

The logo for the journal EMPIRIA, featuring the word in a bold, serif font.

EMPIRIA. Revista de Metodología de las Ciencias Sociales

ISSN: 1139-5737

ISSN: 2174-0682

empiria@poli.uned.es

Universidad Nacional de Educación a Distancia

España

Arribas Macho, José María

CATHY O'NEIL: Armas de Destrucción Matemática. Cómo el Big Data aumenta la desigualdad y amenaza la democracia. Madrid: Capitán Swing, 2017, 269 pp

EMPIRIA. Revista de Metodología de las Ciencias Sociales, núm. 41, 2018, Septiembre-, pp. 199-202

Universidad Nacional de Educación a Distancia

España

Disponible en: <https://www.redalyc.org/articulo.oa?id=297165396011>

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

The logo for redalyc.org, featuring the text "redalyc.org" with a stylized red "r" and "a".

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

CATHY O'NEIL: *Armas de Destrucción Matemática. Como el Big Data aumenta la desigualdad y amenaza la democracia. Madrid: Capitán Swing, 2017, 269 pp.*

Las redes sociales se han convertido en un fenómeno de masas que ha revolucionado los estudios de mercado y de opinión, y en consecuencia se han ido abandonando metodologías como la encuesta estadística representativa en favor de modelos extensivos del tipo *big data*: el estudio masivo de los datos que se obtienen a través de la red. Los llamados estudios de minería de datos se realizan como los antiguos estudios censales, se actúa sobre el conjunto de la población, aunque ahora los datos se producen a partir de internet, y sin tener en cuenta problemas metodológicos como la representatividad de los resultados, la validez, los sesgos, o el marco teórico que está detrás.

El mundo de la publicidad y el *marketing* que habíamos conocido hasta la eclosión de los datos masivos, ha experimentado cambios tan profundos que obligan a replantearse algunos saberes académicos relacionados con los estudios de mercado. El viejo creativo y los gestores de medios tradicionales, son sustituidos por “persuasores *on-line*”, y conceptos como “retorno de la inversión”, “usuarios únicos”, “tasas de conversión”, “clics”, “visitas”, “páginas vistas”, “tiempo de permanencia”, “huella digital”, etc., constituyen un nuevo vocabulario para los sociólogos del consumo. Ahora nos movemos en

un entorno en el que casi todo sucede a través de la red y en las tiendas *on-line*. En los departamentos de investigación de las grandes marcas multinacionales, se ensayan nuevas técnicas de persuasión y el análisis automático sustituye progresivamente al factor humano, de modo que el mayor protagonismo corresponde a la inteligencia artificial o a los diseñadores de algoritmos, los llamados a sí mismos: *científicos de datos*.

Cathy O'Neil, es una de estas diseñadoras de algoritmos, que estuvo trabajando antes de la crisis como analista en D.E.Shaw, un destacado fondo de cobertura que colaboraba con los bancos de inversión, luego estuvo en RiskMetrics Group y en la multi-nacional de venta de alojamientos Expedia. En su libro nos describe con todo lujo de detalles su paso por lo que ahora se conoce como economía digital: “*las operaciones que hacíamos con números se traducían en billones de dólares que pasaban de una cuenta a otra*”, aunque en 2008 cuando llevaba algo más de un año trabajando en el sector, y todo se derrumbó, descubrió que las matemáticas que habían sido su refugio intelectual y profesional, también habían sido un elemento causante de la crisis. En ese contexto conoció el movimiento *Occupy Wall Street*, alter ego del 15M español, y participó en un grupo crítico sobre

banca alternativa. Como apunta O'Neil, en 2008 ya se había producido un cambio de paradigma: el mundo financiero “funcionaba veinticuatro horas al día procesando petabytes de información, en gran parte datos extraídos de las redes sociales o de páginas web de comercio electrónico. Y en lugar de prestar atención a los movimientos de los mercados financieros mundiales, se dedicaba cada vez más a analizar a los seres humanos. Los matemáticos y los especialistas en estadística estudiaban nuestros deseos, nuestros movimientos y nuestro poder adquisitivo. Predecían nuestra solvencia y calculaban nuestro potencial como estudiantes, trabajadores, amantes, o delincuentes”. Lo que ahora conocemos como economía digital proporcionaba ya grandes ganancias en 2008, e indudablemente estuvo detrás de la crisis financiera posterior¹, favoreciendo las maniobras especulativas de los bancos, las hipotecas *subprime* y la quiebra de millones de trabajadores que habían contratado hipotecas sin suficientes garantías.

