

ENSAYO

**LA EDICIÓN DE TEXTOS CIENTÍFICOS VIGILADA:
LAS GRANDES EDITORIALES Y LA MONETIZACIÓN
DE LA INFORMACIÓN DE LOS AUTORES**

Jefferson Pooley

Muhlenberg College, 2400 W. Chew St., Allentown, Pennsylvania, 18104, USA. email: jeffersonpooley@muhlenberg.edu

RESUMEN. Las grandes editoriales del mundo académico basan parcialmente su modelo de negocios en el trabajo no remunerado de los investigadores y en la venta de sus productos derivados a las universidades e instituciones académicas. Así, la labor de revisión y edición de los investigadores se traduce en la venta de estos productos en forma de artículos por separado o suscripciones a precios de usura, considerando además en muchos casos, los cobros por el procesamiento de los artículos (APC, por sus siglas en inglés). Este es un negocio lucrativo al que las grandes editoriales no quieren renunciar. Sin embargo, ahora se añade otro negocio quizá igual o más lucrativo a sus operaciones editoriales: la riqueza detrás de la información propia de los autores o de los datos sobre el comportamiento de estos, contenidos en las publicaciones que poseen o controlan. Las grandes editoriales y corporativos asociados como Clarivate, Elsevier, Springer Nature, Wiley, Taylor & Francis, SAGE y otras, utilizan el llamado “análisis predictivo”, con el que procesan los datos de la producción científica para analizar el comportamiento de los académicos que luego venden a sus clientes (y proveedores): las propias universidades. Estas editoriales se conducen con los académicos tal como lo hace Google con todos sus usuarios: obtiene datos sobre su comportamiento generando información valiosa sobre sus tendencias, hábitos o preferencias. Google monetiza esta información con las distintas empresas que desean vender sus productos a un público dirigido. De la misma forma, las editoriales venden la información recolectada a las universidades y oficinas de gobierno relacionadas con la política científica para la toma de decisiones. El nuevo negocio del “análisis predictivo” se basa en los datos de comportamiento que producen los académicos. Es decir, los investigadores generan datos con cada participación en un artículo o informe revisado por pares. Algunos de esos datos se introducen en los productos principales de las editoriales, en forma de recuentos de descargas y recomendaciones de artículos. La publicación académica es su propia economía de vigilancia emergente. Podemos denominar a una editorial como “vigilante”, si obtiene una parte sustancial de sus ingresos utilizando el análisis predictivo a partir de la información extraída del comportamiento de los investigadores. Como dijo la Dra. Sarah Lamdan de la Facultad de Derecho de la City University of New York (CUNY School of Law): “...tus revistas te están espiando”.

Pooley, J. 2023. La edición de textos científicos vigilada: las grandes editoriales y la monetización de la información de los autores. *CICIMAR Océánides*, 38(1): 9-18.

EN UN PRINCIPIO

Durante el mes de abril de 1998, dos estudiantes de Stanford, Sergey Brin y Larry Page, volaron por todo el mundo para presentar una ponencia sobre su incipiente motor de búsqueda, Google (Büttcher et al., 2016). En su intervención en la Séptima Conferencia Internacional de la World Wide Web (WWW 98), celebrada en Brisbane (Australia), Brin y Page describieron cómo su enfoque (tomando el “gráfico” de enlaces existente en la web como indicador de calidad y relevancia) mejoraba los índices clasificados a mano de Yahoo!, Lycos y otros motores de búsquedas similares (Brin y Page, 1998). Seis meses después, llevaron su idea al mercado y empezaron a trabajar en un garaje cercano. En dos años, Brin y Page habían despachado a sus rivales de los motores de búsqueda, camino de construir el mayor negocio publicitario de la historia del capitalismo.

La historia del origen de Google es bien conocida. Menos conocida es la deuda de Brin y Page con la biblioteconomía y la bibliometría. Como reconocieron en Brisbane, su idea clave, i.e., utilizar la estructura de enlaces de la web como un plebiscito para determinar la relevancia de las búsquedas, la tomaron prestada del análisis de citas. “El gráfico de citas (enlaces) de la Web es un recurso importante que no se utiliza en los motores de búsqueda actuales” (Brin y Page, 1998). Establecieron que el “PageRank” de una página web era una medida de su “importancia de

Fecha de recepción: 05 de enero 2023

citación”, que resulta coincidir con asombrosa coherencia, con lo que los buscadores quieren encontrar. Su enfoque es una extensión de la “literatura [a]cadémica sobre citas” (Brin y Page, 1998).

Los fundadores de Google se habían basado en la bibliometría, un campo que surgió en la década de 1960 para estudiar (entre otras cosas) la red de citas académicas. Como dijo el historiador de la ciencia Derek de Solla Price en un artículo fundamental de 1965, las citas proporcionan una “red mundial total de artículos científicos” (Price, 1965). A principios de la década de 1970, gracias a los avances informáticos, se empezó a utilizar el análisis completo de citas para medir el impacto de las revistas, la productividad científica y la estructura de la influencia académica.

