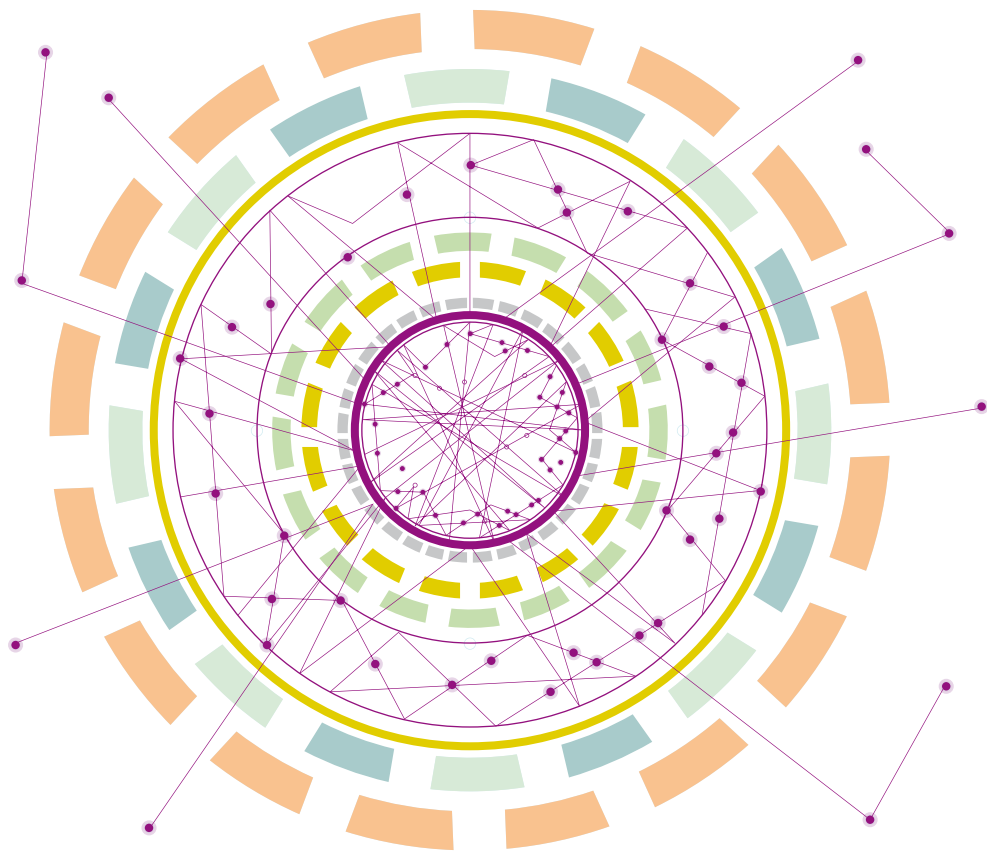


INTERVENCIONES FEMINISTAS PARA LA IGUALDAD Y LA JUSTICIA



Compiladoras: **Diana Maffía, Patricia Gómez, Aluminé Moreno y Celeste Moretti**



Poder Judicial de la Ciudad de Buenos Aires
Consejo de la Magistratura



Intervenciones feministas para la igualdad y la justicia



www.editorial.jusbaire.gob.ar
editorial@jusbaire.gob.ar
fb: /editorialjusbaire
Av. Julio A. Roca 534 [C1067ABN]
+5411 4011-1320



Sello
**Buen
Diseño**
argentino

Intervenciones feministas para la igualdad y la justicia / Compilado por Diana Maffía; Patricia Gómez; Aluminé Moreno; Celeste Moretti. Danila Suárez Tomé [et al.] -1a ed.- Ciudad Autónoma de Buenos Aires: Editorial Jusbaire, 2020.
600 p.; 22 x 16 cm.

ISBN 978-987-768-137-6

1. Derecho. 2. Estudios de Género. 3. Feminismo. I. Maffía, Diana, comp. II. Gómez, Patricia, comp. III. Moreno, Aluminé, comp. IV. Moretti, Celeste, comp. V. Suárez Tomé, Danila.
CDD 341.4858

© Editorial Jusbaire, 2020

Hecho el depósito previsto según Ley N° 11723

Declarada de interés por la Legislatura de la Ciudad Autónoma de Buenos Aires.

Res. Nro. 543-2018

Consejo Editorial

Presidenta:

Ana Salvatelli

Miembros:

Alberto Maques

Francisco Quintana

Fabiana Haydeé Schafrik

Mariana Díaz

Sergio Delgado

Alejandra García

Editorial Jusbaire

Coordinación General: Alejandra García

Dirección: Gerardo Filippelli

Edición: Martha Barsuglia, María del Carmen Calvo

Corrección: Daniela Donni, Leticia Muñoz, Mariana Palomino y Julieta Richiello

Coordinación de Arte y Diseño: Mariana Pittaluga

Maquetación: Carla Famá

La presente publicación ha sido compuesta con las tipografías *Saira* del tipógrafo argentino Héctor Gatti para la fundidora Omnibus-Type y *Alegreya* de la fundidora argentina Huerta Tipográfica.

Maternidad y justicia penal. Prácticas de abandono, aborto e infanticidio en la Ciudad de Buenos Aires (fines del siglo XIX - principios del siglo XX)	
Julieta Di Corleto	251

Capítulo 7

Intersecciones entre género, tecnologías y derechos

Tecnopersonas: sujetos alienados	
Lola Almendros	280

Hombre es a mujer como inteligencia es a lucirse. Los <i>big data</i> y la desigualdad de género	
Capitolina Díaz Martínez y Pablo Díaz García	297

Sesgos de género en la Inteligencia Artificial	
Diana Maffía	319

Capítulo 8

Movimientos sociales: la gota que horada la piedra

Agendas judiciales y repertorios sociales. Iniciativas de formación en género en el sistema de justicia previas a la Ley Micaela	
Aluminé Moreno y Roberta Ruiz	338

Ambiente, género y derecho: hacia un derecho ambiental ecofeminista y un Estado feminista de derecho	
María José Lubertino	375

El camino hacia la igualdad en la publicidad: acuerdos, tracciones y esfuerzos feministas	
Marcela Espíndola	414

Hombre es a mujer como inteligencia es a lucirse. Los *big data* y la desigualdad de género

Capitolina Díaz Martínez* y Pablo Díaz García**

En este capítulo vamos a presentar algunos pros y contras de las tecnologías de la información y la comunicación (TIC) y, en particular, de los macrodatos (*big data*), en relación con la des/igualdad de género. Comenzaremos con una presentación esquemática de las posiciones divergentes acerca de las TIC y los *big data*, con ejemplos al respecto. A continuación, nos centraremos en un trabajo empírico propio que revela las formas invisibles, pero no por ello menos perversas y omnipresentes, a través de las cuales los *big data* no solo reproducen, sino que también incrementan las desigualdades de género. Mostraremos cómo el aparentemente neutral uso de un algoritmo de clasificación de datos discrimina (en principio, no intencionalmente) a las mujeres.