Cathy O'Neil dedica su libro a criticar los modelos matemáticos opacos que sirvieron, y sirven, para hacer funcionar la economía del *big data*, utilizando la calificación de los consumidores y los demandantes de un crédito hipotecario. A esos algoritmos les asigna un nombre de

resonancias bélicas: “armas de destrucción matemática” (ADM). O'Neil parte de criticar el secretismo de estos algoritmos, pues las empresas no solo impiden que sean públicos sino que se las ingenian para que los científicos de datos que los diseñan no tengan acceso al algoritmo completo, sólo les permiten trabajar con fragmentos concretos, pero además están cargados de prejuicios ideológicos. En aquellos casos en los que se utilizan para la administración pública, por ejemplo en temas de seguridad y delincuencia, subvenciones y becas para escuelas, O'Neil denuncia como las clases altas blancas resultan siempre favorecidas. Los modelos, “*a pesar de su reputación de imparcialidad, reflejan objetivos e ideologías*”. Las ADM se utilizan cada vez más en sectores como salud y recursos humanos, estableciendo “*normas generales que ejercen sobre nosotros una fuerza muy similar al poder de la ley*”.

Con la inteligencia artificial, los modelos aprenden, pero la materia prima para hacer previsiones siempre son los datos históricos, y las actualizaciones y ajustes de los modelos constituyen “puntos ciegos” que reflejan las prioridades de las empresas, problema aparte son las cuestiones más técnicas como correlaciones espúreas o sesgos de conformación. En la economía digital, la veracidad del ADM, la validez de la prueba, es el dinero que gana la empresa, pues “*el problema de los beneficios acaba actuando como un valor sustitutivo de la verdad*”.

Los elementos que conforman un modelo ADM, según O'Neil son, de un lado la opacidad y el secretismo, pero también su aplicación a gran escala, a grandes masas de consumidores (las clases altas siguen teniendo trato personalizado en la contratación de créditos y en las ventas) y el daño que ocasionan, o lo que es igual, la injusticia inherente al modelo: “Cuando los mercados se hundieron en 2008 fue cuando comprendí la horrible verdad. Era incluso peor que

¹ Después de los ataques terroristas de 2001 se produjo cierta recesión que las autoridades trataron de resolver con unos tipos de interés más bajos que alimentaron la burbuja inmobiliaria. Los bancos y las inmobiliarias hicieron creer que cualquiera podía conseguir una hipoteca y se lanzaron a la construcción de tierras alejadas y baldías. Por su parte, “los bancos apostaron miles de millones en toda clase de instrumentos financieros vinculados a la prosperidad inmobiliaria”, pero en 2007, los bancos dejaron de prestarse entre ellos, pues comenzaron a darse cuenta de “la peligrosa basura que tenían en sus carteras”, en consecuencia, los tipos de interés interbancario comenzaron a subir

robar dinero bobo de las cuentas de la gente, el sector financiero estaba en el negocio de crear ADM y yo tenía un pequeño papel en todo aquello”. Otro problema que destaca O’Neil es la utilización de la micro-segmentación, las empresas segmentan el mercado de modo que los mensajes que dirigen a los consumidores son diferentes en función del perfil, lo mismo que sucede en el mercado político y de opinión; para combatir esto propone realizar campañas colaborativas denunciando la micro-segmentación, y haciendo ver a los ciudadanos que los mensajes y la información que ofrecen estas empresas son diferentes en función del tipo de consumidor y el perfil del votante. También propone medir el impacto de los modelos y realizar auditorías de los algoritmos.