Dos décadas después, en Brisbane, Brin y Page presentaron a Google como el antídoto académico contra los motores de búsqueda basados en la publicidad. Se quejaron de que empresas como *Yahoo!* no hicieran públicos sus métodos, con el resultado de que la tecnología de búsqueda seguía siendo “en gran medida un arte negro”. Con Google, decían, “tenemos el firme objetivo de impulsar un mayor desarrollo y comprensión en el ámbito académico”. En un apéndice de su charla, ahora famoso (Brin y Page, 2012), los dos estudiantes de posgrado denunciaban el modelo de negocio basado en la publicidad de sus rivales comerciales. Brin y Page escribieron: “Esperamos que los motores de búsqueda financiados por la publicidad

Fecha de aceptación: 16 de febrero 2023

estén inherentemente sesgados hacia los anunciantes y alejados de las necesidades de los consumidores”, un “sesgo particularmente insidioso”, añadieron, ya que es muy difícil de detectar (Brin y Page, 2012). Como se verá adelante, pronto cambiaron de opinión.

Ante el colapso de las empresas “punto.com” en 2001 y las demandas de los inversores, Brin y Page (tomando prestado el verbo de Silicon Valley) dieron un giro. Como ha documentado Shoshana Zuboff, la empresa se lanzó de lleno a los anuncios: *i.e.*, anuncios orientados, informados por el tesoro de datos de los usuarios que la empresa había almacenado (Zuboff, 2019). Para el año 2004, la empresa había salido a bolsa y se encontraba valorada en ese entonces en 27,000 millones de dólares (La Monica, 2004). Aprovechando sus datos de búsqueda y servicios derivados de los usuarios, Google pasó a captar casi el 30% de los ingresos mundiales por publicidad digital (Cramer-Flood, 2021). En la actualidad, el valor de mercado de *Alphabet*, la empresa matriz de Google, ronda los 2 billones de dólares. Construida a partir del análisis de citas académicas, la empresa es el ejemplo definitorio de lo que Zuboff llama “capitalismo de vigilancia”.

LA BIBLIOMETRÍA, LA ALTMETRÍA Y LAS GRANDES EDITORIALES

El campo de la bibliometría, desde su aparición a principios de la década de 1960, ya estaba inmerso en el capitalismo de datos. También en este caso, la historia es bien conocida: Eugene Garfield, un químico en ciernes convertido en empresario científico, creó su empresa de indexación científica, el Instituto de Información Científica (ISI), a mediados de la década de 1950 (Wouters, 2017). En 1964, el ISI de Garfield creó el primer Science Citation Index, una base de datos de artículos publicados y sus citas (Garfield, 2006). Pioneros de la bibliometría como de Solla Price se asociaron con Garfield para explotar la base de datos del servicio, de ahí la red mundial total de artículos científicos de de Solla Price. Otros índices ISI para las ciencias sociales y las artes siguieron en la década de 1970, cuando la empresa de Garfield también comenzó a publicar sus Journal Citation Reports (Baykoucheva, 2019). En 1992, con la World Wide Web en pañales, Garfield vendió ISI a Thomson, el gigante canadiense de la información. La empresa volvió a cambiar de manos en 2016, en una escisión de capital privado llamada Clarivate. El índice de citas de Garfield (ahora llamado Web of Science, o WoS por su acrónimo) fue el centro de la operación de 3,500 millones de dólares.

En aspectos fundamentales, el negocio de Clarivate se parece al de Alphabet. Clarivate, por supuesto, no se alimenta del “parque de bomberos” de la publicidad como Google. Pero ambas empresas extraen datos del comportamiento, que procesan con análisis predictivos. En el caso de Google, todos participamos en la acción, con cada búsqueda y correo electrónico; una vez depurados, los datos se venden a los clien-

tes-anunciantes de la empresa para su visualización selectiva. Los datos de comportamiento de Clarivate se obtienen de un público mucho más reducido, *i.e.*, los académicos, que, a diferencia de Google, también son los principales clientes de la empresa. Pero la estrategia empresarial principal es la misma: extraer datos del comportamiento para alimentar modelos predictivos que, a su vez, se perfeccionan y venden a los clientes. En un caso se trata de términos de búsqueda y en el otro de resúmenes y citas, pero en cualquier caso se trata de extraer dinero de los subproductos del comportamiento (del consumidor o del académico). En lugar de la propensión a comprar de Google, Clarivate vende apuestas sobre la productividad y el impacto futuro de la investigación, entre otros productos de predicción académica.

Este ensayo también se detiene en una predicción: el modelo de negocio de Clarivate está llegando a la edición académica. Google es un competidor, pero los verdaderos competidores de la empresa Clarivate son Elsevier, Springer Nature, Wiley, Taylor & Francis y SAGE. Elsevier, en particular, lleva años introduciéndose en el análisis predictivo (Murakami-Wood, 2009). Por supuesto, los gigantes de la edición se han beneficiado durante mucho tiempo de los académicos y de nuestros empleadores universitarios, empaquetando el trabajo no remunerado de escritura y edición de los académicos para vendérselo de nuevo como suscripciones a precios usureros o cargos por procesamiento de artículos (APC, por sus siglas en inglés). Se trata de un negocio lucrativo al que Elsevier y los demás no renunciarán. Pero están añadiendo otro negocio a sus operaciones editoriales heredadas, al estilo de Clarivate. El tesoro de datos de que disponen las editoriales es, en todo caso, mucho más rico que el gráfico de citas por sí solo (al estilo de Google).