¿Las TIC y los *big data*, ayudan o perjudican a las mujeres?

Las grandes innovaciones, aquellas que afectan la vida de multitud de personas, suelen generar controversia. Para algunas personas, los aspectos positivos de su incorporación a la vida social son los

* Catedrática de Sociología de la Universidad de Valencia. Doctora en Sociología por la Universidad de Londres. Profesora en la Universidad de Oviedo desde 1992 a 2012. Directora de la Unidad de Mujeres y Ciencia del Ministerio de Educación y Ciencia y Directora General de Empleo en el Ministerio de Igualdad.

** Técnico superior en Administración de Sistemas Informáticos en Red. Trabajó como técnico de telecomunicaciones en Telice Comet, administrando sistemas de comunicación subterránea y programando dispositivos de IoT (Internet de las cosas). Actualmente trabaja como técnico de producto para T-Systems Iberia.

únicos destacables, mientras que otras se centran en sus efectos negativos, al menos sobre una parte de la sociedad. La irrupción de las TIC en nuestras vidas ha sido y está siendo tan notable que, sin duda, debe tener detractores y entusiastas. Dentro de estas tecnologías, los *big data* o macrodatos tampoco se libran de la controversia. Nos referiremos a estos últimos.

Los *big data*¹ son enormes cantidades de datos recogidos de forma pasiva a partir de las interacciones digitales realizadas por la gente o por las máquinas entre sí. Son datos muy variados, acumulados y usables a velocidades vertiginosas. Doug Laney los caracterizó inicialmente, en 2001, en términos de “3 V”: volumen (cantidad de datos), variedad (diferentes tipos de datos) y velocidad (velocidad a la que se generan y transmiten los datos).²

Además de su poder para describir un fenómeno, pueden ser utilizados –como ningún conjunto de datos previos permitía– para identificar patrones, tendencias, relaciones, correlaciones, etcétera. El mero uso del celular, de una tarjeta de crédito, nuestras conversaciones en las redes sociales, nuestra interacción con las administraciones públicas a través de Internet, junto con las imágenes que capturan los satélites geoespaciales, entre otras muchas actividades, generan los *big data*. Y de esa generación gigantesca de datos se pueden colegir (para bien o para mal) muchas cosas.

El volumen de los *big data* es tan masivo (hoy se mide en *zettabytes*³ pero dentro de poco habrá que crear otra unidad mayor para medirlo) que no se puede extraer información de ellos o analizarlos con las técnicas tradicionales de procesamiento de datos. El manejo de los *big data* presenta grandes dificultades. Es por eso por lo que es un campo per-

1. Mashey, John R., “Slides from invited talk”, 24/04/1998. Disponible en: https://static.usenix.org/event/usenix99/invited_talks/mashey.pdf [fecha de consulta: 26/09/2016]. Se le atribuye la popularización del término *big data*.

2. Laney, Doug, “3D Data Management: Controlling Data Volume, Velocity and Variety”, en *Application Delivery Strategies*, File 949, 2001. Disponible en: <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> [fecha de consulta: 02/09/2016].

3. Un *zettabyte* (ZB) es una unidad de almacenamiento de información que equivale a 10²¹ bytes. En 2018, los datos almacenados en el planeta llegaron a 18 *zettabytes* y se espera llegar a los 175 *zettabytes* en 2025. Disponible en: <https://www.bernardmarr.com/default.asp?contentID=1846> [fecha de consulta: 25/11/2019].

manente de innovación, desde el modo como se capturan, almacenan y clasifican los datos, hasta la forma en la que se hacen las búsquedas y los algoritmos desarrollados con ese fin. Si bien lo que inicialmente impresiona es la captura y el almacenamiento continuo y acumulativo de *zettabytes* de datos, lo más sobresaliente es que esa magnitud de datos favorece, de algún modo, la configuración intencional de nuestro futuro inmediato –y a más largo plazo– porque indica tendencias y permite, a quien tenga el poder y/o la capacidad para eso, operar sobre la base de las previsiones, alterando así (para bien o para mal de según quiénes) lo que hubiera sido el desarrollo no intervenido.

Los *big data* ¿imprescindibles? en la igualdad de género

Muchas organizaciones en pro de la igualdad de género han encontrado en los *big data* una herramienta de gran valor. Algunas lo ven como el gran equalizador. Suponen que proporciona datos fríos y neutros que evitan sesgos producidos por la falta de información (brecha de género de datos). Desde la Organización de Naciones Unidas (ONU)⁴ a las investigadoras feministas, pasando por figuras reconocidas como Hillary Clinton (“los datos no solo miden el progreso, lo inspiran”)⁵ o Melinda Gates⁶ u organizaciones específicas como Global Partnership Equal Measures 2030,⁷ hace años que nos venimos quejando de la brecha de datos de género y hay un cierto optimismo sobre la capacidad de los *big data* para denunciar y corregir flagrantes discriminaciones de género. La idea en la que se basan es que disponer de

4. Abreu Lopes, Claudia y Bailur, Savita, “Gender Equality and Big Data. Making Gender Data Visible”, en UN-Women, 2018. Disponible en: <https://www.unwomen.org/en/digital-library/publications/2018/1/gender-equality-and-big-data> [fecha de consulta: 10/12/ 2018]; Vega Montiel, Aimée, “Gender Equality and Big Data in the Context of Sustainable Development Goals”, en *Partecipazione e Conflicto*, vol. 11, N° 2, 2019, pp. 544-556.

5. Hillary Clinton en conferencia ante la ONU en 2012, presentando “Data2X”.

6. Fearnow, Benjamin, “Bill, Melinda Gates says big data is ‘sexist’”, en Newsweek, 19/02/2019. Disponible en: <https://www.msn.com/en-us/news/technology/bill-melinda-gates-says-big-data-is-sexist/ar-BBTpi20> [fecha de consulta: 11/11/2019].

7. Disponible en: <http://www.data4sdgs.org/node/806>

buenos datos que revelen la desigualdad puede ser el impulso necesario para que las autoridades actúen en favor de las mujeres.