Su crítica aparece cargada de humanismo, incluso cuando habla de los aspectos más técnicos: *“Los modelos del big data, codifican el pasado: No inventan el futuro. Para inventar el futuro hace falta imaginación moral y eso es algo que solo los seres humanos pueden ofrecer”*. No hace una crítica totalmente destructiva, al contrario, plantea incorporar a los algoritmos mejores valores, dando prioridad a la justicia sobre los beneficios. Llega a proponer un juramento ético de los programadores, así como sistemas regulatorios de los algoritmos, prestando atención a todo aquello que se resiste a la cuantificación. Lo cual no obsta para que su juicio sobre las ADM sea demoledor: “con su promesa de eficiencia y justicia, distorsionan la educación superior, acrecientan la deuda, incitan a las penas de prisión en masa, golpean a los pobres en prácticamente todas las coyunturas y socavan la democracia”.

Desde el punto de vista metodológico, podemos concluir que una de las diferencias fundamentales entre el método del muestreo de la época inicial de los estudios de consumo y la época actual del *big data*, es que entonces los institutos

de investigación de mercados trabajaban sobre cuestiones sociales. Trabajaban para el Estado o para grandes compañías de servicios públicos como telefonía, electricidad, comunicaciones, etc., la mayor parte de ellas nacionalizadas. Ahora los técnicos del *big data* trabajan para grandes compañías multinacionales, o crean su propia *start-up* con el objetivo de venderla y maximizar rápidamente el beneficio económico. Para Bowley o Hilton, padres del muestreo estadístico, lo fundamental era conocer en Inglaterra el número de parados y organizar el subsidio de paro; para compañías como Amazon, lo fundamental es conocer los hábitos y preferencias de los consumidores con objeto de organizar la demanda y aumentar beneficios. A diferentes objetos, diferencia de métodos. Así la prevalencia de los datos va más allá de sus posibilidades técnicas de tratamiento y análisis, pues está más relacionada con su valor mercantil, lo que algunos han llamado la financiarización de los datos. Dos expertos en *big data* como Mayer-Schönberger, V, Cukier, K., lo plantean sin rodeos: “calcular su valor ya no significa meramente sumar lo que se gana mediante su uso primario (...) la mayor parte del valor de la información está latente y surgirá de unos usos secundarios futuros (...). Esto se parece a los obstáculos que había para ponerles precio a los derivados financieros antes del desarrollo de la ecuación Black-Scholes en la década de 1970 (...); colgarle una etiqueta con el precio al valor de la opción de datos representa, como mínimo, una espléndida oportunidad para el sector financiero”. Llegados a este punto, aparece el verdadero *leit motiv* del tratamiento de los datos masivos: su valor mercantil. Los datos son en realidad, derivados financieros, si no constituyen la base real de una investigación científica responsable, si no tienen por finalidad la resolución de problemas ciudadanos, su utilización tampoco requiere de muchas precauciones metodológicas. No es

necesario complicarse la vida con la exactitud, la validez, o los límites del error probable, lo único importante es el balance de resultados. Como apunta O'Neil en el título de su libro, la economía del *big data*, no solo aumenta la desigualdad, también amenaza la democracia. En todo caso nos enfrentamos a un mundo cargado de posibilidades para los sociólogos, lástima que ahora no suceda como en los años 1920 cuando la investigación estadística se hacía fundamentalmente para el Estado o para un sector público que todavía concedía valor al interés general.

BIBLIOGRAFÍA:

ALONSO, L. E., FERNÁNDEZ RODRIGUEZ, C.J. (2018) "Poder y sacrificio. Los nuevos discursos de la empresa". Madrid: Siglo XXI

MAYER- SCHÖNBERGER, V, CUKIER, K. (2013): "Big data (La revolución de los datos masivos)". Turner Noema.

ALSUDAIS, A. (2017). Quantifying the Offline Interactions between Hosts and Guests of Airbnb.

https://www.researchgate.net/publication/316551046_Quantifying_the_Offline_Interactions_between_Hosts_and_Guests_of_Airbnb

COTARELO, R.; GIL, J. (comp.) (2017) "Ciberpolítica. Gobierno abierto, redes, deliberación, democracia. Madrid: INAP (Instituto Nacional de Administración Pública),

José María. Arribas Macho

UNED.