¿Por qué preocuparse por la edición vigilada? Una de las razones es el balance, ya que el comercio de las empresas con futuros académicos, aumentará aún más los beneficios a expensas de los contribuyentes y los estudiantes. La razón más importante es que nuestro comportamiento (una vez alejado de nosotros y abstraído en métricas predictivas) se duplicará en nuestras vidas laborales. Los prejuicios existentes, como la propensión de los académicos a auto-citarse, recibirán una nueva capa de legitimidad algorítmica. En términos más generales, el sistema de recompensas académicas ya está distorsionado por las métricas. En la medida en que las estadísticas y los índices de las editoriales se integran en la concesión de becas, la titularidad y la promoción, y otras decisiones de evaluación, la marea métrica gana poder. El mayor riesgo es que los académicos interioricen una mentalidad analítica, ya fomentada por los recuentos de citas y los factores de impacto.

LA VIGILANCIA COMO SERVICIO

Por muy útil que sea, la noción de “capitalismo de la vigilancia” de Shoshanna Zuboff está demasiado ceñida a un sector relativamente pequeño de la

economía: la publicidad digital. Ese mismo estrecho margen llevó a Zuboff, en *La era del capitalismo de la vigilancia*, a hacer demasiado hincapié en la novedad del negocio de los futuros del comportamiento que atribuye a Google (Zuboff, 2019). Los sectores de los seguros y la calificación crediticia, por mencionar dos, llevan más de cien años vinculando los datos a los beneficios predictivos (Bouk, 2018; Lauer, 2017). Como hemos visto, la ISI de Garfield estaba en el negocio de los datos antes de que nacieran Larry Page y Sergey Brin.

Para llegar al parentesco de los editores con Google o, para el caso, con Hartford, necesitamos un descriptor más amplio. Los juristas Mariano-Florentino Cuéllar y Aziz Huq han propuesto una alternativa pluralizada, “economías de vigilancia”, para referirse a la gama de modelos empresariales que tratan de monetizar los datos sobre el comportamiento (Cuéllar y Huq, 2010). “A medida que más industrias encuentren formas de incorporar los excedentes del comportamiento a sus modelos de negocio”, escriben, “la parte de la economía que se engloba bajo este término aumentará, quizá drásticamente”. Cuéllar y Huq destacan el pluralismo: Los contornos específicos de cualquier economía de la vigilancia variarán en función de las normas y reglamentos específicos de cada sector. En otras palabras, no es necesario llevar demasiado lejos la analogía con Google. Es poco probable que los negocios de datos basados en las citas y descargas de los académicos, emulen el modelo de Google basado en la publicidad. Es más probable que los grandes editores, junto con Clarivate y otros posibles actores, se apoyen en su actual estrategia de suscripción, con productos de datos licenciados a universidades y otros clientes del sector de la investigación. En cualquier caso, se quedarán con el excedente de comportamiento que produzcan los académicos. Como dijo la Dra. Sarah Lamdan de la Facultad de Derecho de la City University of New York (CUNY School of Law): “... tus revistas te están espiando” (Lamdan, 2021).

Las editoriales se encuentran en una posición envidiable, ya que los investigadores generan datos con cada artículo o informe de revisión por pares. Algunos de esos datos se incorporan a los productos principales de las editoriales, como los recuentos de descargas y las recomendaciones de artículos. Pero tenemos motivos para creer, basándonos únicamente en los productos de datos existentes, que los editores están descremando los residuos del comportamiento de los académicos ante la perspectiva de una futura monetización. En un importante artículo reciente, el académico de CTS Jathan Sadowski discrepaba del tópico de que los datos son el “nuevo petróleo” (Sadowski, 2019). En su opinión, los datos son materia prima para otros productos, fáciles de intercambiar por dinero. Los datos son a menudo una mercancía de este tipo, admite Sadowski y el extenso sector del corretaje de datos es un buen ejemplo de ello. Pero también es útil pensar en los datos como capital, en el

sentido específico del término “capital”, desarrollado por el sociólogo francés Pierre Bourdieu. El capital de datos se asemeja en su forma al capital cultural de Bourdieu: aunque una apreciación aprendida del arte abstracto puede, en determinadas condiciones, conducir a un trabajo lucrativo, el valor de ese capital cultural no es meramente, ni siquiera principalmente, monetario. Del mismo modo, el capital de datos puede convertirse en dólares en algunos contextos. Pero su valor para los propietarios puede residir en otra parte. Las empresas pueden utilizar los datos para orientar la estrategia, perfeccionar los flujos de trabajo o entrenar modelos, entre otras cosas. Al igual que el capital social o cultural, la acumulación de datos tiene una cualidad prospectiva: un incentivo para acumularlos con la expectativa de un valor futuro. La publicación académica es su propia economía de vigilancia emergente. Podemos llamar a una empresa editorial de vigilancia si obtiene una proporción sustancial de sus ingresos de productos de predicción, alimentados por datos extraídos del comportamiento de los investigadores. Según esta definición, ya tenemos entre nosotros, editores de vigilancia.

