

Ensayo

Pensar la tecnología digital con perspectiva de género

Grupo GIFT

**Tomás Balmaceda, Karina
Pedace, Diego Lawler, Diana
Pérez y Maximiliano Zeller**

Noviembre 2021

Pensar la tecnología digital con perspectiva de género

Autores: Grupo GIFT: Tomás Balmaceda, Karina Pedace, Diego Lawler, Diana Pérez y Maximiliano Zeller

Coordinadoras: Carolina Aguerre y Maia Levy Daniel

Diseño: Mónica Castellanos

Licencia Internacional Pública de Atribución/ReconocimientoNoComercial-SinDerivados 4.0 de Creative Commons.



Las opiniones expresadas en las publicaciones incumben únicamente a los/as autores/as. No tienen intención de reflejar las opiniones o perspectivas del CETyS, CLD ni de ninguna otra organización involucrada en el proyecto.

Pensar la tecnología digital con perspectiva de género

Noviembre 2021

Grupo GIFT

Tomás Balmaceda, Karina Pedace, Diego Lawler, Diana Pérez y Maximiliano Zeller

Contenido

1. Introducción: Tiempos interesantes para pensar	5
2. Pensar la tecnología digital	7
2.1. Pensar la tecnología sin determinismos	7
2.2. Pensar la tecnología sin neutralidad valorativa	10
2.3. Pensar la tecnología con perspectiva de género	11
3. Desarrollos en IA: géneros, sesgos y discriminación	18
3.1. De qué hablamos cuando hablamos de algoritmos.....	18
3.2. De qué hablamos cuando hablamos de sesgos	18
3.3. Algunos sesgos en los sistemas de algoritmo en sentido amplio	20
4. Los sesgos en acción	27
4.1. Sistemas de recomendaciones	27
4.2. Sistemas que involucran el procesamiento del lenguaje natural.....	32
5. Conclusiones.....	39
6. Bibliografía.....	41

1.

Introducción: Tiempos interesantes para pensar

Vivimos en tiempos interesantes: se calcula que en enero de 2022 casi cinco mil millones de personas en todo el globo -esto es, alrededor del 60% de la población- se habrá conectado al menos una vez a la semana a Internet, la mayoría de ellos mediante un teléfono inteligente¹. Una de las actividades predilectas de estas personas cuando están en línea es conectarse a una red social: Facebook cuenta en junio de 2021 con 2.853 millones de perfiles activos mensuales, seguido por YouTube, con 2.291 millones e Instagram, 1.386 millones. WhatsApp, que nació como un mensajero instantáneo pero fue mutando a una plataforma de contenidos que incluye grupos y la posibilidad de compartir “historias”, es utilizado por 2000 millones de personas. Ninguna de estas redes sociales existía hace dos décadas cuando las personas que se conectaban a Internet rondaban los 533 millones, apenas el 8.7% de la población²... ¿qué sucedió en ese lapso de tiempo? Sin dudas una serie de transformaciones muy profundas que impactaron en todos los sectores de la sociedad, incluso los más tradicionales, y que modificaron nuestra vida cotidiana, la manera de vivir nuestros vínculos sociales, familiares y sexo-afectivos y que incluso inciden en nuestra identidad.

Sin embargo, estas transformaciones no implican que se hayan dado cambios radicales en el balance de las fuerzas dentro la sociedad. De hecho, y en muchos sentidos, se consolidó el *status quo* vigente hasta entonces, con situaciones de inequidad e injusticia por parte de los actores poderosos en detrimento de sectores marginalizados o minimizados.

En este trabajo queremos llamar la atención, justamente, sobre cómo la tecnología se ha vuelto un facilitador más de opresión contra un sector de la sociedad. Nuestro foco estará puesto, específicamente, en las tecnologías digitales y las cuestiones de géneros. Nos referimos con la primera expresión a un amplio subconjunto de artefactos tecnológicos que utiliza microprocesadores y códigos para crear aplicaciones a partir de sistemas de algoritmos y que típicamente llegan a la gente en computadoras, teléfonos inteligentes y otros tipos de dispositivos similares. Y hablaremos de cuestiones de géneros para referirnos a una mirada crítica de las consecuencias que tiene en el entramado social el sistema sexo-genérico hegemónico, una construcción sociocultural que asigna significados y valores a las personas por características como su sexo, identidad y orientación sexual. Preferimos hablar de “géneros” y no de “género” para poner explícitamente en cuestión la manera binaria en la que este sistema de representaciones asigna características sociales, culturales, políticas, jurídicas y económicas a la esfera “de lo masculino” y “de lo femenino”.

Con este objetivo en mente, este trabajo tiene tres partes. En la primera brindaremos un breve marco conceptual para poder encuadrar nuestras reflexiones dentro de esquemas más amplios de la filosofía de la tecnología. En este planteo haremos una elucidación conceptual de un concepto clave, el de algoritmo, y esbozaremos un posible mapa de las corrientes surgidas

¹ Hootsuite's Global Report <https://www.hootsuite.com/es/recursos/tendencias-digitales-2021>

² Peterson, M. P. (Ed.). (2005). Maps and the Internet. Elsevier.

del encuentro entre tecnología y géneros. Luego, en la segunda parte, analizaremos el concepto de sesgo y cómo gravita en los desarrollos tecnológicos, potencialmente perjudicando a aquellos sectores que no son los dominantes. Una vez que contemos con estas herramientas las usaremos, en la tercera parte, para analizar casos concretos y recientes, con la esperanza de ponerlos bajo una nueva luz, una que habilite nuevas ideas y, tal vez, el inicio del camino que nos lleve a respuestas y soluciones.

Somos un grupo de filósofas y filósofos que pensamos estas cuestiones desde la Argentina mientras atravesamos la crisis provocada por la pandemia de COVID-19. Si bien nos proponemos inspirar reflexiones y conversaciones que excedan lo coyuntural, no podemos evadir nuestras circunstancias. Creemos que no se pueden entender los cambios sociales en el vacío así que decidimos echar mano a casos reales para ilustrar algunas de nuestras reflexiones. En todas las instancias intentamos hacer una reconstrucción fiel de los hechos, pero nuestro trabajo dista de ser histórico o periodístico de modo que estos sucesos servirán como forma de acceso a las temáticas que queremos abordar. Además, como creemos que la verdadera reflexión surge del análisis y el diálogo, decidimos convocar a distintas figuras de la región para que nos enriquezcan con sus comentarios y pensamientos. Pero este texto coral no busca clausurar nada, sino que invita a que se nutra de cada lectura y abra nuevas preguntas y caminos. Queda hecha, así, la invitación a pensar en comunidad.

2.

Pensar la tecnología digital

2.1. Pensar la tecnología sin determinismos

Hay un hecho incontestable: el impacto de las innovaciones tecnológicas recorre todos los niveles de nuestra vida cotidiana; no hace falta más que reparar en los teléfonos inteligentes y el siempre creciente menú de aplicaciones, en las posibilidades que nos ofrece internet y la web, en las computadoras personales, los automóviles o los artefactos domésticos, por nombrar algunos ejemplos. Nuestra vida cotidiana se transforma constantemente por la incorporación, difusión y uso de distintas tecnologías.

Los cambios tecnológicos o la introducción de innovaciones tecnológicas parecen estar asociados a la materialización de cambios sociales. Sin embargo, hay quienes van más allá de la caracterización de los impactos de la tecnología en la vida cotidiana y sostienen que la tecnología es una fuerza única que posee la capacidad de cambiar nuestras concepciones sobre el mundo, transformar las estructuras y el funcionamiento de nuestras instituciones, modificar las relaciones de producción y trastocar la totalidad de la estructura social, alterando los vínculos económicos y las relaciones de poder.

Quienes piensan de este modo sostienen la tesis de que la tecnología es el agente causal más relevante de los cambios sociales. Esta tesis se conoce como la tesis del determinismo tecnológico y supone una interpretación específica de la relación entre tecnología y sociedad.³ Según ella, la tecnología es considerada un factor externo a la sociedad, que le da forma a la sociedad no siendo parte de ella. En este sentido la tecnología sería una fuerza autónoma que se ejerce unidireccionalmente, que condiciona pero que no está condicionada.⁴ La lógica interna del desarrollo de la tecnología, esto es, el cambio tecnológico, no se encontraría influida por la intervención humana individual o social. De esta sugerencia se sigue que si la tecnología es una fuerza que se desarrolla autónomamente, entonces se halla fuera del control humano. Las personas solo tendrían que lidiar con las consecuencias del inexorable desarrollo tecnológico pero no podrían modelarlo según sus deseos, intereses y valores.⁵ Estos discursos sobre el carácter imparabile e inevitable del desarrollo tecnológico arrinconan la posibilidad de intervenir o regular la generación de innovaciones tecnológicas, dándoles a ellas un aire fatalista que alimenta tanto los imaginarios de los enfoques tecnofóbicos así como los de los tecnofílicos, generando o distopías sombrías o utopías diáfanas.

Sin embargo, el constructivismo social de la tecnología ha mostrado empíricamente, a través de innumerables estudios de caso, que las tecnologías no determinan unidireccionalmente ni

³ El sociólogo y economista norteamericano Thorstein Veblen (1857-1929) fue el primero en emplear esta expresión. Las variantes del determinismo tecnológico están recogidas en Smith y Marx (1997).

⁴ Hay un paralelo con el determinismo climático del siglo XIX, que afirmaba que el clima, un factor independiente y externo sobre el que las sociedades no tienen influencia, daba forma a la naturaleza de la sociedad.

⁵ El locus clásico de esta tesis es Winner (1979).

cambios ni formas sociales, sino que sus efectos dependen precisamente de las configuraciones sociales y culturales en que tienen lugar; además, estas configuraciones no solo catalizan los efectos de las tecnologías, sino que influyen significativamente en los desarrollos tecnológicos mismos. Hay factores sociales, culturales, organizativos y políticos, entre otros, que influyen en los desarrollos y las innovaciones tecnológicas.⁶

Estos estudios de caso resaltan el carácter contingente y situado de los desarrollos tecnológicos; al mismo tiempo, muestran las fuertes influencias del contexto en el diseño y desarrollo de las tecnologías.⁷ El análisis de las controversias tecnológicas (episodios donde se plantean varias opciones tecnológicas para resolver un problema similar o en los que una innovación se propone como alternativa a una tecnología existente) evidencia la injerencia de factores de múltiples naturalezas y proveniencias en la estabilización y predominancia de una solución tecnológica frente a otras alternativas.⁸

La moraleja de los estudios empíricos del constructivismo social es que el cambio tecnológico no se inscribe en un sendero autónomo de desarrollo, sino que se encuentra modelado por diferentes agentes y factores heterogéneos, donde los factores estrictamente técnicos no poseen necesariamente el peso decisivo. Como señalan los estudios constructivistas, las tecnologías son productos contingentes que reflejan el medio social en que fueron producidas, y se hallan modeladas por un interjuego de factores técnicos, profesionales, políticos, intereses, sesgos, creencias, herramientas de diseño, conocimientos, materias primas disponibles, habilidades, etc.⁹

Un aspecto interesante de esta perspectiva de análisis es que convierte a las tecnologías en algo completamente heterogéneo. El análisis empírico de los procesos de innovación tecnológica muestra que los sistemas y artefactos tecnológicos son entidades que incorporan contenidos tecnológicos, así como tensiones y compromisos sociales, económicos, políticos, profesionales, conocimientos especializados, intereses y sesgos de distintas procedencias, habilidades y expectativas. De esta manera, el mundo tecnológico y el mundo social están ensamblados en los sistemas y artefactos tecnológicos.

Ahora bien, esto no significa que la emergencia de ciertas tecnologías no tenga consecuencias sociales. Tampoco implica suscribir la sugerencia de que las innovaciones tecnológicas se expliquen completamente por causas sociales. Rechazar la tesis del determinismo tecnológico no supone suscribir la tesis del determinismo social; algo que ocurriría si se afirmara lo anterior. Ambos determinismos perciben a la tecnología como una caja negra (es decir, un proceso inaccesible del que solo se puede conocer su resultado final pero no los procesos intermedios), que responde a condicionamientos sociales, por ejemplo, económicos, culturales, políticos o ideológicos, o que meramente genera impactos sociales de diferente clase.

Evitar los determinismos conlleva situar la emergencia de los desarrollos tecnológicos en con-

⁶ Cfr. Aibar y Bijker (1997).

⁷ Hughes (1983), por ejemplo, estudia los sistemas tecnológicos de generación y transporte de energía eléctrica de varias ciudades y pone en evidencia cómo una tecnología similar se desarrolla y difunde en contextos sociales diferentes.

⁸ Véase, por ejemplo, el estudio de Pinch y Bijker (1987) donde se analizó el desarrollo de la bicicleta durante la segunda mitad del siglo XIX.