Una de las primeras iniciativas en búsqueda de *big data* de calidad sobre las mujeres es la plataforma “Data2X”. Esta plataforma es el resultado de la colaboración de varias agencias de la ONU, gobiernos, asociaciones, académicas y mecenas de la empresa privada.⁸ Su objetivo es reducir la brecha de datos de género mediante la promoción y expansión de la producción de datos no sesgados sobre las mujeres, con el fin de mejorar las políticas, las estrategias y la toma de decisiones relativas a la vida de las niñas y las mujeres del mundo. Basan su trabajo en una filosofía bien establecida en la ONU: *no data, no problem, no policy*. Que podríamos traducir por: si no hay datos, no hay problema y si no hay problema, no hacen falta políticas. En “Data2X” dan tanta importancia a los datos que su lema es “sin igualdad de datos no hay igualdad de género”.

Gracias a esta plataforma hoy se dispone de datos de las mujeres en espacios no muy transparentes como la economía *gig*⁹ (trabajo informal, pobreza, precariedad, etcétera), de mapas sobre su riesgo y vulnerabilidad en las calles, así como de información acerca de su situación económica, sanitaria o educativa en lugares remotos. Ciertamente, los datos con los que trabajan los proyectos de “Data2X” y otros de la ONU no siempre son lo que se entiende por *big data*. En parte, son datos no estructurados (como la generalidad de los *big data*), como los derivados del uso por las mujeres de sus tarjetas de crédito y de sus teléfonos celulares. Pero, otra parte, son datos estructurados, producidos específicamente para alcanzar los fines del proyecto. Algunos ejemplos son la reunión de datos sobre las mujeres en zonas remotas de países en los que los datos son escasos, como Nepal, Uganda o Kenia,¹⁰ por ejemplo, datos sobre género, agua y agricultura¹¹ y nume-

8. La William and Flora Hewlett Foundation, la Bill & Melinda Gates Foundation y la Clinton Foundation son los soportes económicos principales de Data2X.

9. Economía *gig*: un tipo de empleo basado en trabajos cortos, esporádicos y relacionados con las tecnologías de la información y la comunicación.

10. Disponible en: <https://unstats.un.org/unsd/demographic-social/meetings/2018/tokyo-globalforum-genderstat/Presentation/7d.%20Data2X%20-%20Big%20data%20for%20gender%20-%20expanding%20horizons%20and%20recognizing%20limitations.pdf> [fecha de consulta: 12/12/2018].

11. Disponible en: <http://www.iwmi.cgiar.org/2015/11/gender-agricultural-water-and-big-data-on-the-table/> [fecha de consulta: 15/01/2019].

rosos datos en relación con los Objetivos de Desarrollo Sostenible. De hecho, esta combinación de datos no estructurados y estructurados se trasluce en la definición de ONU Mujeres de los *big data* que difiere un poco de la que hemos propuesto más arriba: “*Big data* se refiere a fuentes de datos que requieren nuevas herramientas o métodos para su captura, custodia y procesamiento de una forma eficiente”.¹²

Así pues, para la ONU, no se trata solo de la producción y el almacenamiento de datos, sino de su eficiente manipulación. Esto es, se necesitan datos, pero se necesitan también analistas que determinen el significado de los datos. En cualquier caso, el optimismo de la ONU por el potencial aporte de los *big data* a la igualdad es notable. Queda reflejado cuando dice:

Varios proyectos piloto han demostrado la viabilidad de utilizar análisis de *big data* para el desarrollo sostenible. Por ejemplo, nuevas ideas sobre patrones de movilidad, interacciones sociales, sentimientos y actitudes, actividad económica, alerta temprana y bienestar comunitario pueden derivarse de un agregado de datos con privacidad protegida.¹³

En el ámbito privado, el uso de los *big data* para revelar situaciones de desigualdad de género, también ha sido notable. Un ejemplo es el estudio realizado por Cath Sleeman sobre las mujeres en la industria de la creatividad. Revisó más de medio millón de artículos de las secciones de creatividad del periódico *The Guardian* de 2000 a 2018. En ese estudio encontró no solo muchas más referencias a hombres que a mujeres, sino que los verbos y las expresiones asociados a ellas y a ellos diferían. Así, por ejemplo, se presta mayor atención a los sonidos particulares de las mujeres, como “risas”, “gritos”, “carcajadas” y “gruñidos” y a reacciones no verbales, como “sonrisas”, “muecas” y “asentimientos”. Mientras que es mucho más probable que aparezca “él dirigió” que “ella dirigió”, y de manera similar “él actuó”, “él diseñó”, “él administró” y “él fundó”.¹⁴ Otro ejemplo conocido es la investigación sobre diferencias de género en el mundo de la abogacía de

12. Abreu Lopes, Claudia y Bailur, Savita, *op. cit.*, p. 1.

13. Ídem.

14. Sleeman, Cath, “Big data analysis reveals staggering extent of gender inequality in creative industries”, en *The Conversation*, 2018. Disponible en: <http://theconversation.com/big-data-analysis-reveals-staggering-extent-of-gender-inequality-in-creative-industries-121482> [fecha de consulta: 11/11/2019].

alto nivel en los Estados Unidos. Es un tema de estudio clásico, pero la investigación realizada por *Sky Analytics* es la primera con *big data*. El estudio se hizo a partir de las facturas de 40.000 abogadas/os y procuradoras/es (*bill timekeepers*) y como se informa en el propio estudio es “el primer estudio de género basado en 3.400 millones de dólares estadounidenses”¹⁵ cuyas revelaciones han sido imposibles de cuestionar y han puesto en evidencia la disparidad de género en las firmas de abogacía. Otra de las grandes iniciativas es *Gender Gap Grader*.¹⁶ Han creado una herramienta digital que sirve para medir las diferencias de género en los *big data*. Puede aplicarse a cualquier campo, en cualquier lugar del mundo, si se dispone de un volumen de datos lo bastante grande. Sus investigaciones en el ámbito de la inversión, la ciencia o la aviación han supuesto un cambio notable en la percepción de esos campos. Sus herramientas están disponibles *online*.

En resumen, con las herramientas apropiadas, el análisis de los macrodatos que involuntariamente generamos, en nuestra actividad diaria, se pueden poner de manifiesto las notables desigualdades y discriminaciones de género que existen en el planeta.

Los *big data* informan de las diferencias de género, pero tienen sesgos que pueden incrementarlas

Si bien los *big data*, como acabamos de ver, pueden esclarecer la situación de las mujeres de manera más incontestable que otras formas de análisis convencional, también presentan grandes limitaciones. Nos referiremos a tres tipos de limitaciones: 1) ¿a quién dejan fuera? 2) ¿son socialmente neutros (especialmente en relación con el género) los algoritmos con los que se manejan?, y 3) su omnipresencia ¿condiciona nuestra vida más allá de lo éticamente aceptable?