LA EDITORIAL (“ENCHILADA”) COMPLETA

La editorial holandesa Elsevier se fundó a finales del siglo XIX, pero no fue hasta la década de 1970 cuando empezó a lanzar y adquirir revistas a un ritmo frenético. El modelo de Elsevier fue Pergamon, la editorial científica de la posguerra fundada por el audaz checo Robert Maxwell (Buranyi, 2017). En 1965, más o menos cuando apareció por primera vez el *Science Citation Index* de Garfield, Pergamon publicaba 150 revistas. Elsevier siguió los pasos de Maxwell y a finales de los setenta publicaba 35 títulos al año. Ambas empresas subieron agresivamente sus precios de suscripción, obteniendo enormes beneficios de la señalización de prestigio del Factor de Impacto de las Revistas de Garfield. Maxwell vendió Pergamon a Elsevier en 1991, meses antes de su escabrosa muerte (Buranyi, 2017).

Ese mismo año Elsevier adquirió The Lancet y puso a prueba lo que se convertiría en ScienceDirect, su plataforma web de distribución de revistas (Elsevier, 2005; Zijlstra, 1994). En 1993, la editorial holandesa se fusionó con Reed International, un fabricante de papel británico convertido en conglomerado de medios de comunicación. En 2015, la empresa cambió su nombre por el de RELX Group, tras dos décadas de adquisiciones, desinversiones y lanzamientos de productos, incluido Scopus en 2004, la respuesta de Elsevier a la Web of Science de ISI. El nombre, “más corto y moderno”, explica RELX (2016), es un guiño a la “transformación” de la empresa, que ha pasado de ser una editorial a un “negocio impulsado por la tecnología, los contenidos y el análisis”. ¿La estrategia de RELX? El “desarrollo orgánico de herramientas cada vez más sofisticadas de análisis y toma de decisiones basadas en la infor-

mación” (RELX Group, 2016). En otras palabras, Elsevier iba a convertirse en un editor de vigilancia.

Desde entonces, mediante adquisiciones y lanzamientos de productos, Elsevier ha hecho honor a su auto-descripción. Al moverse por todo el ciclo de vida de la investigación, la empresa se ha posicionado para cosechar excedentes de comportamiento en cada etapa (Posada y Chen, 2017; 2018). ¿Seguimiento de los resultados de laboratorio? Elsevier cuenta con Hivebench, adquirida en 2016. ¿Software de citas e intercambio de datos? Mendeley, adquirida en 2013. ¿Publicación de documentos de trabajo o preimpresos? SSRN y bepress, 2016 y 2017, respectivamente.

Las “soluciones” de Elsevier para la fase posterior a la publicación del flujo de trabajo académico se basan en Scopus y sus 81 millones de registros. ¿Siente curiosidad por el impacto? Plum Analytics, una empresa de altmetría, adquirida en 2017. ¿Quiere hacer un seguimiento de los investigadores de su universidad y su trabajo? Existe el “sistema de gestión de la información de investigación” Pure, adquirido en 2012. ¿Medir el rendimiento de los investigadores? SciVal, escisión de Scopus en 2009, que incorpora el servicio de seguimiento de medios Newsflo, adquirido en 2015.

Elsevier, por retomar una expresión informática, es ahora una editorial de “pila completa”. Sus productos abarcan el ciclo de vida de la investigación, desde la mesa de laboratorio hasta la puntuación de impacto, e incluso (mediante las herramientas de búsqueda de subvenciones de Pure) de vuelta a la mesa de laboratorio, para empezar de nuevo. Algunos de sus productos son, podríamos decir, servicios con beneficios: Mendeley, por ejemplo, o incluso la plataforma de distribución de revistas ScienceDirect, ofrecen a los clientes gestión de referencias o acceso a revistas y facilitan datos de comportamiento a Elsevier. Productos como SciVal y Pure, en la parte superior de la cadena de datos, venden los datos procesados a los investigadores y a sus empleadores, en forma de “inteligencia de la investigación”. Incluso el visor de PDF de la empresa, integrado en ScienceDirect y otros productos, extrae información detallada sobre los lectores (Anónimo, 2023 a).

Facebook, Google y ByteDance, por ejemplo, tienen que regalar sus servicios de cara al consumidor para atraer a los usuarios que producen datos. Si no pagas por ello, dice el adagio de Silicon Valley, entonces eres el producto. Para Elsevier y sus homólogos, en cambio, es un buen negocio: somos el producto y estamos pagando (mucho) por él. De hecho, es probable que los beneficios inesperados de las suscripciones y el APC en el negocio editorial “heredado” de Elsevier hayan financiado su “juerga de adquisiciones” de analítica durante una década. Como tuiteó recientemente Björn Brembs: “el sobrepago masivo de las editoriales académicas les ha permitido comprar tecnología de vigilancia que cubre todo el flujo de trabajo y que se puede utilizar no sólo para combinarla

con nuestros datos privados y venderla, sino también para tomar decisiones de empleo algorítmicas (a.k.a. ‘evidence-led’)” (Brembs, 2021). Esto es un insulto sobre insulto: Desplumarnos una vez para volver a desplumarnos de nuevo, primero en la biblioteca y luego en la oficina de evaluación.