⁹ “Se dice a veces que tenemos los políticos que nos merecemos. Pero si esto es verdadero, entonces también tenemos las tecnologías que nos merecemos. Nuestras tecnologías reflejan nuestras sociedades. Reproducen y encarnan la interacción compleja de factores profesionales, técnicos, económicos y políticos” (Bijker y Law, 1992).

textos específicos y teniendo lugar a través de un conjunto de mediaciones.¹⁰ La tecnología y los sistemas tecnológicos no son externos a la sociedad: por el contrario, son endógenos a ella. La tecnología y el medio social y cultural se reflejan mutuamente, estableciéndose entre ambos polos una relación bidireccional y compleja, de co-producción. Por consiguiente, así como la sociedad no está configurada por los desarrollos tecnológicos tampoco es el caso de que la tecnología sea un producto exclusivo de la sociedad, puesto que los agentes sociales (individuales y colectivos) interactúan al interior de marcos tecnológicos determinados. Así, la medida en que la tecnología está socialmente configurada es análoga a la medida en la que la sociedad está tecnológicamente mediada.



Danila Suárez Tomé

Instituto de Investigaciones Filosóficas-SADAF-CONICET/ UBA

<https://www.aacademica.org/danila.suarez.tome>

“En primer lugar, es bueno tener en cuenta que las relaciones entre la tecnología y la configuración del género y de las relaciones generizadas es mucho más profunda de lo que se puede llegar a suponer. Tenemos la idea de que la tecnología puede reproducir, perpetuar y amplificar diferencias de géneros pre-existentes “naturalmente”, pero nos cuesta más entender cómo puede generarlas en sí misma, y hasta ser una parte constitutiva del género entendido como un dispositivo de poder social que produce determinados individuos que se inteligibilizan y ordenan en una matriz de sexualidad cisheteronormada. Teniendo eso en cuenta, creo que la idea harawayana de comprender a los fenómenos de la realidad como nudos semiótico-materiales, en donde no se puede pensar sólo en términos de naturaleza o de cultura, sino en términos epistémica y ontológicamente híbridos de naturculturas, constituye una heurística poderosa para indagar y problematizar la intersección entre género, tecnología y sociedad”.

“Mi mayor inquietud, pensando desde esta base, es cómo el avance tecnológico en países con mayor desarrollo puede impactar en las ontologías y epistemologías sociales de modo tal en que se genere una diferencia cada vez más marcada con aquellos países en donde las condiciones materiales son más desventajosas. El poder social, y particularmente en el caso del género el poder social identitario, ha tenido siempre un impacto profundo en la constitución de ontologías sociales jerarquizadas y opresivas. Es de esperar, por lo tanto, que un desarrollo geopolíticamente desigual del avance tecnológico tenga efectos de desigualdad en las ontologías sociales, siendo el género una de sus aristas más estructurantes”.

¹⁰ Un ejemplo estudiado es el de las diferencias en el diseño de las barreras antimarea desarrolladas en la costa holandesa y la construidas en la costa inglesa, a solo unas pocas decenas de kilómetros de distancia, en circunstancias geofísicas muy similares y bajo un mismo corpus de conocimientos científicos en dinámica de fluidos o en resistencia de materiales (Bijker, 1994).

2.2. Pensar la tecnología sin neutralidad valorativa

Este ensamblaje entre lo social y lo tecnológico que tiene lugar en las prácticas humanas de producción del mundo artificial pone en cuestión la tesis de que las tecnologías son valorativamente neutrales. Se trata de la creencia ingenua que suele expresarse con la frase popular “la tecnología no es ni buena ni mala: depende de cómo se la use”. Sin embargo, como se ha mostrado arriba, las tecnologías configuran las distintas maneras en que actuamos y experimentamos el mundo; por ejemplo, las tecnologías de diagnóstico prenatal dan forma a las decisiones que tomamos sobre la continuidad o no de un embarazo y las redes sociales construyen un medio ambiente con características propias y configuran nuestras interacciones con otras personas. Así, las tecnologías median nuestras relaciones con el mundo y entre los seres humanos, contribuyen a dar forma a nuestras percepciones, acciones, experiencias y prácticas.

Diseñar un artefacto o sistema tecnológico supone diseñar al mismo tiempo a quienes van a utilizarlo, es decir, a las interacciones entre seres humanos y artefactos, así como sus experiencias y percepciones en una situación material y social determinada. Imaginemos, por ejemplo, artefactos coercitivos como los molinetes que impiden el ingreso al subterráneo sin el ticket correspondiente o los autos que no encienden si la conductora no tiene puesto su cinturón de seguridad, así como tecnologías más persuasivas como las aplicaciones para controlar el tiempo que pasamos en nuestro teléfono inteligente o las aplicaciones de control de ingesta de calorías. Pero también es cierto que cuando usamos un artefacto podemos rediseñarlo de maneras no previstas, reconfigurando los contextos de interacción, experiencias y prácticas. Por consiguiente, los diseños, además de rasgos estéticos, funcionales y materiales, incorporan valores desde su diseño y, posteriormente, se nutren de valores adicionales a través de lo que hacen las personas con las tecnologías en sus vidas diarias.¹¹

Los valores son incorporados a los artefactos y sistemas tecnológicos a través del diseño y sus planes de usos. Diseñar es una actividad humana que consiste en realizar una operación cognitiva-conceptual por medio de la cual se crean modelos de estructuras materiales y mecanismos que realicen funciones propuestas por los diseñadores para resolver una situación que está planteada como un problema. Esta operación cognitiva-conceptual comprende representaciones de opciones de estructuras materiales y mecanismos, así como de cursos de acción, deliberación sobre esos cursos de acción, estructuras materiales y mecanismos, establecimiento de ponderaciones y elección de unos en desmedro de otros de acuerdo a un conjunto de valores y objetivos previamente formulados.¹² Por ejemplo, consideremos a los algoritmos sobre los que se asientan muchas de las aplicaciones disponibles en el mercado. Los parámetros operacionales son especificados por las personas encargadas del desarrollo y son configurados por la gente que los adoptan con ciertos resultados deseados en mente, que privilegian algunos valores e intereses sobre otros. Así hay algoritmos que arman perfiles que terminan discriminando a distintos individuos y/o grupos sociales. Los algoritmos están cargados de valores.¹³

¹¹ Este punto de vista es una advertencia al error que supone considerar la funcionalidad y el uso en abstracto de las prácticas humanas de diseñar esas funciones e imaginar y realizar esos usos. Cuando esas relaciones se consideran en abstracto y se descontextualizan, se produce el error de considerar a las tecnologías como meros medios neutrales que se inscriben en planes de acciones completamente externos a ellas.

¹² Cfr. Caja de herramientas humanísticas (2019) <https://guia.ai/wp-content/uploads/2020/05/Caja-de-herramientas-Humanistas.pdf>

¹³ Es posible pensar en otros ejemplos más allá del campo de las tecnologías digitales: los diques holandeses, por ejemplo, están diseñados para la seguridad de los habitantes de las zonas inundables o los reductores de velocidad, comúnmente conocidos como ‘lomos de burro’, que se disponen en las calles donde hay escuelas y que fueron especialmente diseñados para disminuir la velocidad de los vehículos que circulan por esa calle y brindar seguridad a quienes ingresan y salen de esas escuelas. En ambos casos, la seguridad de las personas que viven en las zonas inundadas o que transitan por esas calles está incorporada en las funciones de los artefactos diseñados. Otro ejemplo muy conocido son los puentes diseñados por Edgar Moses en Long Island, Nueva York, que impedían el paso del transporte público pero sí de vehículos particulares, imposibilitando que las poblaciones con menor poder adquisitivo pudieran disfrutar de la playa (Winner, 1980).

La elección del diseño de un artefacto o proceso tecnológico para su producción e introducción en el mundo real es una respuesta a la pregunta de cuál es la solución adecuada a un problema determinado en cuya definición intervienen muchos actores que despliegan diferentes creencias e intereses. Esta pregunta es importante porque conduce a interrogarse cómo se toman las decisiones tecnológicas, qué dimensiones intervienen en la creación de un artefacto o sistema tecnológico, y cómo se valoran esas dimensiones en la decisión tomada en relación con los intereses de las personas intervinientes, los problemas que resuelven y las modificaciones que introducen en el mundo. Cabe señalar que en el contexto deliberativo del diseño intervienen valoraciones intrínsecas y extrínsecas. Las valoraciones intrínsecas son propiamente tecnológicas y las valoraciones extrínsecas conciernen a los propósitos humanos incorporados en los artefactos y sistemas tecnológicos.¹⁴

Las valoraciones intrínsecas están vinculadas a la eficacia y eficiencia de los artefactos y sistemas tecnológicos, para que funcionen y lo hagan con el menor coste posible de recursos (materiales, económicos, etc.) y consecuencias no queridas (por ejemplo, matar moscas a cañonazos es efectivo, pero tiene una multitud de consecuencias no queridas). Las valoraciones extrínsecas están relacionadas con lo que significa que el artefacto o sistema tecnológico sea una solución adecuada a un problema planteado. Esto no depende solamente de que el artefacto o sistema sea eficaz y eficiente; por el contrario, depende de cómo se inserta en el entorno de las personas y de una sociedad determinada, y de la valoración que hacen quienes lo diseñan, quienes lo producen, quienes lo usan, etc.; una valoración que es siempre situada, donde cuentan más los valores relacionados con cómo se quiere vivir que los aspectos estrictamente tecnológicos. Esta valoración está fuertemente asociada a las particularidades de los grupos sociales así como a sus identidades prácticas y, por tanto, a sus fines, intereses y puntos de vista sobre el mundo: por ejemplo, un grupo social impulsado por una fuerte conciencia y compromiso ecológico puede rechazar la instalación de una central nuclear en su región, a pesar de que ésta sea la opción tecnológicamente más adecuada para resolver el suministro de energía en esas condiciones socioeconómicas y ambientales.

2.3. Pensar la tecnología con perspectiva de género

Las relaciones entre la tecnología y los géneros son un ámbito donde se muestra la configuración social de la tecnología, así como la mediación tecnológica de la sociedad. La tecnología y la sociedad, mutuamente ensambladas, configuran prácticas donde las diferencias de género son modeladas, ejemplificadas y difundidas por la tecnología.



Cristina Bernabéu Franch

Universidad Autónoma de Madrid

<https://portalcientifico.uam.es/ipublic/agent-personal/profile/iMarinaID/04-271191/name/BERNABEU-FRANCH-CRISTINA>

¹⁴ Cfr. Caja de herramientas humanísticas (2019) <https://proyectoguia.lat/wp-content/uploads/2020/05/Caja-de-herramientas-Humanistas.pdf>

“Los entornos artefactuales que habitamos actualmente, incluidos los digitales, no son neutrales desde el punto de vista del género; están connotados masculina y/o femeninamente. Esta hipótesis no está sin embargo comprometida únicamente con el hecho de que los artefactos están generizados, en tanto que reproducen el género, sino especialmente con la idea de que lo producen —esto es, que “generizan”.

En este sentido, se puede afirmar que colaboran activamente en el establecimiento de la división y la diferencia sexual.

El hecho de que los artefactos técnicos sean un factor central en la generización de la sociedad y de la cultura, pone de manifiesto la importancia y la urgencia de intervenir sobre la realidad artefactual para transformar las relaciones de poder desigualitarias y perseguir la justicia social.

Es en este marco en el que creo que hay que comprender el análisis sobre las diferencias que toma el proceso de generización de la tecnología entre Europa y Latinoamérica. Si bien los movimientos feministas comparten, desde una perspectiva global, objetivos comunes (que tienen que ver con la igualdad entre personas), sus luchas adquieren también características propias, en función del contexto en el que se insertan. En este sentido, creo que en Latinoamérica, en comparación con Europa u otros lugares del Norte global (en los que se está produciendo un proceso de individualización cada vez mayor) el movimiento feminista está sabiendo utilizar las tecnologías digitales, no sólo para mantener y reforzar los espacios comunitarios de resistencia, sino también para generar procesos de toma de conciencia dentro del feminismo”.

Las relaciones entre la tecnología y los géneros fueron estudiadas desde los años 70 como parte de la división sexual del trabajo, donde ésta era considerada también una división sexual técnica. Los artefactos hogareños, por ejemplo, fueron fuente y signo de la dominación patriarcal (Bernabéu 2020), estructurando tareas percibidas como improductivas, al mismo tiempo que se infravalora ese conjunto de artefactos tecnológicos, diseñados para tareas “sin épica” (Zafra 2013).

El constructivismo social de la tecnología, retratado más arriba, señala que los artefactos, los sistemas tecnológicos y los géneros se construyen mutuamente. Los artefactos, como los relojes, los teléfonos inteligentes, las píldoras anticonceptivas, la vestimenta y los sistemas tecnológicos invisibilizan o hipervisibilizan ciertos grupos sociales; no solo expresan una realidad generizada, esto es, no son solo productos de unas relaciones sociales generizadas previas, sino que dan forma a una realidad generizada, esto es, una realidad organizada alrededor de la diferencia sexo-genérica.