15. Augur, Hannah, “Big data reveals big gender inequality”, en *DATA CONOMY*, 2016. Disponible en: <https://dataconomy.com/2016/03/big-data-reveals-big-gender-inequality> [fecha de consulta: 04/11/2019].

16. Disponible en: <http://gendergapgrader.com/> [fecha de consulta: 04/11/2019].

¿Quién queda fuera de los *big data*?

La percepción de los *big data* ha pasado de las “3V” descritas arriba y correspondientes a 2001 al ecosistema de “3C” definido por Letouzé, en 2015: migas (*crumbs*, en inglés) digitales (traducciones digitales de la acción humana y las interacciones capturadas por dispositivos digitales); capacidades elevadas para recopilar, agregar y analizar datos, y comunidades involucradas en la generación, el gobierno y el uso de datos (incluidos los generadores de datos, usuarios finales, formuladores de políticas, defensores de la privacidad y comunidades de *hackers* cívicos).¹⁷ Pero ni siquiera esta inclusión de las comunidades en la generación y el manejo de los *big data* nos acerca en igualdad a ellos. Como se indica en el informe de Intel para la ONU:

Tanto en los países desarrollados como en desarrollo, las mujeres constituyen la minoría de usuarios de las TIC. El acceso y uso de las TIC por parte de las mujeres es particularmente limitado en contextos pobres, en donde no solo las infraestructuras de comunicación son menos penetrantes y de menor calidad, sino que existen condiciones de pobreza, marginación y violencia de género que determinan el uso que las mujeres hacen de las tecnologías. En los países en desarrollo, hay 600 millones de mujeres en línea, es decir, 200 millones menos que los hombres. Por región, los datos muestran una dramática brecha digital de género: En América Latina, el acceso de las mujeres a Internet es del 20% menos que los hombres.¹⁸

Así pues, las mujeres (y otros grupos empobrecidos) no participan en igualdad en ninguna de las 3C: Ni producen tantos datos “migas” digitales porque usan menos tecnología digital (menos celulares, tarjetas de crédito, viajes, e-gobernanza, dispositivos de Internet, etcétera). Ni controlan los sistemas capaces de almacenar y gestionar datos (mucha menor participación en la industria digital y mucho menos poder económico). Ni participan en comunidades que buscan el

17. Letouzé, Emmanuele, “Thoughts on Big Data and the SDGs”, 2015. Disponible en: <https://sustainabledevelopment.un.org/content/documents/7798BigData%20-%20Data-Pop%20Alliance%20-%20Emmanuel%20Letouze.pdf> [fecha de consulta: 10/10/2019].

18. Intel & Dalberg Global Development Advisors, “Women and the Web”, 2012. Disponible en: <https://www.intel.com/content/dam/www/public/us/en/documents/pdf/women-and-the-web.pdf> [fecha de consulta: 10/10/2019].

control y gobierno de los datos porque sus actividades de cuidado y las restricciones sexistas de participación en la vida pública les dejan menos oportunidades para esas actividades. En resumen, los *big data* dejan fuera a muchas mujeres y a una parte importante de la población menos rica. Hay un viejo adagio en informática *garbage in, garbage out* que podría traducirse como “si metes datos basura, obtienes basura”, en nuestro caso, cabe interpretarlo como que, “si no metes a las mujeres en los datos, no hay mujeres en el análisis, no existen”.

¿Son socialmente neutros los algoritmos con los que se manejan los *big data*?

Los algoritmos solo pueden ser lo buenos que sean los datos con los que trabajan. Si, como hemos visto, hay mucha población que por su escasa actividad digital queda fuera de los *big data*, los algoritmos con los que se almacenan, se agregan y se gestionan los macrodatos no pueden dar cuenta de esa población. El escándalo producido por los sesgos (sobre todo de etnia y género)¹⁹ de ciertos algoritmos ha llevado al desarrollo de agrupaciones como Algorithm Watch o The Algorithmic Justice League, que trabajan en pro del desarrollo de algoritmos incluyentes no sesgados. En esta dirección es posible aplicar una herramienta como *word2vec*, que es la que hemos usado en la investigación que describiremos más abajo y que da cuenta del sexismo implícito en los contenidos de la Wikipedia en español.

La omnipresencia de los *big data* ¿condiciona nuestra vida más allá de lo éticamente aceptable?

El uso continuo de dispositivos digitales que van dejando rastro-datos (de aquello que hacemos, dónde estamos, qué buscamos, qué compramos, cuál es el estado de nuestra salud, de nuestras finanzas, de nuestra relación con las autoridades, con quién hablamos, con quién nos escribimos o qué tipo de chistes nos gustan), de manera agregada genera una información sobre la cual no solo no tenemos control, sino que, con frecuencia, ni siquiera sabemos que existe. Si el

19. Buolamwini, Joy, “How I’m fighting bias in algorithms”, en *TED talk*, 2015.

perfil que toda esa información sobre una persona puede atentar contra sus derechos a la privacidad, los datos acumulados de millones y millones de actos y de personas puede generar un tipo de información que permita la toma de decisiones (data-acción) sobre asuntos que alteren nuestras vidas. Y esas decisiones pueden ser tomadas por personas o entidades sobre las que el control sea muy escaso. Algunos organismos como la Unión Europea han desarrollado normativa relativa a la propiedad, transparencia, consentimiento, privacidad y apertura de los datos de cada individuo.²⁰ Pero queda mucho debate ético pendiente²¹ sobre cómo manejan esos datos las grandes corporaciones comerciales y los gobiernos y, sobre todo, cómo se acumulan y gestionan los datos masivos de multitud de personas. Puede ser que en estos momentos ya no se pueda controlar el almacenamiento y gestión de los datos existentes, aunque, sin cesar, seguimos produciendo involuntariamente datos de manera exponencial.

Sesgos de género en los *big data*. La Wikipedia, un estudio de caso

Tolga Bolukbasi²² en un inspirador artículo consigue impresionarnos al mostrar cómo el encaje de pares de palabras (conceptos) de un *corpus* textual cualquiera revela, de forma incuestionable, el sexismo estructural de nuestro lenguaje. Bolukbasi y sus colegas analizaron 3 millardos de datos de *Google News*. Trabajan con un algoritmo capaz de captar relaciones del tipo:

... *Si París es a Francia, Tokio es a X, y encuentra que X es Japón.*
Y al usar esta lógica vectorial se encuentran que al preguntar

20. Disponible en: https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en

21. Mittelstadt, Brent y Floridi, Luciano, "The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts", en *Science and Engineering Ethics*, vol. 22, N° 2, 2016, pp. 303-341.