Los productos de predicción de Elsevier clasifican y procesan los datos extraídos de diversas maneras. La empresa promociona lo que denomina Fingerprint Engine, que aplica técnicas de aprendizaje automático a un océano de textos académicos: resúmenes de artículos, sí, pero también patentes, anuncios de financiación y propuestas (Elsevier, n.f. a). El modelo, presumiblemente entrenado con ejemplos codificados por humanos (¿palabras clave de artículos designados por académicos?), asigna palabras clave (por ejemplo, “Resistencia a los medicamentos”) a los documentos, junto con lo que equivale a una puntuación ponderada (por ejemplo, 73%). La lista de términos y puntuaciones es, según la empresa, una “huella digital”. El motor se utiliza en diversos productos, como Expert Lookup (para encontrar revisores), JournalFinder de la empresa y su software Pure, de gestión de la investigación a nivel universitario. En este último caso, son los académicos los que obtienen la huella digital: Pure aplica tecnología semántica y 10 vocabularios de palabras clave específicos de la investigación para analizar las publicaciones y las concesiones de subvenciones de un investigador y transformarlas en un Fingerprint™ único: un índice visual diferenciado de conceptos y una lista ponderada de términos estructurados.

Las técnicas de aprendizaje automático que utiliza Elsevier coinciden con otras actividades de análisis predictivo de RELX dirigidas a clientes corporativos y jurídicos, como LexisNexis Risk Solutions. Aunque RELX no proporciona cifras de ingresos específicas para sus productos de predicción académica, las revelaciones de la empresa a la SEC en 2020 indican que más de un tercio de los ingresos de Elsevier proceden de bases de datos y productos de referencia electrónica, un negocio, afirma la empresa, en el que “seguimos impulsando un buen crecimiento a través del desarrollo de contenidos y la mejora del aprendizaje automático y la funcionalidad basada en el procesamiento del lenguaje natural” (Elsevier, n.f. b).

Muchos de los rivales de Elsevier también parecen estar lanzándose al mercado de la analítica con una estrategia similar de recopilación de datos de toda la pila de investigación. Taylor & Francis, por ejemplo, es una unidad de Informa, un conglomerado con sede en el Reino Unido cuyas raíces se remontan a Lloyd’s List, la revista de inteligencia marítima del siglo XVIII. En su informe anual de 2020 la empresa escribió que tiene la intención de “utilizar y analizar más profundamente los datos de primera parte” que se encuentran en Taylor & Francis y otras divisiones, para “desarrollar nuevos servicios basados en datos duros y conocimientos de datos de comportamiento”

(Informa Group, 2021). El año pasado, Informa adquirió Faculty of 1000, junto con su plataforma de publicación OA (Open Access) F1000Research (Anónimo, 2020). Para no quedarse atrás, Wiley compró Hindawi, una gran editorial independiente de OA, junto con su plataforma Phenom. La compra de Hindawi siguió a la adquisición en 2016 por parte de Wiley de Atypon, una empresa de software orientada al investigador cuya plataforma en línea, Literatum, Wiley adoptó recientemente en toda su cartera de revistas (Hunter, 2016). “Conoce a tu lector”, dice Atypon sobre Literatum. “Construye informes sobre la marcha y obtén visualización del uso del contenido y el comportamiento del sitio de los usuarios en tiempo real”. Springer Nature, por citar un tercer ejemplo, se encuentra bajo el mismo paraguas corporativo de Holtzbrinck que Digital Science, que incuba *startups* y lanza productos a lo largo del ciclo de vida de la investigación, como el competidor de Web of Science/Scopus Dimensions, el repositorio de datos *figshare*, el rastreador de impacto Altmetric y muchos otros.

Recientemente se produjo una fatídica convergencia: Elsevier anunció un programa piloto para incorporar algunas revistas de Wiley y Taylor & Francis a ScienceDirect de Elsevier (Elsevier, 2022). Si el programa piloto conduce a algo duradero, estaremos un paso más cerca de lo que Leslie Chan ha llamado la “plataformización de la infraestructura académica” (Chan, 2019).

Los grandes oligopolios editoriales no son las únicas empresas que buscan beneficiarse del comportamiento de los investigadores. Está, por supuesto, la propia Clarivate, cuya compra de ProQuest por 5,300 millones de dólares se cerró a finales de 2021, el mismo día que Wiley anunció su compra de Knowledge Unlatched, una empresa con ánimo de lucro (Fund, 2021; Pooley, 2021). Las dos redes sociales académicas respaldadas por capital riesgo, Academia y ResearchGate, vuelven a empaquetar la actividad de los investigadores en los sitios a través de análisis de usuarios; los observadores llevan años especulando con que las empresas crearán productos analíticos basados en sus tesoros de datos. ResearchGate ya vende una herramienta de búsqueda de empleo y publicidad segmentada (“Mejore sus opciones de segmentación con sofisticados anuncios secuenciales”). Las empresas de vigilancia parasitarias de otras facetas de la enseñanza superior sin ánimo de lucro (la vida de los estudiantes, por ejemplo, o el aula) también están creciendo. Las empresas de gestión de programas en línea (OPM, por sus siglas en inglés), un negocio en el que también participa Wiley (McKenzie, 2018), están saliendo a bolsa con valoraciones multimillonarias basadas, según los informes, en el valor de sus decenas de millones de perfiles de “alumnos” (Hill, 2021). Lo mismo ocurre con la empresa de capital riesgo EAB, que promociona su software de asesoramiento académico basado en datos como el primer “sistema de gestión de estudiantes” a nivel empresarial. Incluso la propia Google podría, en cualquier

momento, decidir monetizar su motor de búsqueda Google Scholar, lo que supondría una vuelta a sus raíces bibliométricas.