Las investigaciones sobre las relaciones de la tecnología con los géneros han estado motivadas especialmente por el análisis de las maneras en que la tecnología produce una realidad generizada y han estado promovidas en su mayoría por investigadoras y activistas feministas que explicitan la asociación del desarrollo tecnológico con el género masculino y conciben a la tecnología como un potente instrumento para ejercer una distribución asimétrica de poder entre los géneros. Desde esta concepción se advierten, por ejemplo, los diversos modos en que

la tecnología influye en el mercado de trabajo y genera y naturaliza situaciones de explotación y exclusión de las mujeres en la sociedad actual (como las que se reflejan en los distintos tipos de segregación ocupacional, tales como la horizontal, con sectores altamente feminizados como los vinculados con tareas de cuidados y la segregación vertical, que impacta en las posibilidades de ascenso a cargos jerárquicos de toma de decisión; en la exclusión de las mujeres de ciertos ámbitos educativos; o en la promoción de ciertos cuerpos como objeto de consumo).



Natalia Monti y María Julia Giorgelli

**Centro de Protección de Datos Personales
Defensoría del Pueblo de la Ciudad de Buenos Aires**

<https://cpdp.defensoria.org.ar/>

“Cuando nos sumergimos en el campo de la Ciencia y Tecnología en Argentina es llamativo observar que las posiciones de poder de decisión están ocupadas en un 76% por hombres, cuando bajamos de escalafón este porcentaje se reduce, pero siguen siendo espacios dominados, en general, por hombres¹⁵.

Estos datos son aportados solo como una muestra que creemos nos define socialmente. En este sentido, entendemos que cuando trasladamos esa realidad a otros sectores se mantiene esta correlación. Es decir, si nos hallamos con una materialidad concreta en los espacios de la ciencia y la tecnología todo lo que se construya en base a eso se encontrará sesgado: desde las líneas de poder de decisión hasta los diseños de programas de inteligencia artificial (IA)”.

Desde los años 80 en adelante, los estudios feministas sobre la tecnología suscriben la premisa de que la tecnología no es neutral desde el punto de vista del género; por el contrario, las cuestiones de género anidan en el desarrollo de los sistemas tecnológicos o artefactos desde el mismo momento de su diseño hasta su implementación, difusión y uso en los distintos grupos sociales (Wajcman, 1991, 2004 y 2009).¹⁶

15 Datos propios recolectados y sistematizados de información que surge de sitios web oficiales del Gobierno Nacional, la Ciudad Autónoma de Buenos Aires y las provincias de la Argentina, en diciembre de 2021.

16 Wajcman (1991) realizó estudios donde se preocupa por mostrar que los artefactos tecnológicos cotidianos, por ejemplo, los anticonceptivos y los automóviles, están generizados, esto es, conformados por las relaciones, las identidades y los significados de género. Las jerarquías de las diferencias sexuales anidarían en el mismo diseño, difusión y uso de las tecnologías. A su vez, Bernabéu (2020) defiende que los artefactos “invitan, proponen, incitan o imponen la realización de una serie de actividades, prácticas e interacciones específicamente generizadas -vinculadas a la reproducción, la sexualidad, el aspecto físico, la apariencia o la orientación sexual- (...) Hay artefactos masculinos y artefactos femeninos y artefactos que son utilizados masculina o femeninamente”.



Laura Ación

Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires-CONICET, Argentina

Consultora técnica de ARPHAI - Gestión epidemiológica basada en inteligencia artificial y ciencia de datos

<http://www.ciecti.org.ar/en/arphai/>

“La acción principal para el desarrollo de tecnologías digitales con perspectiva de género en Latinoamérica es incluir recursos humanos interseccionales en los equipos que desarrollan tecnología. Necesitamos equipos de trabajo diversos que sean lo más parecido posible a la sociedad para la cual se crean las tecnologías digitales. Es necesario incluir en posiciones tanto técnicas como de toma de decisiones personas cuyas intersecciones no están habitualmente representadas en el desarrollo tecnológico, tales como mujeres que se identifican con pueblos originarios, hombres trans, migrantes o personas con discapacidad, entre otras muchas intersecciones. Suele argumentarse que no hay recursos humanos interseccionales suficientemente formados para participar de estos equipos. Sin embargo, el número de recursos humanos altamente capacitados se amplía cuando se considera la multidisciplinaridad necesaria para desarrollar tecnología”.

“Toda acción que allane el camino e incentive la formación de recursos humanos interseccionales para el desarrollo de tecnologías digitales es urgente”.

Los temas de las agendas de investigación varían según las diferentes motivaciones de los movimientos feministas, dirigiéndose a los estudios empíricos sobre cómo los sistemas o artefactos tecnológicos construyen usos generizados y perfiles de usuarias, o al modo en que las tecnologías generan asimetrías y reproducen desigualdades de género corporizando y distribuyendo poder, o volcándose al activismo político con el propósito de modelar las nuevas tecnologías para ponerlas al servicio de la liberación de los géneros oprimidos (Vergés Bosch, 2013).¹⁷ En algunos casos, la revisión crítica e impugnación de los sistemas tecnológicos patriarcales van de la mano de un programa de acción política de sustitución del régimen económico, social y político capitalista actual, puesto que se sostiene que la tecnología “encarna y (re) produce las desigualdades de género en múltiples intersecciones con otras opresiones”, por ejemplo, opresiones de clase social y económica (Torrano y Fischetti, 2020).¹⁸

17 El trabajo de Vergés Bosch muestra cómo antes de los años 80, los trabajos feministas sobre tecnología presuponían en su mayoría la neutralidad de la tecnología y cuestionaban su implementación para acentuar las desigualdades de género y fortalecer el patriarcado. Gran parte de los estudios estaban dirigidos a mostrar la exclusión o subrepresentación de las mujeres en el desarrollo de las tecnologías.

18 Crenshaw (1989) acuñó el término “interseccionalidad” para referirse al modo en que el poder ejercido sobre las minorías y las desventajas que éstas sufren se expresan y legitiman a través de las múltiples fuerzas sociales, identidades sociales e instrumentos ideológicos.



Constanza Gómez Mont

Fundadora y Presidenta de C Minds

<https://www.cminds.co/>

“El objetivo principal de alcanzar la igualdad en las tecnologías digitales debe venir desde la transversalidad. La desigualdad va más allá de la discriminación algorítmica, esta diferencia va desde el acceso a una educación de calidad, hasta los sesgos de género que existen alrededor del desarrollo profesional de las mujeres latinoamericanas. La redefinición de las estructuras de poder desiguales es fundamental, así como una visión más holística de la problemática que incluye estrategias para tener y retener a más mujeres en puestos de liderazgo dentro de la economía digital, la generación de incentivos y estrategias para proyectos de educación de tecnologías digitales dirigidos específicamente a las mujeres latinoamericanas y otras acciones puntuales que permitan atender los retos desde una perspectiva multidimensional.

Solucionar la falta de participación femenina en la fuerza laboral de la región es fundamental para alcanzar el objetivo final, la equidad de género. La diferencia dentro del campo de las tecnologías digitales es muy grande; menos del 14% de los investigadores de IA, menos del 1% de las empresas que contratan expertos en IA son mujeres y solamente el 22% de las personas expertas en toda el área son mujeres”.

A partir de los años 90, con la irrupción de las tecnologías de la información y la comunicación (TIC), y enfáticamente en las últimas dos décadas, con la presencia ubicua de esas tecnologías en la vida cotidiana, los estudios feministas se concentran en explotar la oportunidad que representan las TIC para el rediseño, reapropiación y agenciamiento de estas tecnologías con la finalidad de hibridar las filiaciones identitarias, deconstruir las categorías de género y empoderar a las mujeres para el protagonismo a través del desarrollo de estas tecnologías. En la última década, este movimiento se alimenta de las teorías postfeministas, un movimiento que cuestiona el feminismo de la igualdad y se abre a la reivindicación de identidades que trascienden categorías prefijadas binarias como sexo y género. Así emergen diferentes movimientos como las cyborgfeministas, las ciberfeministas, las tecnofeministas y las queer feministas, entre otras (Vergés Bosch 2013).

Las cyborgfeministas están inspiradas en el trabajo pionero de Donna Haraway sobre la condición cyborg (1985). Las imágenes más extendidas en la cultura occidental asocian a la naturaleza con una feminidad pasiva, mientras que vinculan las producciones científicas y tecnológicas con una cultura masculina activa. Así, la empresa científica y tecnológica, “objetiva y libre de valores”, se encontraría anclada en la esfera de lo masculino. Sin embargo, la condición cyborg rompe esta dualidad y ese estereotipo. Esta figura híbrida de organismo y máquina es empleada para desarmar la dualidad natural-artificial y deconstruir otras relaciones binarias como la

de mujer-hombre. Lo híbrido y no lo factorizado en polos puros es lo que está extendido en el mundo; la condición cyborg deja de lado los dualismos y funda la posibilidad de explorar políticamente los espacios de liberación abiertos por esta constitución híbrida.

Las ciberfeministas no solo cuestionan las identidades de género, atacando todo esencialismo, sino que, al mismo tiempo, combaten las desigualdades entre los géneros en el desarrollo y acceso a las TIC, con el objetivo de transformar social y políticamente las estructuras económicas y de poder sobre las que se asientan esas tecnologías (Plant 1997). La premisa subyacente en este impulso transformador consiste en ver a la revolución digital como una oportunidad para desarticular las prácticas y fundamentos de poder institucional patriarcal. De este modo se percibe al ciberespacio como un ámbito donde pueden desnaturalizarse y desbiologizarse las diferencias actuales entre los géneros, generando un poder liberador de las relaciones sociales actuales para las mujeres (Sollfrank et al. 1999).

El tecnofeminismo interpela especialmente las relaciones de género en los procesos de diseño, difusión y uso de las tecnologías (Wacjman 2004). Para esta corriente son los propios artefactos y sistemas tecnológicos los que están modelados e impregnados de relaciones, significados e identidades de género (Wacjman 1991). Por consiguiente, las asimetrías y jerarquías provenientes de la diferencia sexual, naturalizada en la sociedad patriarcal, intervienen en el diseño de la tecnología, constriñen su desarrollo y condicionan sus usos y difusión. El tecnofeminismo suma el abordaje de la biomedicina al estudio de las tecnologías de la información y la comunicación.

Las postcolonialistas critican los impulsos etnocéntricos presentes en algunos movimientos feministas, que universalizan una mirada occidental anclada en un cierto grupo sociocultural y económico (Gajjala 1999; Galpin 2002). En esta crítica a la universalización de un conjunto de valores particulares, abogan por realizar estudios feministas localizados, donde se ponga en evidencia que las desigualdades de género están relacionadas con otras desigualdades asociadas con la posición cultural, económica o política de los grupos particulares de mujeres.



Data Género

Observatorio de datos con perspectiva de género

<https://www.datagenero.org/>

“Desde DataGénero - Observatorio de Datos con Perspectiva de Género Asociación Civil nos parece importante remarcar que en lo que respecta a la ciencia y la tecnología todavía queda mucho camino por recorrer hacia la igualdad de género.

Desde la industria tecnológica se dice que hay un 30% de mujeres trabajando actualmente, pero no se especifica en qué roles están, sin siquiera mencionar el porcentaje de personas LGBTI+, que es casi inexistente. Los roles de programación

“más dura” o de ingeniería suelen estar ocupados casi exclusivamente por varones cis, blancos y heterosexuales, lo que genera que muchas veces las soluciones tecnológicas carezcan de perspectiva de género y de empatía por la diversidad. Estos varones cis, que toman decisiones, lideran empresas y crean la tecnología que usamos día a día, además pertenecen en general a una clase media universitaria, ganando mayores sueldos que la media en nuestro país y por lo tanto en muchas ocasiones no están sensibilizados con las problemáticas de otros géneros, otras clases socioeconómicas, es decir: de otras realidades.

Creemos firmemente que tenemos que comenzar a pensar en otras formas de hacer ciencia y tecnología, las normas y prácticas que heredamos fueron pensadas para un cierto tipo de persona (varón, blanco, cis, hetero, clase acomodada).

Si queremos desafiar esas herencias, tenemos que patear el tablero y comenzar a pensar en formas más inclusivas y sensibles que dejen de importar modelos del Norte Global y que comiencen a pensar en la sociedad como un espacio complejo, con diversidades que deben ser parte de los procesos de creación de ciencia y tecnología en nuestro país”.

Las queer feministas comparten con los movimientos anteriores la idea de que la tecnología y el género mantienen una relación de conformación mutua. Sin embargo, consideran que la batalla feminista es una batalla contra toda forma de caracterizar una identidad en términos esencialistas y sostienen que, si bien la tecnología está generizada, el análisis debe focalizarse en cómo el género es un proceso fluido de construcción y desconstrucción mediado por la tecnología (Landstrom 2007). Así como muchos de los movimientos anteriores promueven abrir la caja negra de la tecnología en consonancia con el constructivismo social en tecnología, las queer feministas proponen ahora abrir la caja negra del género.

3.