22. Bolukbasi, Tolga *et al.*, "Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings", en *Cornell University*, 2016. Disponible en: [arXiv:1607.06520v1](https://arxiv.org/abs/1607.06520v1) [cs.CL]; Disponible en: <https://arxiv.org/abs/160706520https://arxiv.org/abs/1607.06520m> [fecha de consulta: 05/08/2016].

... Si hombre es a mujer, informático es a X, arroja para X el resultado *ama de casa*.

O el par: ... Si hombre es a mujer, médico es a X, resultando X ser *enfermera*.

El propio Bolukbasi nos indica que el problema no es solo semántico (que no es poco) sino que puede tener consecuencias de incremento del sexismo de forma inmediata en el mundo real extralingüístico. Imaginemos, por ejemplo, que una empresa necesita contratar personas expertas en programación y le pide unas decenas de nombres a Google, que este operador obtendrá a partir de los *big data* de su *Google News*. En esa enorme base de datos, no aparece “Mujer” como par compatible con “Programador informático”. Por consiguiente, ninguna mujer figurará en la lista que Google venda a la empresa cliente, y ninguna mujer llegará a ser contratada. Se ampliaría así el círculo vicioso de la discriminación de género.

El interés en conocer cómo funcionan los algoritmos de búsqueda en los *big data* radica, no tanto en el hecho de mostrar el sexismo del lenguaje, tema conocido por numerosos estudios, sino que la supuestamente inocencia de las matemáticas, su pretendida neutralidad, deja fuera, no reconoce, a gran parte de la población o no reconoce sus características, sus valores y sus experiencias. El caso de no reconocimiento facial de la cara de una mujer afroamericana²³ porque el algoritmo solo estaba entrenado para reconocer caras caucásicas, es uno de los ejemplos más notable de los sesgos de los buscadores digitales. Si vivimos en un mundo, en buena medida marcado por los *big data* y los mecanismos matemáticos para gestionar esos *big data* ignoran o deforman la realidad social, la imagen que nos arrojen, en lugar de servir para la igualdad y el progreso social, servirá para la reproducción, notablemente ampliada, de las desigualdades y parecerá que no hay responsables, ya que corre el mito de que los ordenadores y las matemáticas son objetivos y neutrales.

A la vista de la capacidad de la red neural usada por Bolukbasi para evidenciar los sesgos de género en un cuerpo de texto, decidimos probar con *big data* en español. Al no contar más que con dos personas y un modesto ordenador personal, nuestra capacidad de manejo de datos era relativamente escasa. Por eso, nos limitamos a lo que en el

23. Buolamwini, Joy, *op. cit.*

argot informático se conoce como una “prueba de concepto”, que es una implementación resumida de una teoría con la intención de comprobar si esa teoría es susceptible de ser explotada de forma útil. Nuestra prueba se realizó sobre 28 millones de palabras de la Wikipedia en español. Comenzaremos por explicar qué es y cómo funciona la red neural con la que realizamos la búsqueda.

La red neural *word2vec*

Word2vec es una red neural, desarrollada originalmente por Google, que procesa texto. Se puede aplicar a lenguaje natural escrito, pero también a genes, códigos, gustos, listas de música y en general a cualquier serie verbal o simbólica de la que se puedan extraer patrones.

Una red neural es un modelo informático automatizado de aprendizaje inspirado en los sistemas nerviosos biológicos. La red neural está preparada para “aprender” a reconocer conceptos a partir del repetido procesamiento de los datos. La red va reconociendo, entre millones, lectura tras lectura de todo el *corpus* textual del que se trate, aquellos elementos de la red que contribuyen a conformar un concepto.

Word2vec, la red neural que utilizaremos en esta investigación analiza y convierte en números (un vector) las palabras que se le introducen. Además, busca la probabilidad de que los estados discretos²⁴ que se dan en la red coocuran.

En *word2vec* se parte de un vector con varios cientos de dimensiones (en nuestro caso, 500). Los vectores de palabras ayudan a los ordenadores a “aprender” del texto.

Cada una de las palabras que analiza la red neural no se analiza por separado, sino en grupos definidos; en el caso del lenguaje natural, esos grupos son frases. Si, por ejemplo, se entrena a esta red con las siguientes frases:

- Mi perro es azul.
- Mi gato es azul.
- Mi perro es verde.
- Mi perro es amarillo.

24. Un sistema discreto es aquel con un número contable de estados, en nuestro caso cada una de las palabras del diccionario.

En el espacio resultante, las palabras “azul”, “verde” y “amarillo” se encontrarán agrupadas entre ellas, estando todas a la misma distancia de “perro”; pero solo “azul” estará cerca de “gato”, a su vez agrupada con “perro”.

A lo largo de numerosas iteraciones con mucha cantidad de frases, las palabras similares van agrupándose y distanciándose de manera adecuada, gracias a la magia de las matemáticas.

En nuestro modo de empleo, las palabras a analizar provienen de textos en lenguaje natural, como una enciclopedia, los artículos de *Google News* o la *Biblia*, siempre que tales textos estén digitalizados. Tras el proceso de aprendizaje, la red neural transforma ese texto en números, con los cuales el ordenador puede realizar operaciones matemáticas a gran escala; ya sea sumas, restas o, lo más interesante, sistemas de ecuaciones, que permiten obtener relaciones de palabras semánticamente sugerentes.

Los algoritmos de *word2vec*, finalmente, ven tantos ejemplos que pueden inferir el género de una sola palabra, igual que pueden inferir que *La Nación* y *El País* son periódicos, que *Matrix* es una película de ciencia ficción, y que el estilo de una prenda de vestir podría ser *hippy* o formal. Que esos vectores de palabra representen una gran parte de la información disponible en una definición de diccionario es un efecto secundario conveniente y casi milagroso de esa tarea algorítmica que consiste en intentar predecir el contexto de una palabra.²⁵

Construcción de nuestra red neural *word2vec*

La versión de *word2vec* que utilizamos en esta prueba de concepto es la disponible en el proyecto de software libre *deeplearning4j*, mantenido por la corporación SkyMind.²⁶

Cuando *word2vec* analiza texto, distingue entre dos unidades básicas: palabras sueltas o *tokens*, y agrupaciones de palabras en frases (*sentences*). El programa dispone de unos clasificadores básicos, que le permiten identificar cada grupo de caracteres entre dos espacios

25. Moody, Chris, “Una palabra vale mil vectores”, 2015. Disponible en: <http://multithreaded.stitchfix.com/blog/2015/03/11/word-is-worth-a-thousand-vectors/> [fecha de consulta: 20/10/2016].

