tanto como a los reguladores. Abordar el problema de forma creativa también puede aportar un mayor conocimiento sobre este tema, principalmente métodos experimentales como pruebas simuladas, lo mismo que analizar los códigos existentes y categorizar los errores que se originan en esos códigos, de acuerdo a si son de naturaleza técnica o legal, mediante estudios de casos que combinen entrevistas y análisis documental.

- *Complejidad legal e institucional.* Las complejidades de los sistemas burocrático-legales que implementan sistemas de gobernanza algorítmica constituyen también una barrera. En especial cómo son adoptados e implantados los sistemas de algoritmos en las estructuras burocráticas. Las diferencias entre lo previsto y lo logrado deberían ser analizadas. También se debería poder evaluar cómo fueron pensados esos sistemas y cuáles fueron las consecuencias de su implementación, fundamentalmente en términos de privacidad y protección de datos.

- *Intereses comerciales y públicos.* Constituyen un problema, por un lado, la falta de equilibrio entre los intereses privados orientados a obtener beneficios por un lado y, por el otro, el interés público socialmente orientado. El desarrollo y la naturaleza comercial de los sistemas de decisión algorítmica, y su más limitada extensión en los sistemas públicos, constituyen una asimetría. Las organizaciones privadas son reacias a los desarrollos más lentos, en aras de asegurar su efectividad y éxito comercial. Este problema se hace más patente en los procesos de privatización del sector de salud pública en algunos países. Las políticas de e-gobierno pueden exponer de una manera más cruda las distintas percepciones

entre los departamentos de tecnologías de la información de carácter público y los consultores privados del sector. Unas variantes de relevamiento de datos secundarios pueden aportar claridad, como los estudios de casos, los más tradicionales análisis legales y económicos, teniendo en cuenta el estudio de las prácticas denominadas “privacidad desde el diseño” (*privacy by design*).

- *Gobernanza efectiva versus derechos individuales.* El foco de consideración de las tecnologías de toma de decisiones basadas en algoritmos debería considerar, más que una efectiva gobernanza, los derechos individuales y el trato justo. Se debería prestar atención a las desigualdades y parcialidades arraigadas en los algoritmos. Debería ser tenida en cuenta la tensión entre efectividad y justicia en los sistemas de gobernanza, subrayando el papel de lo público, preferentemente cómo la tecnología puede fortalecer la transparencia y la participación ciudadana. El seguimiento de las tecnologías denominadas “cadenas de bloques” (*blockchain*), como un mecanismo de gobernanza alternativo para incrementar la transparencia y la participación, puede ser una opción según los expertos.

- *Ausencia de conciencia ética.* Una de las barreras más evidentes para desarrollar una conciencia ética verdadera en el trabajo con algoritmos es la incapacidad de incorporar en la educación tecnológica una adecuada reflexión y un pensamiento ético. La naturalización de la suposición de su carácter neutral constituye un serio problema para enfrentarse a las dimensiones ético políticas de la *big data*. El nivel de conocimiento entre los codificadores, los decisores políticos y el público en general debería ser evaluado y comparado con otros dominios análogos, como la ética médica en el campo de la medicina. Particularmente importante sería relevar las consecuencias éticas de una gobernanza algorítmica en las políticas predictivas y en los perfiles, en los cuales se ocultan sesgos de naturaleza étnica y social que se perpetúan en las estructuras sociales.

- *Privacidad y consentimiento informado.* En relación a la privacidad y el consentimiento informado, son inadecuadas las protecciones a la privacidad, la protección de los derechos humanos, la disonancia entre los sistemas de algoritmos y los marcos regulatorios y legislativos, los conflictos entre el interés público y privado, y el fracaso en informar a la ciudadanía de los múltiples usos de los datos y sus límites. Sería necesario explorar la concepción pública de lo que conocemos como consentimiento informado, el grado de conocimiento por parte de los usuarios de estos mecanismos y su utilización, tanto como de los algoritmos. Las metodologías que más se adecuarían serían las encuestas, las entrevistas y los estudios de laboratorio. Estudios de caso podrían utilizarse para inferir las actitudes públicas sobre la conveniencia *versus* la protección de derechos, constatando si se prefiere sacrificar la privacidad en aras de unos servicios más baratos, y adicionalmente si el concepto de privacidad debe ser redefinido, evaluando posibles nuevos estándares.