Desarrollos en IA: géneros, sesgos y discriminación

Como hemos señalado en la primera parte de este trabajo, el impacto de las tecnologías digitales en nuestra vida cotidiana es tan profundo y se dio en tan poco tiempo que dificulta una verdadera reflexión al respecto porque, por un lado, carecemos de suficiente perspectiva para hacerlo y, por otro, la velocidad con la que avanza motivado por sucesos como la crisis por la COVID-19 puede resultar abrumadora. Quizá sea por estos dos motivos por los que dos términos clave para entender la tecnología digital bajo una perspectiva de género son continuamente utilizados sin tener una definición clara o consensuada. Nos parece relevante, entonces, detenernos a pensar en ellos para poder lograr una taxonomía que nos resulte útil para nuestros análisis y reflexiones.

3.1. De qué hablamos cuando hablamos de algoritmos

Un término clave que aparece cuando hablamos de tecnologías digitales es el de “algoritmo”. Hay dos modos de caracterizar conceptualmente a los algoritmos, a saber, en sentido estrecho o en sentido amplio.

En **sentido estrecho**, un algoritmo es un constructo matemático que se selecciona durante el diseño de un sistema o artefacto tecnológico dada su eficacia pasada en la resolución de tareas similares a las que plantea el problema que ahora se pretende resolver. Son ejemplos de algoritmos en sentido estrecho las redes neuronales profundas, las redes bayesianas, las cadenas de Markov, el modelo Perceptrón simple, etc. (Mittelstadt et al. 2016; Pasquinelli y Joler 2021).

En **sentido amplio**, un algoritmo es un sistema tecnológico tripartito, que comprende datos de entrenamiento, un algoritmo de aprendizaje (el algoritmo en sentido estrecho) y un modelo estadístico como *output* final, y que está diseñado, ensamblado e implementado para determinados propósitos relacionados con la resolución de un problema práctico formulado previamente. Hay cuatro fases clave por las que pasa la producción de un algoritmo en sentido amplio, a saber, (i) la caracterización tanto del problema que se desea resolver como de la solución buscada; (ii) el diseño, formateo y edición de los datos con los que se va a trabajar; (iii) la selección del algoritmo en sentido estrecho; y (iv) la fase de entrenamiento del algoritmo en base a los datos disponibles, y la evaluación del sistema tecnológico hasta su puesta a punto. En cada una de estas fases aparecen problemas y dificultades a resolver, se delibera y se toman decisiones que brindan soluciones, y en cada una de estas fases emergen dificultades algunas de las cuales aparecen identificadas como “sesgos” en la literatura especializada.

3.2. De qué hablamos cuando hablamos de sesgos

Así como, a pesar de su uso extendido, el término “algoritmo” no siempre se usa de la misma forma, i.e. no siempre se usa en el sentido estrecho, técnico; la noción de sesgo también se

usa en contextos muy diversos, con significados no siempre unívocos. La RAE reconoce varios sentidos,¹⁹ entre los que encontramos: (1) torcido, situado oblicuamente; (2) curso o rumbo que toma un negocio, y (3) el sentido estadístico, que seguramente es el que más se ajusta mejor a nuestro tema, y que la RAE define así: “error sistemático en el que se puede incurrir cuando al hacer muestreos o ensayos se seleccionan o favorecen unas respuestas frente a otras”. Como vemos, solo el tercero tiene un sentido negativo (por tratarse de un error).

La noción de sesgo se ha extendido en muchos ámbitos, tal vez por la ubicua incorporación de modelos estadísticos en las más variadas disciplinas. Sin embargo, no es claro que en todos sus usos se recoja exactamente lo mismo, ni que pueda entenderse simplemente como una falla en los datos, que pueda corregirse (como cualquier error, que una vez detectado es susceptible de ser corregido).²⁰

Así, este término se utiliza de forma regular en ámbitos tan dispares como la ingeniería, las ciencias cognitivas y la psicología. Algunos se proyectan en la IA y sus usos, como por ejemplo ciertos sesgos cognitivos que refieren a patrones de procesamiento cognitivo como el sesgo de confirmación, que nos lleva a tomar como cierta de manera irreflexiva la información que confirma nuestras creencias y a rechazar aquella que entra en contradicción con nuestras creencias previas, alterando los modelos racionales ideales de revisión y fijación de creencias. También se reconoce la existencia de sesgos implícitos (*implicit bias*) en todos los sistemas cognitivos humanos,²¹ dado que tendemos a actuar muchas veces por asociaciones que tenemos incorporadas en nuestra cognición de manera implícita, de las que no somos conscientes, y que contradicen las creencias que estamos dispuestos a asumir explícitamente como propias. Estos sesgos implícitos están a la base de las conductas discriminatorias tanto por cuestiones de género como raciales y sociales y abren una multiplicidad de preguntas tanto psicológicas como filosóficas para estudiar (Brownstein 2019).

En el dominio de las tecnologías digitales el uso de la palabra sesgo es más difuso. Por un lado, suele ser la traducción en español de la palabra en inglés *bias*, que en ocasiones también es traducida como *prejuicio*, vocablo que tiene en el hablar cotidiano una connotación negativa. Por otro lado, como los algoritmos en sentido estrecho están íntimamente ligados a modelos estadísticos, algunas personas piensan los sesgos en los algoritmos tal como en ocasiones se piensan los sesgos estadísticos, esto es, como errores que se pueden detectar *ex post facto* en un estudio y que se producen en alguna de sus etapas, ya sea en la recolección, el análisis, la interpretación o la revisión de los datos. Entonces, ¿los sesgos son algo necesariamente negativo? ¿Solo pueden estar vinculados con errores? Quizá sea útil pensarlos a la luz de la definición amplia de algoritmo que propusimos más arriba.

En el apartado que sigue mostraremos cómo en diferentes puntos del desarrollo de sistemas de IA, esto es en la producción de “algoritmos amplios”, se realizan decisiones de diseño que pueden redundar en conductas del artefacto terminado que consideramos sesgadas. Así, entendemos que la aplicación de la noción de sesgo para el caso de los desarrollos en IA, si bien hunde sus raíces en la noción de sesgo estadístico, y está primariamente asociada a los datos

19 <https://dle.rae.es/sesgo>

20 Recordemos, sin embargo, que en estadística un error no es algo que necesariamente pueda ser corregido, sino algo con lo que hay que convivir; qué margen de error tolerar en un modelo estadístico también es una decisión de diseño.

21 Si alguien tiene dudas de poseer estos sesgos implícitos lo invitamos a testearse a sí mismo en <https://implicit.harvard.edu/implicit/takeatest.html>

de partida, también aparece en otros momentos y decisiones de diseño, y también se aplica esta noción al resultado final del proceso de diseño y uso del sistema de IA. Veamos más detenidamente esta cuestión.

3.3. Algunos sesgos en los sistemas de algoritmo en sentido amplio

Existen diferentes taxonomías para catalogar a los múltiples tipos de sesgos existentes en los sistemas tecnológicos digitales. En esta sección, basándonos en nuestras consideraciones anteriores sobre el carácter difuso de esta noción, esbozaremos una posible clasificación, que no pretende ser la única ni ser exhaustiva.²²

Tal como afirman Pasquinelli y Joler (2021), en línea con lo que desarrollamos en la primera parte de este trabajo, los límites de la IA no son simplemente técnicos sino que están “imbricados con sesgos humanos” pero, a su entender, los algoritmos amplifican los efectos de estos sesgos característicamente humanos. En este sentido, argumentan que en el aprendizaje automatizado es necesario distinguir entre:

- sesgos históricos,
- sesgos de datos y
- sesgos algorítmicos.

Como mencionamos arriba, al dar cuenta de las cuatro fases por las que pasa la producción de un algoritmo en sentido amplio, el primer paso para el diseño lo constituye el planteo de un problema a resolver según un propósito práctico para el cual se creará el sistema tecnológico.

(i) Obviamente, el planteo mismo del problema puede estar sesgado, en el sentido de estar inclinado a resolver cierto problema, y no otros, desde cierta perspectiva y con una cierta finalidad. Por ejemplo, se puede decidir diseñar un sistema que prediga embarazo adolescente en lugar de un sistema que busque identificar conductas en varones que permitan predecir ataques sexuales a mujeres menores de edad. Dos caras de un mismo problema.

El sesgo histórico (o sesgo del mundo) ya es evidente en las sociedades antes de la intervención tecnológica. Es el sesgo que está presente en el planteo mismo del problema y en la determinación de los objetivos de diseño. El problema en este caso es la falta de reconocimiento de estos sesgos iniciales, plasmada en la idea de neutralidad valorativa de la tecnología, idea que, como argumentamos más arriba, es insostenible. Los sistemas de IA, así como otros sistemas tecnológicos, cuando se consideran erróneamente como valorativamente neutrales, llevan a la naturalización de este sesgo, o sea, a la integración de la desigualdad en una tecnología aparentemente neutral lo que la hace especialmente dañina (Eubanks, 2018) (véase también Crawford, 2017). Parafraseando a Michelle Alexander, se lo ha llamado “el nuevo código Jim”:

22 Para una lista más detallada de sesgos de la IA véase: Suresh & Guttag, 2020 y Mehrabi et al., 2019, citados en Pasquinelli y Joler 2021.

“el empleo de nuevas tecnologías que reflejan y reproducen desigualdades existentes pero que son promovidas y percibidas como más objetivas o progresistas que los sistemas discriminatorios de épocas previas” (Benjamin, 2019, p. 5).

En este punto es importante notar que todos los sistemas de aprendizaje automatizado, en la medida en que se trata de modelos estadísticos generados a partir de los datos de entrenamiento, contienen lo que se ha dado en llamar este “sesgo histórico” es decir, reproducirán las categorías y comportamientos que se han dado en el pasado (y que recogen los datos de entrenamiento). Es por esto que cuando se trata de tomar decisiones que busquen *cambiar el futuro comportamiento* respecto del pasado, estos sistemas probablemente no sean de gran utilidad.

Los otros tipos de sesgos -de datos y algorítmicos- aparecen en los momentos posteriores del diseño de los sistemas cuando se pasa a las siguientes fases: (ii) los datos, (iii) la selección del algoritmo propiamente dicho, en sentido estrecho, que se entrenará a partir de los datos seleccionados hasta que -como resultado del entrenamiento- pueda realizar la tarea que se le encomendó, y (iv) el modelo estadístico resultante del proceso de entrenamiento que será el usado para que el algoritmo en sentido amplio resuelva la tarea encomendada. En estas tres etapas aparecerán los sesgos de datos y los sesgos algorítmicos, en la medida en que estas tres instancias a su vez suponen una cantidad de toma de decisiones para el diseño. Veámoslo más detenidamente.

(ii) La recolección y selección de los datos de entrenamiento son cruciales para el desarrollo del sistema. Las decisiones sobre el diseño, formateo y edición de los datos incluyen dos elementos: (a) la producción o recolección y (b) el etiquetado o interpretación de los datos.



Juliana Guerra

Derechos Digitales

<https://www.derechosdigitales.org/>

<https://www.derechosdigitales.org/equipo/juliana-guerra/>

“Es muy relevante contar con datos que reflejen mejor las realidades diferenciadas de las poblaciones, por ejemplo, con relación al acceso a tecnologías digitales, reconociendo además las múltiples dimensiones que tiene el acceso. Puede ser simplemente contar con datos desagregados por género, localidad, identidad étnica y cultural; y tipo de dispositivo, tipo y velocidad de conexión, habilidades digitales y conocimiento sobre el funcionamiento de estas tecnologías, entre otras cuestiones. Pero es aún más relevante tener una perspectiva crítica sobre el uso de esos datos: para qué y cómo son recolectados. Esto es, prestar mucha atención a la finalidad de los datos y su caducidad.

Pero, sobre todo, el consentimiento a la hora de su recolección y procesamiento. Esto no se limita a que las personas sujetas de derechos estén informadas y consientan en primera instancia el manejo de sus datos. Implica reconocer los derechos económicos, sociales y culturales por los que trabajan grupos que han sido tradicionalmente marginados y subrepresentados, y adecuar la recolección y procesamiento de datos a dichos objetivos”

El primer paso será la producción o recolección de datos, y para ello es necesario seleccionar una técnica de recolección. En algunas ocasiones los datos son provistos por bases de datos ya establecidas. En otras ocasiones, los datos son provistos por quien solicita el diseño, por ejemplo, si se busca diseñar algún artefacto para colaborar con la toma de decisiones en políticas públicas, los datos los puede proveer el organismo estatal que encarga el diseño, o la empresa que desarrolla algún sistema para mejorar su desempeño con los datos previamente almacenados en los perfiles de los que dispone y que han cedido sus datos al aceptar los “términos y condiciones” de uso.