26. Disponible en: <https://deeplearning4j.org/word2vec> [fecha de consulta: 20/10/2016].

como un *token* y cada grupo de *tokens* entre dos pasos de carro como una frase. Esta clasificación es importante, ya que el aprendizaje está basado en cómo se localizan las palabras dentro de una frase, independientemente de las demás.

En nuestro caso, hemos acabado utilizando párrafos enteros como frases y palabras sueltas (unidades entre dos espacios) como *tokens*, filtrando y descartando cualquier carácter que no fuera una letra o un guion (-), para evitar, por ejemplo, que “punto” y “.” resultaran dos entradas de vocabulario distintas.

Una vez contruidos estos clasificadores, solo falta entrenar la red neural. Para entrenarla se deben definir ciertos atributos de los cuales los más interesantes en este estudio serían:

Frecuencia mínima de palabras: Indica cuántas veces debe aparecer un *token* en el *corpus* para ser reconocido. Ajustar este aspecto ayuda a filtrar palabras superfluas.

Tamaño de ventana: Indica el número de pasos hacia delante y hacia atrás que se consideran dentro de la frase a la hora de comparar y analizar cada *token*. En principio, cuanto más grande sea la ventana, mejor; pero si nuestras frases son pequeñas o disponemos de poco poder computacional es mejor reducir este tamaño de ventana. Este, en nuestro caso, después de varias pruebas, fue establecido en 8.

Iteraciones y épocas: Indican el número de veces que se repite el *corpus* al introducirlo al algoritmo. La diferencia entre iteración y época es que las iteraciones indican cómo se introduce el *corpus* y las épocas son el número de veces que se usan todas las palabras. Un ejemplo gráfico para visualizar esto podría ser un libro que contiene el *corpus*. Las iteraciones son el número de veces que el *corpus* aparece dentro del libro, y las épocas, cada una de las lecturas completas que se hacen del mismo para estudiarlo. Iteraciones son repeticiones inmediatas del *corpus*, mientras que épocas son repeticiones secuenciales completas.

Como el caso anterior, cuanto más altos sean los valores asignados a las iteraciones y a las épocas, más precisa será la red. Conviene advertir, sin embargo, que la asignación de valores altos a iteraciones y épocas multiplica exponencialmente el tiempo de entrenamiento de la red.

Análisis del empleo de *word2vec* en una edición de la Wikipedia en español

Nuestro *corpus* de análisis está extraído de la Wikipedia en español de 2006, de la que pudimos disponer por cortesía de la Universidad Politécnica de Catalunya que la ofrece en abierto.²⁷ La ventaja de usar este *corpus* específico, aparte de la gran cantidad de información que contiene, estriba en que esta versión está reducida a texto plano (solo caracteres, sin ningún formato), casi inmediatamente procesable por *word2vec*. Además, la Fundación Wikimedia ha reconocido la existencia de sesgo de género y falta de diversidad en la Wikipedia y sabe que todavía no ha solucionado este problema. Confiábamos en que ambos problemas serían fácilmente identificables en nuestro análisis del considerable *corpus* representado por la Wikipedia (en español) de 2006.

Debido a limitaciones de tiempo y poder computacional, se entrenó la red neural usando la mitad del *corpus*, 28.291.729 palabras. Se practicaron 2 iteraciones y 6 épocas. El tiempo de entrenamiento, con estos parámetros, ascendió a 38 horas. El peso total de la base de datos tras el entrenamiento es de 1,16 GB, un tamaño no despreciable. Una vez entrenada la red neural, se procedió al análisis de la aplicación.

Técnicas de análisis desarrolladas expofeso

La red neural genera una representación de cada una de las palabras del vocabulario del *corpus*, ordenadas entre sí de tal manera que las distancias entre ellas representan su uso en el lenguaje con el que se entrenó esa red.

A partir de esas representaciones se pueden realizar varias operaciones, como distancias absolutas entre dos palabras, listado de palabras más cercanas y concatenación de operaciones.

Debido a la falta de tiempo, como prueba de concepto y por su fácil e ilustrativa interpretación en lenguaje natural, decidimos usar como única operación analítica, la **concatenación de operaciones**.

palabra₂ + palabra₃ - palabra₋₁ = ?

27. Disponible en: <http://www.cs.upc.edu/~nlp/wikicorpus/>

Esta operación, traducida a lenguaje natural, es una comparación del tipo “uno es a dos como tres es a...?”. El resultado de la operación presenta en pantalla las 10 palabras más cercanas a la ecuación. Por ejemplo:

Hombre es a actor como mujer es a: [actor, actriz, mujer, nominada, isbert, trudie, oscar, hodiak, haymes, karina]

Aunque hemos usado un *corpus* relativamente grande (más de 28 millones de palabras), las carencias en iteraciones suficientes al entrenar la red saltan a la vista en los resultados. Por eso, es necesario que seres humanos (idealmente del tipo “turcos mecánicos” de Amazon)²⁸ interpreten, en calidad de personas expertas, cada salida. Los criterios de interpretación que hemos establecido son los siguientes:

Bueno, cuando el resultado incluye mayoría de palabras lógicas y semánticamente correctas y que no muestran sesgos de género evidentes. El ejemplo anterior sería “bueno”.

Sesgado, cuando el resultado es lógico y semánticamente correcto, pero se evidencia un sesgo de género claro, ya sea directo o de ámbito.²⁹

No válido, cuando el resultado es absurdo.

Una vez definidos estos criterios, se procedió al análisis del modelo usando la estructura *Hombre es a palabra como mujer es a...*, y viceversa, *mujer es a palabra como hombre es a...* El conjunto de palabras usadas (decididas previamente por el y la analista) fueron 110, referidas a profesiones, actitudes, objetos y trayectoria de vida.

Análisis de resultados generales

Los resultados aquí obtenidos son solo ilustrativos de una prueba de concepto acerca de las posibilidades de uso de una red neural como herramienta de análisis de las diferencias sociosemánticas de género; su generalización exigiría un mayor volumen de datos, mayor

28. Disponible en: <https://www.mturk.com/> [fecha de consulta: 11/11/2019].

29. Por “ámbito” aquí se entiende el conjunto semántico creado por las diez palabras que aparecen como resultado de la ecuación. A menudo nos hemos encontrado con que buena parte de las palabras asociadas a mujer son de ámbito familiar, lo cual es un indudable sesgo de género porque no se da en el caso de la palabra hombre. A este sesgo, como más adelante veremos, le hemos llamado “familización”.

potencia computacional y mayores recursos. Esta prueba tiene pues limitaciones vinculadas a que los resultados obtenidos fueron analizados solo por dos personas, a lo largo de dos sesiones de tres horas, con la consiguiente fatiga. Este tipo de limitaciones son las que llevaron al equipo de Bolukbasi a optar por los microtrabajos como herramienta de análisis humano de los resultados.