Conclusiones finales

Es preciso señalar, como anticipamos en los ejes anteriormente reseñados, la importancia de diversas herramientas de investigación pertenecientes a las ciencias sociales para avanzar en una agenda de investigación. Un adecuado trabajo entre las distintas disciplinas puede ser una buena oportunidad para el trabajo conjunto de ingenieros y científicos sociales.

Particularmente importante es la revalorización de la etnografía tradicional como metodología de larga tradición en las ciencias antropológicas. El método etnográfico parece ser un enfoque metodológico apto para entender de manera más detallada y precisa algunos aspectos micro del trabajo tecnológico: la interrelación de las personas y las computadoras, la producción local de representaciones abstractas y las dotaciones de sentido humano arraigadas en las operaciones de creación de algoritmos. Finalmente, y como consecuencia de este intercambio, se podrá reforzar la naturaleza socio-técnica de los algoritmos, la ubicación de todas estas prácticas dentro de un contexto socio-cultural habitualmente diverso y cambiante.

Aproximarse etnográficamente a los algoritmos supone partir de una cultura constituida no solamente por procedimientos racionales, sino por instituciones, personas y contextos que se interrelacionan, casi siempre provisionales, determinados por sentidos que se obtienen de la cultura cotidiana (Seaver, 2017: 10).

231

Bibliografía

BAROCAS S., HOOD, S. y ZIEWITZ, M. (2013): "Governing algorithms: A provocation piece". Disponible en: <http://governingalgorithms.org/resources/provocation-piece/>. Consultado el 5 de marzo de 2018.

BURRELL, J. (2016): "How the Machine 'Thinks:' Understanding Opacity in Machine Learning Algorithms", *Big Data & Society*, vol. 3, n° 1.

CITRON D., PASQUALE, F. (2014): "The scored society: Due process for automated predictions", *Washington Law Review*, pp. 86-101.

DANAHER, J., HOGAN, M., NOONE, C. *et al.* (2017): "Algorithmic governance: Developing a research agenda through the power of collective intelligence", *Big Data & Society*, julio y diciembre, pp. 1-21.

DOMINGOS, P. (2012): "A few useful things to know about machine learning", *Communications of the ACM* 55, n° 10, pp. 35-78.

FLUSSER, V. (2017): *El universo de las imágenes técnicas. Elogio de la superficialidad*, Buenos Aires, Caja Negra Editora.

KITCHIN, R. (2017): "Thinking critically about and researching algorithms", *Information, Communication & Society*, vol. 20, nº 1, pp. 14-29.

MACKENZIE, A. (2007): "Protocols and the irreducible traces of embodiment: The Viterbi algorithm and the mosaic of machine time", en R. Hassan y R. E. Purser (eds.): *Time and temporality in the network society*, Stanford, Stanford University Press, pp. 89-106.

PASQUALE, F. (2015): "The Black Box Society: The Secret Algorithms that Control Money and Information", Cambridge, Harvard University Press.

SEEVER, N (2017): "Algorithms as culture: Some tactics for the ethnography of algorithmic systems", *Big Data & Society*, julio y diciembre, pp. 1-12.

ZARZKY T. (2016): "The trouble with algorithmic decisions: An analytic road map to examine efficiency and fairness in automated and opaque decision making", *Science, Technology & Human Values*, nº 41, pp.118-132.

Cómo citar este artículo

GÓMEZ, C. E. (2019): "Estudios críticos sobre algoritmos: ¿un punto de encuentro entre la ingeniería y las ciencias sociales?", *Revista Iberoamericana de Ciencia, Tecnología y Sociedad –CTS*, vol. 14, nº 41, pp. 215-232.