Marina Benítez Demtschenko

Fundación activismo feminista digital

<https://activismofeministadigital.org/>

“Creo que la privacidad digital es el gran derecho puesto en riesgo en los últimos tres años. Esto responde no sólo a las grandes empresas y sus políticas internas de uso de información de usuarios y usuarias en grados preocupantes, sino también a la falta de conciencia colectiva de este derecho como un derecho fundamental. Lo último, por hábitos que se han consolidado cotidianos en pos de una comunidad digital sin alfabetización digital suficiente, por ejemplo, la sobreexposición propia y ajena. Esto lleva a situaciones posteriores de difícil resolución –preventiva/reparadora-, dado que luego de vulnerada la privacidad, los efectos consiguientes van en escala.

Hay muy poco conocimiento sobre la importancia de los derechos digitales y normalmente se toma conciencia del impacto en la vida analógica, una vez que el daño ya está consumado. El descreimiento del efecto en cascada de la lesión a este tipo de derechos, viene de la mano con la subestimación de los entornos digitales como plano de ejercicio de derechos.

Hay que cambiar el paradigma y entender [los derechos digitales] como derechos fundamentales, a la par de exigir su respeto por terceros. Las ciudadanas comunes requieren alfabetización digital para reconocerse, primero, sujetas de estos derechos.

Luego, encontraremos la forma de hacerlos valer con estrategias colectivas, de apoyo mutuo, de respeto entre pares, y de exigencia a quienes los pongan en riesgo

El problema sigue estando en que no estamos lo suficientemente informadas sobre esto aún. La información se logra con alfabetización digital y ese es el mayor desafío en Latinoamérica, de acá a unos cuantos años por delante”.

En otras ocasiones se recogen datos *ad hoc* por medio del trabajo realizado por personas en el marco de las “economías de los pequeños encargos” en la que se accede a una gran fuerza laboral en línea para, por ejemplo, buscar contenidos en internet para un fin específico o crear contenido, como el que se realiza de manera individual al traducir fragmentos cortos de texto o etiquetando imágenes para algunas tareas.

Dado que los datos pueden recolectarse de diversas maneras, hay conjuntos de datos que ya están disponibles codificados como para ser procesados por el algoritmo y otros que no lo están. En este último caso, habrá además que tomar decisiones concernientes al método de codificación y organización de los datos, etiquetado y clasificación para volverlos accesibles al algoritmo de aprendizaje supervisado.

El sesgo de datos se introduce en estas instancias de recolección y preparación de los datos de entrenamiento por parte de operadores humanos, pero también en la parte más delicada del proceso es el etiquetado de los datos, en el que las taxonomías viejas y conservadoras pueden causar una mirada distorsionada del mundo, tergiversando las diversidades sociales y exacerbando las jerarquías sociales.

(iii) Por su parte, la fase de selección del algoritmo de aprendizaje en sentido estrecho requiere de una *decisión técnica basada en un juicio experto* y en experiencias anteriores en las que se han diseñado algoritmos para resolver problemas similares al que se busca resolver en esta ocasión. Se trata del momento de selección del algoritmo de aprendizaje automatizado -en sentido estrecho- que sea el más apropiado para analizar los datos disponibles y generar a partir de allí una red entrenada para dar respuesta a las preguntas planteadas. Esta instancia involucra dar respuesta a la pregunta en torno a qué arquitectura de algoritmo se utilizará del menú disponible: ¿redes neuronales profundas, redes bayesianas, cadenas de Markov, Perceptron simple, etc.?



Eleonora Cresto

Instituto de Investigaciones Filosóficas-SADAF-CONICET
Editora Asociada de Erkenntnis

<https://iif.conicet.gov.ar/investigadores/investigadores-conicet/eleonora-cresto/>

“Para controlar y corregir los sesgos de los algoritmos de IA se necesita un monitoreo continuo de los resultados en busca de sesgos, para detectarlos tempranamente y

tratar de entender dónde está el problema – proceso que requiere de cierta inversión de recursos. En algunos casos la solución será más fácil que en otros, desde luego, y puede consistir desde introducir cambios en las muestras de entrenamiento de los algoritmos, hasta la introducción de elementos de discriminación positiva por género a la hora de determinar umbrales de predicción (dependiendo de qué algoritmo se trate). En cualquier caso, aún con sus dificultades, el monitoreo y corrección de algoritmos nos permite intervenciones más difíciles de lograr con los seres humanos de carne y hueso, con sus sesgos implícitos y motivaciones no siempre transparentes”.

El sesgo algorítmico es la amplificación adicional del sesgo histórico y del sesgo de datos producida por los algoritmos de aprendizaje automatizado. Según el diagnóstico de Pasquanelli y Joler (2021), el problema del sesgo se origina principalmente por el hecho de que los algoritmos de aprendizaje automatizado están entre los más eficientes para la compresión de información, lo cual da lugar a cuestiones de resolución, difracción y pérdida de información. En efecto, los algoritmos son procedimientos de naturaleza económica, diseñados para obtener un resultado en el menor número de pasos, consumiendo la menor cantidad de recursos: espacio, tiempo, energía y trabajo. La carrera de las compañías de IA está, aún hoy, ocupada en encontrar los algoritmos más rápidos y simples con los cuales capitalizar los datos. Si bien la compresión de información produce una tasa máxima de ganancia en la IA corporativa, desde un punto de vista social produce discriminación y exacerbación de desigualdades estructurales, en un sentido que se evidencia claramente a partir de los ejemplos de sesgo de género en que ejemplificaremos en detalle en la tercera parte de este trabajo.

Finalmente, para realizar el entrenamiento del algoritmo seleccionado a partir de los datos disponibles, el primer paso a realizar es separar los datos de entrenamiento de los datos de evaluación. Este paso es considerado por los especialistas como la “regla de oro”, dado que la verificación de que el algoritmo ya entrenado resuelve como es deseable la tarea a realizar se producirá ingresando los datos seleccionados para la evaluación. Es por ello que resulta crucial que los datos de evaluación sean independientes de los datos de entrenamiento. Cuando esto no ocurre, el modelo resultante puede resultar sesgado, por la incorrecta división entre ambos tipos de datos.

(iv) Por último, entonces, el resultado de este proceso de entrenamiento es un algoritmo que ofrece un modelo estadístico, que permite realizar diversas acciones al sistema, por ejemplo, tomar decisiones acerca de la aplicación o no de categorías a objetos (por ej. si se trata de un mecanismo de identificación de imágenes, o de clasificación de objetos). En cualquier caso, dado que el resultado es un modelo estadístico, los resultados siempre contendrán un *margin de error*, es decir, incluirán indefectiblemente falsos positivos y falsos negativos. Una decisión importante que quienes diseñan toman es *qué margen de error se está dispuesto a tolerar*. Y esta decisión estará basada en una cantidad variable pero grande de *valores* que tienen en cuenta quienes diseñan: la eficiencia del sistema, la relación costo-beneficio en la aplicación de un sistema que produce falsos positivos y falsos negativos, decisiones económicas de la empresa, decisiones políticas del organismo que usará el sistema, etc.



Damián Szmuc

Instituto de Investigaciones Filosóficas-SADAF-CONICET

<https://iif.conicet.gov.ar/becarios/damian-szmuc/>

“Me parece que la presencia de colectivos y representaciones de los sectores sociales vulnerables es esencial en cualquier acción que regule la existencia de estos sesgos. Esto podría canalizarse, por ejemplo, mediante la conformación de comisiones de seguimiento y auditoría independientes (adicionales a las ya presentes en organismos estatales) integradas por una variedad de personas de los sectores sociales perjudicados, activistas y no activistas, cuyo bienestar podría ser especialmente vulnerado por una gestión deficiente de estas problemáticas. Podrían también establecerse normas para que el visto bueno de estas comisiones independientes fuera un requisito esencial para la ejecución de políticas públicas, en vistas de que no puedan ser deliberadamente ignoradas y terminen teniendo un carácter meramente anecdótico”.

En resumen: hemos visto que en las cuatro fases del diseño de un algoritmo en sentido amplio –(i) planteo del problema, (ii) selección de datos, (iii) selección del algoritmo en sentido estrecho y (iv) la puesta a punto del modelo- se toman decisiones que suponen la producción de un sistema tecnológico cargado de valores, en muchos casos los valores vigentes en la sociedad y/o empresa que los produce; en particular, en lo que atañe al tema que nos ocupa en este trabajo, sesgos de género que reproducen y amplifican las características de las sociedades patriarcales en las que vivimos.



Guillermo Anlló y Eleonora Lamm

Oficina Regional de Ciencias para América Latina y el Caribe-UNESCO (Montevideo)

<https://es.unesco.org/fieldoffice/montevideo>

“Mientras continúan los desafíos planteados por la brecha digital y la escasa representación de las mujeres y diversidades sexuales y de género en STEM (Ciencia, Tecnología, Ingeniería y Matemática), la IA y la automatización generan nuevos desafíos para lograr una igualdad de género sustancial.

Los datos muestran que es un hecho demostrado que si la IA y la automatización no se desarrollan y aplican con perspectiva de género, es probable que reproduzcan y refuercen los estereotipos y normas sociales discriminatorias existentes.

La conferencia General de UNESCO aprobó el año pasado la Recomendación sobre la ética de la IA, donde la igualdad de género ocupa un lugar destacado, con un área de acción política específica donde brinda importantes recomendaciones para que la IA no profundice las brechas. Con las líneas de acción propuestas UNESCO procura asegurar que el género sea una parte integral de las herramientas de evaluación del impacto ético que se utilizarán para rectificar las desigualdades en todo el espectro del ciclo de vida de la IA, incluida la nivelación del campo de juego para abrir el acceso a mujeres en campos de vanguardia y orientados al futuro como la IA”.

4.

Los sesgos en acción

Los sesgos de género en los sistemas tecnológicos que hacen uso de la IA son de tipo muy variado. En lo que sigue vamos a condensar varios casos que sucedieron en la última década bajo dos categorías.

- 1 Los sistemas de recomendaciones, que suelen estar orientados al marketing y la publicidad, las políticas públicas y la recomendación de personal en una búsqueda de trabajo, entre otros.
- 2 Los sistemas que involucran el procesamiento del lenguaje natural, que permiten crear dispositivos que “entienden” a sus usuarios, procesan su habla o su escritura y buscan devolver una respuesta coherente y eficiente.

Si bien no se trata de categorías excluyentes, ya que veremos que hay ejemplos de artefactos que realizan simultáneamente estos dos tipos de tareas, presentan dificultades e involucran sesgos de diferente tipo.

4.1. Sistemas de recomendaciones

Dentro de la primera categoría mencionaremos tres ejemplos: el sistema de contratación que presentó en 2014 la compañía estadounidense Amazon, algunos modelos existentes de tecnologías de reconocimiento de rostro por medio de algoritmos de aprendizaje automatizado y la manera en la que se decide a quiénes se mostrarán los anuncios publicitarios de las compañías Google y Meta.

4.1.1. El sistema de contrataciones de Amazon

Uno de los casos más resonantes y paradigmáticos de cómo influyen los sesgos de género de los algoritmos en la vida real es el caso de reclutamiento de personal de Amazon. Durante el año 2014 la compañía decidió automatizar la tarea de contratar nuevo personal, un área en la que encontraba más y más dificultades por el volumen de la tarea, y para ello diseñó un sistema de recomendación mediante algoritmos de aprendizaje automatizado²³. Alentados por el buen resultado de los algoritmos de recomendación de productos de su tienda web, que crea portadas personalizadas para cada usuario, dentro de la empresa creyeron que podrían imitar la mecánica para compilar la información de todas las personas que se presentaban para un

23 <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraped-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>

puesto determinado y brindarle a quien reclute un ranking con los mejores currículums para que tome la decisión que considere más conveniente.

De este modo, quienes crearon el dispositivo en cuestión confiaron en que podían ofrecer a su oficina de personal una herramienta útil para que pudiera desatenderse del esfuerzo que implicaba mirar los cientos de currículums que llegaban a diario y brindarle una forma “rápida, eficaz y objetiva” de identificar a las personas más capacitadas para cada puesto. Al igual que lo que sucede con el sistema de reseñas de productos en su plataforma, cada persona que se presentaba para un trabajo era calificada por este sistema de algoritmos con un puntaje que iba de una a cinco estrellas.

Para crear el sistema, se utilizó el vasto archivo de currículums recibidos de los últimos diez años, indicando cuáles eran los casos de éxito en la contratación. Así se identificaron cerca de 50 mil términos que aparecían y se entrenó a un modelo estadístico para que pudiera identificar patrones que permitieran seleccionar a los mejores candidatos y candidatas.

Pero cuando se implementó el sistema, que tenía como objetivo facilitar la parte más engorrosa de un proceso de selección de personal, se descubrió que de manera sistemática los algoritmos no recomendaban candidatas de género femenino y que las listas con los cinco mejores currículums estaban siempre integradas por varones. Frente a esto, quienes crearon el software decidieron eliminar de las categorías relevantes de análisis el género de quien se presentaba al trabajo, las universidades que explícitamente solo admitían mujeres y hasta la misma palabra “mujer”. De este modo, el sistema no tendría la información explícita de si quien se presentaba lo hacía como hombre o como mujer. Sin embargo, los algoritmos siguieron mostrando preferencia por los currículums masculinos.