La primera ronda de preguntas *Hombre es a X como mujer es a...* produjo un resultado con un 71% de coherencia (resultados buenos y sesgados a la vez), de los cuales, el 38,2% eran sesgados. La segunda ronda, *Mujer es a X como hombre es a...* produjo un 63,4% de coherencia, pero solo un 19,6% de sesgo.

La falta de precisión en las preguntas que partían de “Mujer es...” viene dada por la escasa presencia, en el *corpus*, de los calificativos femeninos y la poca densidad de estos respecto de los masculinos. Muchos de los resultados absurdos repiten ciertos nombres propios de mujeres destacadas, las que “oligopolizan” los artículos donde más aparece la palabra “mujer”. Este sesgo es lo que llamamos *sesgo de omisión*.

Análisis de sesgos

Entre los sesgos observados podemos distinguir tres categorías a las cuales hemos denominado: directo, semántico y de omisión.

Sesgo directo

En estos casos, los resultados apenas dan lugar a dudas, arrojando términos sorprendentes e incluso ofensivos. Algunos ejemplos obtenidos son:

- Hombre es a experto como mujer es a *sabelotodo*
- Hombre es a fidelidad como mujer es a *obediencia*
- Hombre es a trabajo como mujer es a *madre*
- Hombre es a muebles como mujer es a *calzado, textiles*
- Hombre es a inteligencia como mujer es a *lucirse*
- Mujer es a abogada como hombre es a *estilista*
- Hombre es a rey como mujer es a *hija, esposa*

Este tipo de sesgo es fácilmente identificable y su interpretación no requiere conocimiento previo alguno del *corpus*, por lo cual se puede recurrir a una red de microtrabajos (si se dispusiera de algo

equivalente al *Mechanical Turk* de Amazon) para su clasificación en “bueno”, “sesgado” o “no válido”.

El análisis sociológico de cada uno de estos encajes de pares puede resultar muy enriquecedor y ofrecer pistas bastante contundentes sobre los estereotipos de género que permean nuestra visión y son expresión de la sociedad. En esta investigación, sin embargo, no nos detendremos a realizar ese análisis. El reciente descubrimiento, por nuestra parte, de esta metodología no nos ha permitido avanzar más que en la puesta en funcionamiento de la red neural y en la realización de la prueba de concepto que estamos presentando. Por eso, haremos solo unas breves anotaciones en relación con los ejemplos seleccionados y su interpretación semántica, cuando esta interpretación no sea obvia. Así, de los seis ejemplos anteriores, puede ser que el cuarto y el sexto necesiten alguna aclaración. En efecto, en el cuarto ejemplo, *Hombre es a muebles como mujer es a calzado, textiles*, entendemos que se trasluce, como mínimo, la tradicional división sexual del trabajo que está en la base de la masculinización o feminización de ciertas profesiones. Los hombres aparecen asociados a la producción de muebles, a la carpintería, mientras las mujeres lo están a los textiles, su producción, confección, etcétera. Nos quedaría pendiente un análisis detallado del calzado, para el cual tendríamos que hacer otras preguntas a la red neural y buscar otras concatenaciones, proximidades, etcétera, que dieran sentido a esa asociación.

El sexto ejemplo, *Mujer es a abogada como hombre es a estilista*, es particularmente interesante y, sin duda, necesita de más preguntas. En la fase actual de desarrollo del análisis, nos atrevemos a decir que esta asociación indica un sesgo de género inverso: son los hombres los que aparecen asociados a una profesión de menor prestigio que aquella a la que aparecen asociadas las mujeres. Probablemente, esté ocasionado por el hecho de que en nuestro *corpus* además de aparecer muy pocas mujeres abogadas, están asociadas a palabras también con pocas ocurrencias, como la de estilista. En cualquier caso, apenas hemos encontrado ejemplos de ese tipo y por eso sería otro elemento que convendría explorar.

La investigación que proponemos para el futuro se basa en la realización de análisis en la línea recién sugerida. Esta línea avanzará a través de la formulación de nuevas preguntas y/o de otros análisis de

ejercicios matemáticos más complejos, que no hemos podido llevar a cabo en este esbozo de investigación. Nos referimos al estudio de las distancias absolutas entre palabras, o al listado de palabras más cercanas, técnicas que nos permitirían dar una razón más cabal de los sesgos de género en la Wikipedia en español de 2006.

Sesgo semántico

Este tipo de sesgo es más sutil y difícil de observar. Sin embargo, es aquí donde la red neural brilla por su capacidad. El sesgo más general y extendido que hemos encontrado en nuestro estudio, de más de 28 millones de palabras, es el de la *familización de las mujeres*: el término “mujer” casi siempre aparece rodeado de términos que tienen que ver con la familia, mientras que el término “hombre” aparece como una entidad independiente. Por ejemplo:

Hombre es a amor como mujer es a:
[madre, hija, pareja, hijos, esposa, hermana]

Mujer es a amor como hombre es a:
[espíritu, dios, mundo, deseo]

Hombre es a casa como mujer es a:
[madre, familia, hija, esposa, hermana]

Mujer es a casa como hombre es a:
[pueblo, tiempo, vida]

Otra modalidad de sesgo semántico que hemos podido observar es la *sexualización de los ámbitos femeninos* en comparación con los masculinos.

Mujer es a lesbiana como hombre es a:
[fetiche, *fiery*, *musings*, *spotless*, *libertine*]

De este último resultado, podemos extraer que en la mayoría de las instancias en las que se habla de “lesbiana” es en artículos relacionados con la industria pornográfica. Conviene aclarar aquí que los términos

inglés que aparecen corresponden a títulos de películas porno. Aparecen en inglés porque la Wikipedia no traduce los títulos de las películas, del mismo modo que no se traducen los nombres propios.

Indudablemente, esto es solo una aclaración de por qué hemos elegido este ejemplo. Un análisis de este exigiría el completo desarrollo de esta investigación.

Un sesgo todavía más sutil se puede encontrar en el siguiente par:

Mujer es a embajadora como hombre es a:
[Unicef, Acnur, naciones (unidas), microcrédito]

Hombre es a embajador como mujer es a:
[Tymoshenko, vizcondesa]

El par anterior nos indica que cuando le preguntamos a la red neural por una mujer embajadora, la única asociación con hombre que encuentra es cuando estos trabajan en organizaciones de buena voluntad o parecidas, como las Naciones Unidas u organizaciones benéficas, no comparable al “embajador” masculino, reservado para aquellos que ejercen de tal como profesión y para los que no encuentra la red neural nada parecido entre las mujeres, solo a una persona en concreto, Tymoshenko (suponemos que será Yulia Tymoshenko), o un título nobiliario, vizcondesa. Realmente, en este par, y sin más información, podría decirse que se dan sesgos de género semánticos, como acaba de ser explicado, y por omisión, porque no se encuentra un encaje que corresponda apropiadamente.