La comunidad científica está empezando a contraatacar. Existe una página de petición Stop Tracking Science, con más de mil firmas en el último recuento (Anónimo, 2023 b). SPARC North America, el grupo de defensa del Acceso Abierto (OA), ha dado la voz de alarma (Shockey, 2021). La fundación alemana de investigación, Deutsche Forschungsgemeinschaft (DFG), publicó su propio informe-advertencia en octubre: “industrialización del conocimiento a través del seguimiento”, en palabras del informe (DFG-Committee, 2021). Un acuerdo de lectura y publicación para 2020 entre Elsevier y universidades neerlandesas provocó protestas, sobre todo porque la empresa había incluido sus productos de predicción en el acuerdo (de Knecht, 2020).

El sociólogo David Murakami-Wood ya nos lo advirtió en 2009: “Los editores se están convirtiendo, si no el Gran Hermano, al menos en varios pequeños” (Murakami-Wood, 2009). El coro, en los últimos años, se ha hecho más fuerte, con las alertas de Alejandro Posada, George Chen, Lisa Hinchliffe, Leslie Chan, Richard Poynder, Björn Brembs y Renke Siems en *Elephant in the Lab* en el mes de abril 2022 (<https://elephantinthelab.org/>). El problema es que la mayoría de los académicos en activo no tienen ni idea de que se les está empaquetando y vendiendo.

EFFECTOS DE BUCLE

Desviar el dinero de los contribuyentes, de las matrículas y de las dotaciones para acceder a nuestro propio comportamiento es una indignidad financiera y moral. Que estemos pagando a los vendedores por segunda vez, después de los gastos de suscripción y APC que agotan el presupuesto, es un escándalo. El año pasado, Elsevier obtuvo beneficios por valor de 1,400 millones de dólares, con unos ingresos de 3,600 millones, lo que supone un margen de beneficios del 38%. Este lucrativo negocio se basa en el trabajo no remunerado de los académicos, subvencionado por nuestros empleadores universitarios. El producto tipográfico de ese trabajo, según una antigua queja, se nos vende a precios exorbitantes. Ahora Elsevier se queda con la crema del comportamiento y también la vende. En todo caso, los beneficios del primer negocio han financiado la creación del segundo.

Consideremos también el uso previsto de estos productos de vigilancia. Los clientes de muchos de los análisis predictivos vendidos por Elsevier y otros son administradores universitarios y oficinas nacionales de investigación. El objetivo de estos productos es racionalizar las prácticas de valoración y evaluación descendentes que se han impuesto en las últimas décadas, especialmente en el mundo académico anglófono. Algunas de las prácticas, y la mayor parte de la mentalidad, se han tomado prestadas del sector empresarial. En mayor o menor medida, el celo

por la medición está impulsado por la idea de que el principal objetivo de la universidad es hacer crecer las economías regionales y nacionales. Productos como Pure y SciVal se cuentan, o se contarán, entre las herramientas cuantificadas mediante las cuales los valores económicos y de ingeniería configuran lo que entendemos por educación superior. Como mínimo, sus tabulaciones se utilizarán para justificar la “priorización de programas” y otras reasignaciones presupuestarias. Como ha observado Ted Porter, la cuantificación es una forma de tomar decisiones sin que parezca que se decide (Porter, 2020).

En este sentido, las “herramientas de decisión” que venden los editores de vigilancia son máquinas de blanqueo, abstracciones que borran el contexto de nuestras desordenadas realidades académicas. Es cierto que el artículo de investigación estándar, e incluso sus conjuntos de datos subyacentes, ya están abstraídos. Pero las puntuaciones de productividad de los investigadores, por poner un ejemplo, están muy alejadas de nuestras prácticas de creación de conocimiento. Una de las razones por las que esto es importante es que las puntuaciones y los índices algorítmicos pueden camuflar los sesgos que estructuran la vida académica. Pensemos en la dinámica centro-periferia en las líneas Norte-Sur y nativo-inglés: Las brechas que se remontan a la historia geopolítica, incluido el legado del colonialismo europeo, pueden quedar aún más enterradas bajo el peso de las métricas patentadas.

El problema no es meramente de camuflaje. Con toda la autoridad que se concede a las medidas cuantitativas, incluidas las decisiones de financiación y contratación, la puntuación predictiva podría agravar los sesgos encubiertos. Como han demostrado varios estudiosos, las métricas y las clasificaciones ayudan a representar el mundo que pretenden simplemente describir (Espeland y Sauder, 2007; 2008; Fourcade y Johns, 2020). Así, los hablantes nativos de inglés podrían parecer más propensos a producir artículos de impacto, basándose en los datos de citas anteriores utilizados para entrenar un algoritmo predictivo, una medida que, a su vez, podría justificar la concesión de una subvención. Esta dinámica de ventaja acumulativa serviría para ampliar las disparidades existentes, un efecto Matthew a escala de Scopus.