Así, el novedoso sistema, que había sido anunciado con bombos y platillos dentro de la compañía, debió ser desactivado. Una investigación posterior determinó que la misma base de datos con la que se entrenó al algoritmo estaba constituida por las decisiones que habían tomado las personas encargadas de la contratación de personal en el pasado y que, por lo tanto, reflejaba una conducta en la que se había discriminado a las mujeres: en los últimos 10 años Amazon había contratado ingenieros de género masculino en la gran mayoría de los casos y sus características iban mucho más allá de los marcadores explícitos. No había, por lo tanto, errores propios del sistema, sino que éste reflejó un sesgo preexistente en las prácticas de contratación de la empresa, que estuvo incorporado en los datos seleccionados para el entrenamiento, dada la tasa previa de preferencia por varones para los puestos, y que se vio amplificado por el uso del algoritmo que los datos entrenaron.

4.1.2. Reconocimiento facial y de voces

Una de las aplicaciones de la IA que más está creciendo en los últimos años son los sistemas de reconocimiento de personas. Si bien el uso de la tecnología para identificar individuos es de larga data, como el viejo método de las huellas digitales o el iris, desde comienzos de siglo avanzó a pasos agigantados el uso de algoritmos que procesan imágenes de cámaras para analizar rostros. Es la manera en la que, por ejemplo, millones de personas a diario desbloquean las pantallas de sus teléfonos celulares o la forma en la que muchas líneas aéreas permiten el ingreso a sus vuelos. Es casi imposible transitar hoy por una ciudad desarrollada sin

ser capturado por cámaras, ya sea en la vía pública o en algún espacio privado como un local comercial o un banco.

El reconocimiento facial es cada vez más utilizado por distintas empresas con diferentes motivos, que van desde la seguridad hasta la contratación de personal. En los últimos años esta tecnología ha avanzado a gran velocidad, lo mismo que la implementación de su uso. Y como toda nueva tecnología, a medida que se ha ido masificando han crecido los peligros que lleva asociada y se ha vuelto cada vez más “natural”, en tanto las personas dejan de sorprenderse por la multiplicación de cámaras en espacios públicos, privados o de trabajo y lo toman como un elemento más del paisaje urbano.



Tobías Schleider

Universidad Nacional de Mar del Plata, Universidad Nacional del Sur, Instituto Latinoamericano de Seguridad y Democracia, Consultor Internacional en Seguridad Democrática.

<https://www.ilsed.org/>

“Entre los derechos vinculados con lo digital más vulnerados en la región está el derecho a la imagen, que afirma que la captación, difusión, reproducción o publicación de una persona solo puede hacerse con su consentimiento (o por manda legal que lo supla). Y uno de los mecanismos más frecuentes y con mayor crecimiento para vulnerarlo con frecuencia e insistencia es el uso de la videovigilancia, en especial con la incorporación de la tecnología de reconocimiento facial.

En nuestra región, las autoridades no suelen brindar detalles sobre cuáles son los términos de los convenios celebrados con las empresas proveedoras de los sistemas informáticos que dan base a estas iniciativas, ni de los resultados de su puesta en práctica, ni de los eventuales ajustes que se realizan (o no) en virtud de ellos. Por otra parte, es habitual que integrantes de fuerzas de seguridad y funcionarios del Poder Ejecutivo, del Poder Judicial y del Ministerio Público Fiscal sean instruidos en el empleo de herramientas tecnológicas cuyo control se encuentra en poder de un grupo minúsculo de empresas privadas, a espaldas del conocimiento de la ciudadanía. En pocas ocasiones el destino de las imágenes captadas puede conocerse con precisión.

Informaciones básicas como las relativas a su almacenamiento y la seguridad informática involucrada, su difusión voluntaria o por negligencia, su uso o incluso su comercialización, son opacas o, directamente, están ocultas. El ámbito de despliegue del derecho a la imagen es, por ende, cada vez más reducido. Y se conecta con otros derechos y garantías vulnerados, como el estado de inocencia, el debido proceso legal, la libertad de expresión, la intimidad y la protección de datos personales de índole diversa”.

Más allá de las implicancias que puede tener este monitoreo constante de los individuos, y de los supuestos que están en la base de la identificación de las personas, los sistemas de reconocimiento facial están muy lejos de ser perfectos y cada vez más muestran sus limitaciones e inconvenientes.

Uno de sus problemas principales es que estos sistemas muestran una tasa de error superior con los rostros de las mujeres que con los hombres. Originariamente se pensó que la razón de este sesgo era que los algoritmos de aprendizaje automatizado habían sido entrenados con bases de datos que no son representativas de la población a la cual se le aplican, pero aun cuando esas bases se encuentran equilibradas para ambos géneros, la diferencia se mantiene.²⁴

Así otra posible explicación del origen del sesgo contra la población femenina en esta tecnología parece ser que el algoritmo es muy sensible a pequeñas modificaciones tanto en el rostro, como en el pelo o debido al uso de maquillaje, cambios más frecuentes en rostros de género femenino que masculino. En una investigación reciente²⁵ se ha encontrado que los rostros femeninos son mal identificados hasta en un 18% de veces más que los masculinos.

En una realidad social que descansa cada vez más en sistemas de seguridad basados en el reconocimiento de rostros -tanto para funciones de teléfonos inteligentes como también para poder ingresar a ciertos edificios, acceder a espectáculos o eventos deportivos, tomar un avión o hacer un trámite administrativo- que nuestra cara no sea correctamente identificada acarrea varios problemas. No solo eso: en los últimos años también creció el uso de estos sistemas por parte de fuerzas del orden, lo que redundó en detenciones de personas en medios de transporte o espacios públicos.²⁶

Pero no solo el reconocimiento facial está sesgado en contra del género femenino, sino también lo está otro mecanismo cada vez más popular dentro de los dispositivos: el reconocimiento de la voz. Si bien Google asegura que tiene una tasa de precisión mayor a 95% a la hora de procesar la voz humana, un análisis más minucioso revela que es un 13% más preciso para voces de varones que para voces de mujeres.²⁷

Los inconvenientes de este sesgo quedaron a la luz con el caso de una mujer irlandesa que se presentó a un trabajo en el que reprobó una prueba de dominio del inglés oral evaluado por un sistema de IA, que le otorgó puntaje de 74 sobre 90, siendo ella hablante nativa de dicho idioma.²⁸

Al analizar por qué las voces de las mujeres no son procesadas con el mismo nivel de acierto que las de los hombres, una vez más se comprueba que los sistemas fueron entrenados con datos sesgados. En el caso de Google, por ejemplo, se sabe que se usan las populares charlas de divulgación TED que están compuestas por un 70% de oradores de género masculino.

4.1.3. Segmentación y entrega de anuncios

Cuando, en 2019, la compañía Facebook, que fue rebautizada en 2021 como Meta, celebró

24 <https://arxiv.org/abs/2002.00065>

25 <https://embed.widencdn.net/pdf/plus/watchguard/x3zt3dnmbh/Facial-Recognition-Article.pdf?u=sune5b>

26 <https://www.npr.org/2020/06/24/882683463/the-computer-got-it-wrong-how-facial-recognition-led-to-a-false-arrest-in-michig>

27 <http://www.ethicsinnlp.org/workshop/pdf/EthNLP06.pdf>

28 <https://www.theguardian.com/australia-news/2017/aug/08/computer-says-no-irish-vet-fails-oral-english-test-needed-to-stay-in-australia>

quince años de vida, su CEO y fundador Mark Zuckerberg publicó una carta abierta en la que detalló el modelo de negocios de la compañía a raíz de “muchas preguntas” que estuvo recibiendo. En el texto, el empresario explicó que “las personas siempre nos dicen que, si van a ver anuncios, prefieren que sean relevantes. Esto significa que tenemos que entender lo que les interesa. De modo que, partiendo de las páginas que les gustan y los contenidos en los que hacen clic, además de otras señales, creamos categorías -por ejemplo, gente que vive en España y a la que le gusta las páginas sobre jardinería-, y los anunciantes pagan por mostrar su publicidad a esa categoría de personas”.²⁹ De acuerdo con lo explicado “si bien la publicidad dirigida a grupos específicos existía mucho antes de Internet, la publicidad on line permite una segmentación mucho más precisa y, por lo tanto, anuncios más relevantes”.

Así, el modelo de negocios de Meta, y de compañías como Google, depende en gran medida de la publicidad, de presentar anuncios publicitarios lo más personalizados posibles, en una acción de segmentación que realizan sistemas de algoritmos basándose en las conductas registradas de cada perfil. A esto se lo llama “entrega de anuncios” y aquí también se comprueban sesgos de género.

En 2015 investigadores de la Universidad Carnegie Mellon crearon perfiles ficticios de consumidores y los pusieron a visitar sitios web de empleo y comportarse como lo harían si estuviesen buscando un trabajo. La mitad de los 1000 perfiles de usuario creados eran de género masculino y la otra mitad de género femenino. Luego de registrar la actividad de búsqueda de trabajo, la empresa Google empezó a mostrar anuncios donde se solicitaban distintos empleos. Lo que pudieron observar es que el tipo de anuncios que se mostraban a los perfiles masculinos difería ampliamente del tipo mostrado a los perfiles femeninos. Por ejemplo, uno de los resultados de la investigación fue que un anuncio de ejecutivos bien pagado se les mostró 1.816 veces a hombres y solo 311 veces a mujeres.

Lo mismo se corroboró con Meta³⁰, incluyendo no solo sesgos de género sino también de etnia. De acuerdo con el estudio, se mostraron cinco anuncios en los que se buscaban conserjes a una audiencia que era 65% femenina y 75% negra. También se mostraron cinco anuncios de empleos en la industria maderera a usuarios que eran 90% hombres y 70% blancos. Y en los casos más extremos, se mostraron anuncios de empleados de supermercados a audiencias que eran en un 85% mujeres y oportunidades de conducir un taxi a audiencias que eran 75% afrodescendientes. Si bien, al igual que en caso de Amazon, no estaba explícito el dato de género o raza, de todas formas, los algoritmos lograban identificar a sus objetivos indirectamente mediante otros datos proporcionados, como por ejemplo el código postal (ya que la composición étnica varía según el barrio donde se viva) o sus estudios o intereses deportivos y musicales.

Los sesgos iban mucho más allá de lo evidente: por ejemplo, en los casos de anuncios con imágenes, elementos de la fotografía como una pelota de fútbol o un perfume eran claves para entender a quién le mostrarían la publicidad. En el primer caso, el objeto estereotípicamente masculino era dirigido a audiencias identificadas como hombres mientras que el segundo, un objeto estereotípicamente femenino, a mujeres. Se trataba de anuncios que, a excepción de ese elemento, eran iguales pero que cuando tenían imágenes “masculinas” llegaban a una audiencia que era 60% masculina, mientras que el anuncio “femenino” fue mostrado a una audiencia 65% femenina. Los resultados indican que el sistema implementado por Facebook

²⁹ <https://www.elmundo.es/economia/empresas/2019/01/25/5c4a2034fdddfa4918b45a9.html>

³⁰ <https://arxiv.org/pdf/1904.02095.pdf>

analiza de forma preventiva los anuncios para determinar quién debería verlos y toma esas decisiones utilizando estereotipos de género.

Al ser algoritmos dentro del ámbito privado, las empresas prefieren no dar explicaciones sobre cómo funcionan exactamente a menos que sean obligadas por el Estado. Eso fue lo que sucedió a partir del año 2019 con una serie de demandas presentadas en la Justicia estadounidense y europea vinculadas con conocer detalles de la política de los anuncios en temas sensibles como el empleo, la vivienda o los créditos.³¹ En dichos anuncios se estableció que no se puede especificar como parámetros la edad, el género o el código postal, así como otras variables que se puedan correlacionar con los sesgos detectados.

Los anuncios mostrados de forma sesgada según el género o la etnia están lejos de ser una excepción a la regla. Si pensamos en todo el tiempo que pasamos frente a las pantallas, casi toda nuestra actividad se encuentra dominada bajo las reglas de los distintos algoritmos que deciden qué es lo que nos muestran (y por lo tanto qué es lo que nos ocultan también). Como vemos con estos simples ejemplos, estos sistemas son los responsables de que veamos a qué trabajo podemos apuntar, pero también con quiénes nos relacionamos, qué noticias o informaciones son más importantes, todas cuestiones que de alguna manera moldean nuestra forma de ver y de ser en el mundo y es por eso que es fundamental poder entender cómo funcionan, especialmente si reproducen sesgos y valores que no compartimos.

4.2. Sistemas que involucran el procesamiento del lenguaje natural

En cuanto a la segunda categoría de sesgos, la que incluye sistemas que involucran el procesamiento del lenguaje natural, encontramos numerosos casos recientes, como lo que sucede con la función “autocompletar” que ofrecen distintas plataformas; la marcada feminización de los asistentes de voz y la generización de los robots sociales.