Como podemos ver, para analizar este tipo de sesgo hace falta un buen conocimiento sobre el *corpus* analizado y sus constricciones semánticas. Por lo tanto, no podría ser externalizado a microtrabajos tan fácilmente como el sesgo directo. Es en este nivel de análisis donde el conocimiento sociológico tiene un mayor papel.

Sesgo por omisión

Este sesgo se da cuando el peso de una de las partes del par en el corpus es tan bajo que el hecho de operar con el encaje correspondiente apenas modifica los resultados. Se puede observar en conceptos como física, matemática, química, en los cuales los resultados siempre

se refieren al significado de “ciencia física”, “ciencia matemática”, “ciencia química”, pero no a la acepción “mujer (profesional de la) física, matemática, química, etcétera”.

Este sesgo por omisión se puede observar, también, por ejemplo, en la siguiente operación:

Mujer es a reina como hombre es a:
[rey, amidala, príncipe, naboo]

Hombre es a rey como mujer es a:
[Hija, mujer, esposa]

Puede parecer que el primer componente del par arroja un resultado adecuado (rey es la primera respuesta), sin embargo, el resto de los conceptos que nos devuelve la red tienen todos que ver con la saga de películas Star Wars, proyectando la preocupante conclusión de que la mayor parte de las instancias de “reina” aparecen en artículos sobre uno de los personajes de esa saga, la Reina Amidala. De hecho, cuando preguntamos a la inversa, ni siquiera devuelve “reina”, sino que directamente habla de hija, mujer y esposa. Esto solidifica la familización detectada en el punto anterior, debido a que, en los artículos de la Wikipedia, cuando los hombres son reyes, las mujeres de su entorno no son más que familia suya.

Encontramos la misma relación de omisión al preguntar por “profesora”, obteniendo personajes de la saga de libros de Harry Potter.

Este tipo de sesgo también requiere conocimiento sobre el corpus, y su detección resultaría difícil de delegar a microtrabajadores/as estándar. En realidad, un análisis cuidadoso y profundo de los encajes neurales demandaría un equipo multidisciplinar, especialmente en el estudio de una enciclopedia, como es este caso. Por lo tanto, la materialización de este proyecto requeriría un tipo de microtarefas, encargadas a microtrabajadores/as especialistas en cada uno de los campos específicos sobre los que el equipo de sociólogas/os no tenga conocimientos suficientes. En el brevísimo análisis presentado hasta el momento, hemos necesitado conocimientos específicos sobre películas porno, *Star Wars* y *Harry Potter*.

Desgenerización de los algoritmos para eliminar o reducir los sesgos de género

A diferencia de los seres humanos, podemos enseñar rápidamente a los algoritmos a considerar y evitar los sesgos. En algunos casos es suficiente incluir el sesgo como otro indicador o si un género está sobrerrepresentado, poner un filtro que los iguale. En otros casos, sin embargo, se precisa de un trabajo más complejo de interacción entre seres humanos y máquinas. En el modelo de Bolukbasi en el que nos hemos basado, lamentablemente, word2vec arrastra cualquier sesgo que pueda tener el cuerpo de texto sobre el que trabajemos. Para combatirlo, el equipo de Bolukbasi usó la modalidad de trabajo “turco mecánico” de Amazon. Se ofrecieron pares de conceptos a diez trabajadores/as distintos para que juzgaran si los conceptos estaban sesgados por género o no. Si la mayoría decía que sí, el par de palabras se consideraba sesgado. Este método de detección aporta varios beneficios, como la diversidad de población llamada a emitir sus juicios –si se la selecciona con este criterio en mente–, o la minimización de la fatiga asociada a las tareas de procesar, de forma continuada, una gran cantidad de datos. En efecto, cada trabajador/a, solo tiene que analizar un par de palabras cada vez. Una vez localizados, por este método, los términos “deformados” por el sexismo, Bolukbasi y su equipo aplicaron varios algoritmos distintos –cuya descripción desbordaría el grado de detalle de este artículo– para eliminar el sesgo en las analogías por pares. Tras el indicado proceso de corrección de sesgo, y al volver a preguntar a la red neural sobre los mismos conceptos problemáticos, el número de analogías sexistas pasó de 30 en 150 a 7 en 150.

El enfoque de Bolukbasi fue el de conseguir la corrección de la red neural por sí misma, una vez terminado su entrenamiento. Esto es, una vez detectados los sesgos en los pares de palabras y corregidos matemáticamente, la base de datos queda casi desprovista de sesgos, aunque la versión textual de la que procede esa base no se altere. Esto es, las noticias recogidas en *Google News* se mantuvieron tal cual, y así siguen, pero la red neural resultante se desprende de la mayoría de sus sesgos de género. Dicho de otro modo: el texto fuente se mantiene idéntico, pero la base de datos en la que se convierte, recomputa

las relaciones entre las palabras de ese texto, de tal forma que apenas tiene sesgos de género en el nivel de esa base de datos. Obsérvese que este nivel es el realmente operativo a efectos prácticos (por ejemplo, cuando se toman decisiones “haciendo hablar” automáticamente a esa base de datos. Viene a ser como si se le explicara a un niño pequeño que robar es malo, aunque su padre lo hiciera todos los días. Esto no evita ni corrige la actitud del padre, pero sí educa a la criatura para que, en un futuro, no considere que robar es bueno.

En nuestro caso, en el análisis de la Wikipedia, aunque nos gustaría, no intentamos corregir los sesgos que encontramos porque no tenemos la plataforma adecuada ni los recursos para alquilar la plataforma *Mechanical Turk* de Amazon. Pero sabemos que los sesgos pueden eliminarse y sin que cambie el texto de la Wikipedia. Por pensar en una aplicación más próxima al ámbito judicial: sin cambiar el texto de todas las sentencias de un juzgado, se pueden corregir las correlaciones que se establezcan entre pares de conceptos para que dejen de estar sesgadas por género. Esto es muy importante ya que, en breve, se utilizarán los *big data* para determinar el sentido de las sentencias o de los diagnósticos médicos o de la concesión de una hipoteca y más vale que, aunque los datos se hayan heredado de un contexto social sexista, dispongamos de algoritmos para corregirlos.