Los efectos en bucle de la puntuación algorítmica pueden incluir jugar con la medida. Como dice la Ley de Goodhart, cuando una medida se convierte en un objetivo, deja de ser una buena medida (Anónimo, 2023 c). Los académicos, al igual que otros sujetos de las mediciones clasificadas, pueden “optimizar” sus artículos para atraer al algoritmo. Si las subvenciones, los ascensos y el reconocimiento les siguen, este comportamiento reforzará un sistema de recompensas ya meteorizado. En palabras de Tarleton Gillespie, podemos ajustar nuestro trabajo para que sea reconocible algorítmicamente, o incluso para vernos a nosotros

mismos a través del prisma de los análisis predictivos de Elsevier (Gillespie, 2017).

REFERENCIAS

- Anónimo. (2020). Informa PLC Press Release. Informa.com. <https://www.informa.com/globalassets/documents/investor-relations/2020/informa-adds-f1000-research.pdf>
- Anónimo. (2023 a). Reading scientific literature can also reveal privacy, Elsevier PDF reader collects user information. Inf.news (s/f). <https://inf.news/en/tech/ebd522a6fc3a471730373ba697f55ce6.html>
- Anónimo. (2023 b) Stop tracking science. (s/f). Stop-trackingscience.Eu. <https://stoptrackingscience.eu/>
- Anónimo. (2023 c). Wikipedia contributors. Goodhart’s law. Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Goodhart%27s_law&oldid=1132595374
- Baykoucheva, S. (2019). Eugene Garfield’s ideas and legacy and their impact on the culture of research. *Publications*, 7(2), 43. <https://doi.org/10.3390/publications7020043>
- Bouk, D. (2018). How our days became numbered: Risk and the rise of the statistical individual. University of Chicago Press.
- Brembs, B. (2021). Massive over-payment of academic publishers has enabled them to buy surveillance technology covering the entire workflow that can be used not only to be combined with our private data and sold, but also to make algorithmic (a.k.a. “evidence-led”) employment decisions. Twitter. <https://twitter.com/brembs/status/1440942564094402560?s=21>
- Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1–7), 107–117. [https://doi.org/10.1016/s0169-7552\(98\)00110-x](https://doi.org/10.1016/s0169-7552(98)00110-x)
- Brin, S., & Page, L. (2012). Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer Networks*, 56(18), 3825–3833. <https://doi.org/10.1016/j.comnet.2012.10.007>
- Buranyi, S. (2017). Is the staggeringly profitable business of scientific publishing bad for science? *The guardian*. <https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science>
- Büttcher, S., Clarke, C. L. A., & Cormack, G. V. (2016). *Information retrieval: Implementing and evaluating search engines*. MIT Press.