4.2.1. La función autocompletar

En 2013 ONU Mujeres, la organización de las Naciones Unidas dedicada a promover la igualdad de género y el empoderamiento de las mujeres, denunció que la función “autocompletar” de Google sugería palabras y frases ofensivas o con un claro estereotipo de género cuando se escribía en la barra de búsqueda cosas tales como “¿Por qué las mujeres negras...”. De acuerdo a lo recabado por la investigadora Safiya Noble las primeras sugerencias para continuar la frase e iniciar la búsqueda eran “¿Por qué las mujeres negras están tan enojadas?” “¿Por qué las mujeres negras son tan gritonas?” “¿Por qué las mujeres negras son tan malvadas?” y “¿Por qué las mujeres negras son tan atractivas?”³²

Cuando los resultados de este estudio se hicieron públicos, la compañía decidió quitar cualquier sugerencia para completar esa frase a la hora de una búsqueda a la vez que cualquier fra-

31 <https://about.fb.com/news/2019/03/protecting-against-discrimination-in-ads/>

32 Safiya Noble, Algorithms of Opression, (2018)

se que complete expresiones como “los negros son...”, “los asiáticos son...”, “los homosexuales son...”, “los latinos son...”, “los musulmanes son...” y “los judíos son...”, entre otros. La estrategia, entonces, fue eliminar o reducir funciones en el caso de grupos minorizados, en vez de buscar alternativas superadoras. Lo mismo sucedió en 2015 cuando el algoritmo de reconocimiento de objetos etiquetaba como “gorilas” a personas afrodescendientes: simplemente dejó de usar “gorila” como una opción.

Nótese que esta estrategia consistente simplemente en eliminar de la vista del usuario los casos que son potencialmente conflictivos, no corrige el algoritmo, que sigue actuando de la misma forma: solo se inhibe su exhibición al usuario. Es por eso que podemos seguir encontrando casos discriminatorios, si los buscamos. Sin necesidad de hacer una búsqueda muy exhaustiva, en junio de 2021 hemos encontrado que el sistema de autocompletado de Google completa “los doctores son...” con los siguientes adjetivos: científicos, servidores públicos, licenciados, héroes, mujeriegos, millonarios, felices... En cambio, si tipeamos “las doctoras son...” no aparece nada, o sea que las mujeres doctoras no son ni científicas, ni servidoras públicas, ni heroínas, etc. Asimismo, “las enanas son... enojonas”, en tanto “los enanos son... discapacitados”; “las enfermeras son... fáciles”, mientras que “los enfermeros son... profesionales”, etc. Como vemos el algoritmo nunca fue corregido, y el sesgo de género sigue presente. Inhibir la expresión pública del resultado del algoritmo no soluciona el problema de fondo.

En segundo lugar, en el ámbito de la filosofía y de los estudios de género, así como en los movimientos feministas, se han identificado sesgos sexistas, patriarcales y androcéntricos³³ en el lenguaje natural y sus usos, que reflejan las realidades sociales en las que esos lenguajes se usan. Desde hace más de 50 años se han propuesto una diversidad de políticas de uso “políticamente correcto” del lenguaje destinadas a eliminar del ámbito público usos discriminatorios, sea por razones de género-sexo, etnia, religión, condición social, etc. En el caso específico de las discriminaciones de identidades sexo-genéricas estos movimientos han promovido el uso de un “lenguaje inclusivo”. Ahora bien, dado que los datos de entrenamiento de los que parten los sistemas de procesamiento del lenguaje natural emplean bases de datos que utilizan una enorme cantidad de usos del lenguaje natural en donde los sesgos sexistas, androcéntricos y patriarcales predominan, todos los sistemas basados en ellos replican estos sesgos en sus tareas específicas.

Algunos de los casos detectados involucran los sistemas de autocorrección por ejemplo, los mensajes en WhatsApp que usan formas masculinas en lugar de femeninas (Wellner 2020), o traductores automáticos, en los cuales al traducir de una lengua, como el inglés en donde los sustantivos no tienen carga de género (véase Scotto y Pérez 2020) como “*teacher*” a una lengua como el español con más carga de género, i.e. donde contamos con una forma masculina “maestro” y una forma femenina “maestra”, se reproducen los sesgos de género de la sociedad usuaria de ese lenguaje, traduciendo “*nurse*” por enfermera y “*doctor*” por doctor (Wellner 2020 recoge estudios realizados que muestran estos sesgos para traducciones del inglés al hebreo y refiere a un estudio de 2017 de Quartz Magazine donde se verificaron situaciones similares en turco, chino y finlandés).

Hay algunos cambios fáciles de realizar para que el lenguaje no presente estos sesgos discrimi-

33 Que el lenguaje sea “androcéntrico” supone poner en el centro al hombre e invisibilizar a la mujer (por ejemplo, refiriéndose a un grupo de individuos de ambos sexos, usando el género masculino: “los senadores”); que sea “patriarcal” implica reproducir lingüísticamente los supuestos de un régimen social patriarcal (por ej. en frases hechas como “Señores y señoras” o giros como “jefe de familia”); el uso “sexista” supone la superioridad del hombre sobre la mujer (por ej. si hablamos de “los empresarios” y “las secretarías”, cuando hay CEOs mujeres y secretarios varones).

minatorios cuando se trata simplemente de eliminar ciertas expresiones del lenguaje, por ej. eliminando el lenguaje peyorativo (*slurs*). Pero hay muchos términos que solo contextualmente son discriminatorios (como “negro”, no podemos simplemente eliminarlo de toda oración del español). Y los casos más complejos son aquellos en los que los cambios necesarios para hacer inclusivo al lenguaje suponen alteraciones gramaticales, por ejemplo, cuando están involucrados los pronombres personales, o cuando se usa -como ocurre en español- el genérico masculino. Es lo que sucede con expresiones como “los diputados” para referirse a un conjunto de diputados y diputadas (incluso cuando, de acuerdo a la legislación vigente en la Argentina, se debe respetar la paridad y tener al menos un 50% de representantes femeninas).



Laura Alonso Alemany

Grupo de Procesamiento del Lenguaje Natural, Facultad de Matemática, Astronomía, Física y Computación, Universidad Nacional de Córdoba, Argentina
<https://sites.google.com/view/etica-practica-cd/inicio>

“Con el uso creciente de aplicaciones de inteligencia artificial hemos visto crecer también los casos de discriminación de género. Por ejemplo, vimos cómo se calculaban riesgos crediticios mayores para mujeres porque los datos históricos no tenían suficientes casos de mujeres que hubieran devuelto sus créditos en tiempo y forma... ¡porque históricamente no se otorgaban créditos a mujeres! De esta forma, la automatización amplifica la inequidad, tal como desarrolla Virginia Eubanks en La automatización de la desigualdad (Eubanks 2021).

Pero además de amplificar la inequidad, las ubicuas aplicaciones de inteligencia artificial naturalizan estas desigualdades, dándoles una pátina de objetividad por el hecho de estar basadas en mecanismos y datos que imaginamos imparciales, desprovistos de subjetividad. Nada más lejos de la realidad. Estos mecanismos no dejan de ser artefactos desarrollados por grupos dominantes, que por lo tanto reproducen, amplifican y finalmente naturalizan sus valores, silenciando los valores de comunidades minoritarias y minorizadas.

De la misma forma que los sistemas automáticos no son objetivos, no existe una receta universal y objetiva para determinar cuáles de sus comportamientos son perniciosos. Lo que se considera discriminación de género en una cultura puede considerarse una libertad de género en otra, como por ejemplo la monogamia en culturas de base cristiana o de base islámica (Vasallo 2018). Es necesario empoderar a los diferentes grupos sociales locales para que inspeccionen el comportamiento de las aplicaciones, descubran los que les resultan perniciosos, los sistematicen y los denuncien. En ese sentido, en América Latina se están desarrollando las habilidades técnicas y los vínculos con la experticia y el territorio para naturalizar el involucramiento con estos sistemas”.



Cecilia Danesi

UBA / Universidad de Perugia/ Universidad de Salamanca

<https://www.ceciliadanesi.com/>

“Es importante contar con una legislación: existen muchos principios éticos como las directrices de la Unión Europea o las indicaciones de UNESCO pero al no ser vinculantes no tienen la fuerza de la ley no hay obligatoriedad en su cumplimiento y, por lo tanto, tampoco hay consecuencias. Esta legislación debería ser regional porque las plataformas no reconocen fronteras. Creo que la regulación de la Unión Europea de octubre de 2020 es muy interesante para investigar, porque toma principios éticos del desarrollo de la IA y los vuelve ley. El talón de Aquiles de América Latina es la falta de uniformidad y coordinación en sus leyes”.

4.2.2. Asistentes de voz

Los asistentes virtuales que utilizan la voz como comando se han vuelto muy populares en los últimos años gracias a las posibilidades que presenta poder interactuar con tecnología sin necesidad de tocar ningún dispositivo ni ocupar la vista en eso, permitiendo realizar en simultáneo otras tareas, como manejar un coche, cocinar o seguir trabajando sin interrupciones. El atractivo y potencial del reconocimiento automático del habla son tan altos que muchas de las grandes compañías de tecnología tienen proyectos vinculados con él. Para la empresa argentina Globant, por ejemplo, nos acercamos a una era “post pantalla” en la que todas las compañías de tecnología deben estar preparadas para que la interfaz del usuario se dé a través de la voz.³⁴

Sin embargo, es notorio comprobar que los proyectos más populares del momento -Siri, en los sistemas iOS de Apple; Alexa, del ecosistema de la empresa Amazon y Cortana, de Microsoft- feminizaron a sus asistentes de voz, dándole un nombre, una personalidad y un tono femenino en la voz que viene por *default*. Sus bautismos no han sido caprichosos. Siri fue bautizada por uno de sus creadores, el ingeniero noruego Dig Kittlaus, quien sugirió el nombre porque en nórdico significa “mujer hermosa que conduce a la victoria”; mientras que Cortana se inspiró en la saga de videojuegos Halo, en la que así se llama una inteligencia artificial encarnada en una sensual mujer de piel azul que recuerda a la reina egipcia Nefertiti. Alexa, en cambio, surgió del estudio de nombres de mujeres que fueran fáciles de distinguir para los algoritmos y que no fueran muy populares para evitar confusiones en los diálogos cotidianos. Jeff Bezos, el CEO de Amazon, estaba en contra de esa denominación y quería que las personas simplemente dijeran “Amazon” para interactuar con el software pero rápidamente fue disuadido por miembros de

34 The Sentinel Report 2019

su equipo, quienes le explicaron que los usuarios se sentirían más cómodos interactuando con alguien con nombre de mujer.³⁵

Esto tiene sus consecuencias: en los últimos años se multiplicaron los estudios que encontraron que personas de todo el mundo tenían un trato violento con sus asistentes virtuales de voz. En efecto, se ha establecido a través de exhaustivas investigaciones que la feminización de los asistentes virtuales genera que sean objeto de violencia. Esto implica que hay usuarios que asumen conductas violentas hacia estos dispositivos. Más allá del hecho trivial de que estos aparatos no son seres sintientes que puedan ser víctimas de violencia -a diferencia de las mujeres y otras identidades sexo-genéricas de carne y hueso-, es importante destacar varios hechos ligados a su existencia que reflejan la sociedad patriarcal en la que vivimos. Por un lado, es llamativo que estas plataformas, que se están multiplicando a nuestro alrededor y con las que es común crear lazos de mayor intimidad que con otros dispositivos, son maltratadas si son presentadas como mujeres. Estudios publicados detectaron que era común que los usuarios hicieran comentarios sexistas y dieran órdenes denigrantes a sus dispositivos, incluso sabiendo que eran imposibles de cumplir: la intención parecía ser molestar a esa voz femenina y comprobar qué cosas estaba dispuesta a hacer para complacer.³⁶ Estos pedidos son procesados e interpretados por las plataformas pero hay distintas estrategias programadas para ellos: mientras que algunos asistentes simplemente están programados para responder o decir que el comando no se comprende, en otras ocasiones se prestan al juego.

En el caso de Siri, por ejemplo, la frase “¡Me sonrojaría si pudiese!” fue la que los programadores decidieron poner como respuesta a insinuaciones con la palabra “*bitch*”. La decisión quedó en el ojo de la tormenta cuando UNESCO publicó un informe sobre tecnología y género que se llamó exactamente así (“*I’d Blush If I Could*”), forzando a que en la siguiente actualización de Siri se eliminara esa frase y se la reemplazara por un “No sé qué responder a eso”. Alexa, por su parte, hasta 2018 respondía “¡Gracias por el comentario!” cuando se le decía “puta” y ahora responde lacónica: “No estoy segura de qué es lo que esperás con esas palabras”. Como vemos, las decisiones de diseño relativas a cómo deben responder estos dispositivos al diálogo humano pueden amplificar o ayudar a corregir situaciones sociales de inequidad o violencia de género.

Los asistentes virtuales, además, reciben nombre y personalidad femenina cuando son diseñados para tareas hogareñas y de asistencia personal, pero si se trata de plataformas para pedir asesoramiento financiero sobre ámbitos como el de los automóviles y las motos, por ejemplo, sus nombres y personalidades son masculinas. Los estudios culturales feministas señalaron ya hace muchos años cómo los entornos artefactuales están generizados³⁷ pero la fascinación que nos producen las tecnologías digitales de inteligencia artificial en ocasiones nos hace olvidar cómo ellas también consolidan las normatividades de género.

4.2.3. Robots sociales

Aunque estamos más acostumbrados a interactuar con sistemas como Siri o Alexa, que no tienen una forma física definida, sino que son plataformas que pueden estar en nuestro telé-

35 Balmaceda, T. “Inteligencia artificial y patriarcal” - Página/12, 7 de enero de 2021

36 Silvervarg, Annika, et al. “The effect of visual gender on abuse in conversation with ECAs.” International Conference on Intelligent Virtual Agents. Springer, Berlin, Heidelberg, 2012.

37 Cfr. Wajcman 2004.

fono o en un parlante inteligente, también existen robots sociales, que son artefactos físicos con sistemas internos que fueron diseñados para interactuar y comunicarse con las personas de manera intuitiva y agradable siguiendo normas y patrones sociales. Se trata de dispositivos que suelen tener una apariencia humanoide y que tienen en su programación comportamientos que generan empatía en las personas que interactúan con ellos.

Los humanos tenemos una tendencia a proyectar sobre ciertos objetos, que tienen características en su apariencia y comportamiento, como si contasen con propiedades humanas como la intencionalidad o incluso personalidad sin dejar de reconocer que son máquinas. En los últimos años, varios estudios determinaron que también proyectamos en ellos nuestras concepciones sobre el género, algo reforzado por la creciente generización de los robots sociales.

Los indicadores de género se incorporan al diseño de robots de múltiples formas, desde la silueta de su forma hasta el nombre con el que es bautizado, el tono de su voz, las facciones de lo que entendemos como rostro y los colores con los que es fabricado.³⁸ No es necesario que haya múltiples elementos, ya que incluso un único rasgo es suficiente para desencadenar una interpretación de género que va acompañada de un conjunto de expectativas normativas de comportamiento.

Los robots sociales, además, son diseñados y creados para tareas específicas.³⁹ Y, como muchos tipos y ámbitos de trabajo están estereotipados para ciertos géneros, una y otra vez se refuerza la feminización de ciertas tareas, como robots sociales con nombre de mujer para tareas de cuidado o domésticas mientras que aquellos artefactos pensados para la seguridad de un centro comercial o un aeropuerto, por ejemplo, presentan rasgos masculinos.⁴⁰ De acuerdo con quienes diseñan estos dispositivos, debe existir una coincidencia entre lo que las personas esperan del rol que desempeña el robot social y su apariencia. El supuesto subyacente es que la apariencia del robot puede influir en la voluntad de las personas para cumplir con sus instrucciones e interactuar mejor. Nadie parece estar pensando que hacer coincidir la apariencia de los robots con las supuestas expectativas de los humanos refuerza las desigualdades de género y que, incluso si esas expectativas fueran ciertas y la interacción resultase más provechosa, diseñar robots masculinos para tareas que estereotípicamente son femeninas puede ayudar a combatir prejuicios y alentar la aceptación de varones humanos en esos roles.⁴¹

Otra opción, claro, es diseñar robots sin género pero hasta ahora ha demostrado ser un gran desafío no solo en el diseño sino también en la recepción que tiene en las personas, que toman incluso la característica más leve (por ejemplo, cómo “le suena” un nombre que ha sido inventado para no tener género o el tamaño de sus ojos) para automáticamente asignarle género.⁴² Para evitar los estereotipos humanos, algunos diseñadores de robots han optado por rechazar la antropomorfización pero sin perder la empatía que esto genera. El RIBA-II,

38 Kittmann et al., 2015

39 Dejamos de lado, en este trabajo, el análisis de los robots creados con fines sexuales, que generan mucha reflexión y debate pero que, por ahora, no cuentan con funciones robustas digitales vinculadas con el análisis del lenguaje o la conducta y que, por lo tanto, consideramos que no revisten relevancia para nuestros objetivos puntuales aquí.

40 Nass & Moon, 2000; Goetz et al., 2003; Carpenter et al., 2009; Eyssel & Hegel, 2012; Otterbacher & Talias, 2017

41 Eyssel y Hegel, 2012

42 Powers y Kiesler, 2006

por ejemplo, es un robot japonés de cuidados diseñado para levantar a adultos mayores o personas enfermas de su cama que se parece a un gran oso de peluche.

Si los humanos proyectamos en los robots expectativas de género sobre cómo deberían actuar las entidades “masculinas” y “femeninas”, surgen algunas preguntas:

¿Se debe promover la generización de este tipo de artefactos? ¿O deberíamos evitarla? ¿La generización cumple un rol en la eficacia de estos dispositivos en sus tareas y en la aceptación de los seres humanos de los mismos o refuerza los estereotipos de género que amplifican las desigualdades sociales?

5.

Conclusiones

Hemos realizado un recorrido por diversos asuntos teóricos y explorado ejemplos en los que se ha visto claramente cómo las sociedades configuran dispositivos tecnológicos que a su vez, cuando se lanzan al mercado, reconfiguran las sociedades que los crearon. Nunca resulta previsible la dimensión del impacto que las nuevas tecnologías tendrán en nuestras vidas y nuestras sociedades, pero como sabemos, tampoco es posible ni deseable detener el avance tecnológico. Sabiendo que las interacciones e influencias múltiples pueden ser de importantes dimensiones, debemos prestar mucha atención a estos posibles cambios, y adoptar en todo momento una actitud reflexiva y cauta respecto de las tecnologías emergentes.

Aplaudimos la agricultura, pero sabemos el daño que causa la quema de bosques autóctonos para ganar campos para el cultivo. Avanzamos en tecnologías que nos permiten la ganadería intensiva y las granjas piscícolas, pero sabemos del efecto invernadero que producen los gases emitidos por el ganado y la pérdida de biodiversidad que genera el exceso de antibióticos en las aguas. Con las tecnologías digitales pasa igual: sabemos que nos permiten conectarnos y simplificar tareas de una manera impensable hace cuatro décadas, pero también sabemos que muchas de ellas generan mecanismos de control social, división de nuestras sociedades difíciles de zanjar, *fake news* y alientan movimientos anti-vacunas que en tiempos de COVID-19 ponen en peligro nuestras vidas.

La filosofía, como dijimos al comenzar este trabajo, propone la adopción sistemática de una actitud reflexiva, crítica, propone seguir haciendo preguntas cuando ya nadie las hace. El rol de la filosofía es mantener despiertas a las sociedades para que nada resulte pasivamente aceptado, y la posibilidad de modificar el mundo, nuestras comunidades y nosotros mismos siempre sea una opción viable, a través del pensamiento reflexivo y la pregunta crítica.

Como mostramos a través de numerosos ejemplos en la parte III de este documento, las tecnologías digitales, entre muchos otros cambios en nuestras sociedades, están acentuando las inequidades de género, reforzando sistemáticamente conductas patriarcales, en parte porque seguimos viviendo en una sociedad patriarcal y estas tecnologías son fruto de estas sociedades, en parte porque las tecnologías mismas profundizan las desigualdades por los argumentos que esgrimimos en la parte II de este trabajo, pero también en parte porque no estamos reflexionando lo suficiente acerca de su diseño e implementación, como para mitigar estos efectos desde el diseño mismo.

Una forma siempre eficaz de generar esta actitud crítica y reflexiva es apuntar a una pluralidad de voces. Integrar distintas perspectivas políticas, filosóficas, disciplinares y sexo-genéricas en el diseño e implementación de estas tecnologías, y en las decisiones en el ámbito de las políticas públicas para adoptarlas, es un primer paso para atender a diversas inquietudes y fomentar el diálogo sobre estos temas. También esta pluralidad de voces debería estar presente en el diseño de legislaciones actualizadas que contemplen estos fenómenos y protejan a la ciudadanía de las posibles violaciones de derechos fundamentales que estas tecnologías pueden amplificar.

Tal vez sea cierto que tenemos la tecnología que nos merecemos. Pero también es cierto que podemos merecer algo mejor. Desde nuestra voz filosófica creemos haber contribuido a las discusiones en la esfera pública aportando al debate polifónico que debemos darnos. Entre otras cosas, que así sea depende del ejercicio crítico de nuestra imaginación.

6.

Bibliografía

Aibar, E. y W. Wijker. "Constructing a City: The Cerdá Plan for the Extension of Barcelona". *Science, Technology and Human Values*, vol. 22, n.1, pp. 3-30, 1997.

Bernabéu, C. (2020). *La colaboración de los artefactos en la producción de la normatividad de género*. Tesis doctoral. Universidad Autónoma de Madrid.

Bimber, Bruce. "Las tres caras del determinismo tecnológico". Smith, Merritt Roe y Marx, Leo, (eds.), *Historia y determinismo tecnológico*, Madrid, Alianza, 1996, pp. 95-115.

Bijker, Wiebe. "SocioHistorical Technology Studies, illustrated with examples from coastal engineering and hydraulic technology". En S. Jasanoff, G. E. Markle, J. C. Petersen, y T. Pinch (eds.). *Handbook of Science, Technology and Society*. London, Sage, ppp. 229-256, 1994.

Bijker, Wiebe; Hughes, Thomas y Pinch, Trevor, (eds.). *The Social Construction of Technological Systems*. Cambridge, MIT Press, 1987.

Bijker, Wiebe. *Of Bicycles, Bakelites, and Bulbs. Toward a Theory of Sociotechnical Change*. Cambridge (MA), MIT Press, 1995.

Bijker, Wiebe y J. Law (eds.). *Shaping Technology/Building Society*. Cambridge (MA). MIT Press

Brownstein, Michael, "Implicit Bias", *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2019/entries/implicit-bias/>>.

Buolamwini, J., Gebru, T. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." *Proceedings of Machine Learning Research* 81:1-15, 2018, Conference on Fairness, Accountability, and Transparency.

Crenshaw, K W. *On Intersectionality: Essential Writings*. New York. The New Press, 2017

Eubanks, V. *La automatización de la desigualdad. Herramientas de tecnología avanzada para supervisar y castigar a los pobres*, Capitán Swing, 2021.

Deva, S. *Open Global Rights. Afrontar el sesgo de género en la inteligencia artificial y la automatización*. 2020. Disponible en: <https://www.openglobalrights.org/addressing-gender-bias-in-artificial-intelligence-and-automation/?lang=Spanish>.

Hughes, Thomas. *Networks of Power: Electrification in Western Society*. Baltimore. John Hopkins University Press, 1983.

Kroes, P., P. Verbeek (eds.). *The Moral Status of Technical Artefacts*. Springer, 2014.

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

Noble, S. (2018) *Algorithms of oppression*. NYU Press, New York.

Pasquinelli, M., Joler, V. *El Nooscopio de manifiesto. laFuga*, 25, 2021.

Silvervarg, Annika, et al. "The effect of visual gender on abuse in conversation with ECAs." *International Conference on Intelligent Virtual Agents. Springer, Berlin, Heidelberg*, 2012.

Torrano, A., & Fischetti, N. *Filosofía feminista de la técnica y la tecnología. Notas para una academia latinoamericana activista. Pensando-Revista de Filosofía*, 11(23), 2020, 54-67.

Trescott, Martha. *Dynamos and Virgins Revisited. Women and technological change in history*, Metuchen: Scarecrow Press, 1979.

UNESCO, E. C. *I'd blush if I could: Closing gender divides in digital skills through education*, 2019.

Vasallo, B. *Pensamiento Monógamo: Terror Poliamoroso, La Oveja Roja*, 2018.

Vergés Bosch, Núria. "Teorías feministas de la tecnología: Evolución y principales debates", *Repositorio digital Universitat de Barcelona*, 2013 Disponible en: <http://diposit.ub.edu/dspace/handle/2445/45624>

Wajcman, Judy. *Feminism confronts technology, Pennsylvania: Pennsylvania State University Press*, 1991.

Wajcman, Judy. *TechnoFeminism, Cambridge, Massachusetts, Polity Press*, 2004.

Wajcman, Judy. "From Women and Technology to Gendered Technoscience", en *Information, Communication & Society*, 10:3, 2007, pp. 287-298.

Wajcman, Judy. "Feminist theories of technology", *Cambridge Journal of Economics*, Vol. 34, 2009, pp. 143-152.

Webster, Juliet. *Shaping Women's work. Gender, Employment and Information Technology*, London, Routledge, 1996.

Winner, Langdon. *Tecnología autónoma. La tecnología incontrolada como objeto del pensamiento político. Barcelona, Gustavo Gili*, 1979.

Descargo de responsabilidad. Las opiniones expresadas en la publicación incumben únicamente a los/as autores/as. No tienen intención de reflejar las opiniones o perspectivas del CETyS, CLD ni de ninguna otra organización involucrada en el proyecto.