- Chan, L. (2019). Platform Capitalism and the Governance of Knowledge Infrastructure. Digital Initiative Symposium 2019, April 29-30, University of San Diego, CA
- Cramer-Flood, E. (2021). "Duopoly Still Rules the Global Digital Ad Market, but Alibaba and Amazon Are on the Prowl." *eMarketer*, May 10, 2021. <https://www.emarketer.com/content/duopolys-till-rules-global-digital-ad-market-alibaba-amazon-on-prowl>.
- Cuéllar, M. F., & Huq, A. Z. (2010). Economies of Surveillance. Book Review. *Harvard Law Review*, 133, 1280–1336.
- de Knecht, S. (2020). Dutch open science deal primarily benefits Elsevier. *ScienceGuide*. <https://www.scienceguide.nl/2020/06/open-science-deal-benefits-elsevier/>
- DFG-Committee (2021). Data tracking in research: aggregation and use or sale of usage data by academic publishers. A briefing paper of the Committee on Scientific Library Services and Information Systems of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation). *Zenodo*. <https://doi.org/10.5281/ZENODO.5937995>
- Elsevier. (2005). *A Short History of Elsevier*. London: Reed Elsevier. <http://www.ask-force.org/web/Seralini/Elsevier-Short-History-2005.pdf>.
- Elsevier. (2022). Discover research: Wiley, ACS, RSC, Taylor & Francis, Elsevier. (s/f). *Sciencedirect.com*. <https://www.sciencedirect.com/org>
- Elsevier. (n.f. a). "Elsevier Fingerprint Engine." Accessed November 16, 2021. <https://www.elsevier.com/solutions/elsevier-fingerprint-engine>.
- Elsevier. (n.f. b). "[Pure] Features." Accessed November 16, 2021. <https://www.elsevier.com/solutions/pure/features>.
- Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American journal of sociology*, 113(1), 1–40. <https://doi.org/10.1086/517897>
- Espeland, W. N., & Stevens, M. L. (2008). A sociology of quantification. *Archives Europeennes de Sociologie. European Journal of Sociology. Europaisches Archiv Für Soziologie*, 49(3), 401–436. <https://doi.org/10.1017/s0003975609000150>
- Fourcade, M., & Johns, F. (2020). Loops, ladders and links: the recursivity of social and machine learning. *Theory and Society*, 49(5–6), 803–832. <https://doi.org/10.1007/s11186-020-09409-x>
- Fund, S. (2021). Wiley acquires open access innovator knowledge unlatched. *Knowledge Unlatched*, A Wiley Brand, Wiley-VCH GmbH – A Company of John Wiley & Sons, Inc. <https://openresearch.community/posts/wiley-acquires-open-access-innovator-knowledge-unlatched>
- Garfield, E. (2006). Commentary: Fifty years of citation indexing. *International Journal of Epidemiology*, 35(5), 1127–1128. <https://doi.org/10.1093/ije/dyl190>
- Gillespie, T. (2017). Algorithmically recognizable: Santorum's Google problem, and Google's Santorum problem. *Information, Communication and Society*, 20(1), 63–80. <https://doi.org/10.1080/1369118x.2016.1199721>
- Hill, P. (2021). Coursera, 2U, and the emerging education platform market. *PhilOnEdTech*. <https://philonedtech.com/coursera-2u-and-the-emerging-education-platform-market/>
- Hunter, S. (2016). Wiley signs definitive agreement to acquire atypon. *Atypon: Online Publishing Platform & Web Development Tools; Atypon Systems, LLC*. <https://www.atypon.com/news/wiley-signs-definitive-agreement-to-acquire-atypon/>
- Informa Group. (2021). *Depth & Data: Informa Group Annual Report and Accounts 2020*. London: Informa Group.
- Lamdan, S. (2021). Your journals are spying on you: Research surveillance in library products. <https://scholarworks.iu.edu/dspace/handle/2022/26873>
- La Monica, P. R. (2004). "Google Jumps 18% in Debut." *CNN Money*, August 19, 2004. <https://money.cnn.com/2004/08/19/technology/goog/>
- Lauer, J. (2017). *Creditworthy: A history of consumer surveillance and financial identity in America*. Columbia University Press.
- McKenzie, L. (2018). Wiley makes a play to lead OPM space, and signals change ahead. *Insidehighered.com*. <https://www.insidehighered.com/digital-learning/article/2018/10/10/wiley-makes-play-lead-opm-space-and-signals-change-ahead>
- Murakami-Wood, D. (2009). Spies in the information economy: academic publishers and the trade in personal information. *ACME*, 8(3), 484–493. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-200719>
- Pooley, J. (2021). Knowledge unlatched strikes again and again. *Jeffpooley.com*. <https://www.jeffpooley.com/2021/01/knowledge-unlatched-strikes-again-and-again/>
- Porter, T. M. (2020). *Trust in numbers: The pursuit of*

objectivity in science and public life. Princeton University Press.

Posada, A., & Chen, G. (2017). Publishers increasingly in control of scholarly infrastructure and this is why we should care. The knowledge Gap. <http://knowledgegap.org/index.php/sub-projects/rent-seeking-and-financialization-of-the-academic-publishing-industry/preliminary-findings/>

Posada, A., & Chen, G. (2018). Inequality in knowledge production: The integration of academic infrastructure by big publishers. 22nd International Conference on Electronic Publishing.

Price, D. J. de S. (1965). Networks of Scientific Papers: The pattern of bibliographic references indicates the nature of the scientific research front. *Science* (New York, N.Y.), 149(3683), 510–515. <https://doi.org/10.1126/science.149.3683.510>

RELX Group. (2016). Annual Report and Financial Statements 2015. London: RELX Group.

Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society*, 6(1), 205395171882054. <https://doi.org/10.1177/2053951718820549>

[org/10.1177/2053951718820549](https://doi.org/10.1177/2053951718820549)

Shockey, N. (2021). Addressing the alarming systems of surveillance built by library vendors. SPARC. <https://sparcopen.org/news/2021/addressing-the-alarming-systems-of-surveillance-built-by-library-vendors/>

Zijlstra, J. (1994). The University Licensing Program (TULIP): A Large Scale Experiment in Bringing Electronic Journals to the Desktop. *Serials*, 7(2), 169–172.

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

Wouters, P. (2017). Eugene Garfield (1925–2017). *Nature*, 543(7646), 492. <https://doi.org/10.1038/543492a>

Una primera versión del presente ensayo se publicó en: Pooley, J., (2022). Surveillance Publishing. *The Journal of Electronic Publishing* 25 (1). <https://doi.org/10.3998/jep.1874>.

La presente versión proviene de: Pooley, J. (2022). Surveillance Publishing. *Elephant in the Lab*. <https://doi.org/10.5281/zenodo.6384605>. Traducción del Comité Editorial *CICIMAR Océánides*.

Copyright (c) 2023 Jefferson Pooley



Este texto está protegido por una licencia [Creative Commons 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Usted es libre para Compartir —copiar y redistribuir el material en cualquier medio o formato- y Adaptar el documento- remezclar, transformar y crear a partir del material— para cualquier propósito, incluso para fines comerciales, siempre que cumpla la condición de:

Atribución: Usted debe dar crédito a la obra original de manera adecuada, proporcionar un enlace a la licencia, e indicar si se han realizado cambios. Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que tiene el apoyo del licenciante o lo recibe por el uso que hace de la obra.

[Resumen de licencia](#) - [Texto completo de la licencia](#)

