

## ***ETIFICACIÓN, LA TRANSFORMACIÓN DIGITAL DE LO MORAL\****

*Patrici Calvo\*\**

<http://orcid.org/0000-0002-3228-9019>

*calvop@fis.uji.es*

**RESUMEN** *Actualmente, la convergencia de tres ámbitos de aplicación tecnológica —Big Data, Inteligencia Artificial e Internet de las Cosas— permite hiperconectar, algoritmizar y dataficar los distintos ámbitos de actividad humana para optimizar sus procesos productivos, comunicativos, asistenciales y/o relacionales, y, con ello, aumentar su eficiencia, eficacia y sostenibilidad. El éxito cosechado en sectores como el industrial o el retail y las enormes expectativas que abre su implementación en otros como el político o el asistencial, empero, ha extendido la idea de que incluso es posible discernir mejor qué es lo justo y bueno a través del análisis y procesamiento algorítmico de datos masivos sobre lo moral procedentes de los individuos hiperconectados. Sin embargo, este enfoque dataficado de ética alberga dudas razonables respecto a su viabilidad, posibilidad y eficacia. Pero, sobre todo, produce inquietud su aceptación y puesta en marcha acrítica por las importantes anomalías que produce en el mundo de la vida, especialmente en la concreción y perdurabilidad del saber moral. Este estudio pretende mostrar los retos, límites y consecuencias que subyacen a la dataficación del ámbito moral.*

**Palabras clave** *Algoritmización; dataficación; ética discursiva; hiperconectividad; sociedad digital.*

\* Artículo sometido em 26/09/2018 e aprobado em 01/04/2019. Este estudio es parte del Proyecto de Investigación Científica y Desarrollo Tecnológico FFI2016-76753-c2-2-p, financiado por el Ministerio de Economía y Competitividad, y uji-a2016-04, financiado por la Universitat Jaume I.

\*\* Universitat Jaume I. Castellón de la Plana, Comunidad Valenciana, España.

**ABSTRACT** *The convergence of three areas of technological application – Big Data, Artificial Intelligence and the Internet of Things – means the different area of human activity can be hyperconnected, algorithmised and datafied to optimise processes involving production, communication, healthcare and relationships. The aim is to increase their efficiency, effectiveness and sustainability. The success achieved in sectors like industry and retail and the huge expectations aroused by the implementation of this technology in other areas, such as politics and healthcare, have encouraged the idea that it might even be possible to find a better way of working out what is good and fair in the algorithmic analysis and processing of Big Data concerning morality drawn from hyperconnected individuals. There are reasonable doubts, however, concerning the possibility, viability and effectiveness of this datafied approach to ethics. Above all, a critical acceptance and implementation of it raises concerns about the considerable anomalies it causes in the world of life, particularly in the accumulation and durability of moral knowledge. This study attempts to show the challenges, limits and consequences underlying the datafication of morality.*

**Keywords** *Algorithmisation; datafication; discursive ethics; hyper-connectivity; digital society.*

## 1. Introducción

Actualmente, la sociedad presencia imparable y acriticamente la *hiperconectivización, dataficación y algoritmización* de todas las esferas de actividad humana; una tendencia —por imposición tecnológica y económica— que está transformando los procesos productivos, educativos, asistenciales, comunicativos, selectivos, decisorios y democráticos, y, lamentablemente, no siempre para mejor.

Este curso de acción ha sido posible gracias a la posibilidad e implementación del Internet de las Cosas (IoT). Definido por la Comisión Europea como el conjunto de “Cosas con identidades y personalidades virtuales que operan en espacios inteligentes usando interfaces inteligentes para conectarse y comunicarse dentro de contextos sociales, ambientales y de usuario” (European Commission, 2008, p. 4), el IoT ha permitido la convergencia de las diversas y versátiles tecnologías de aplicación, como las *Key Enabling Technologies* (KET)<sup>1</sup>, la Inteligencia Artificial (AI) y el Big Data

1 Según la European Commission (2009), las Tecnologías Facilitadoras Esenciales (KET) aluden a, principalmente, la micro y nanoelectrónica, las nanotecnologías, la biotecnología, los materiales avanzados y la fotónica.

(BD), y, con ello, la posibilidad de recopilar y procesar mediante modelos matemáticos datos masivos de la actividad de los ciudadanos hiperconectados en cualquier esfera de actividad en la que se muevan para, entre otras cosas, tomar decisiones supuestamente racionales que permitan la optimización de los procesos implicados.<sup>2</sup>

Este hecho, junto al discurso embriagador que acompaña la aplicación y desarrollo del IoT en las diferentes esferas de actividad humana,<sup>3</sup> amplificado a través de los medios de comunicación en forma de noticia, publicidad o información científica, ha impulsado la rápida expansión y aceptación de la *hiperconectividad digital* de la sociedad que lo posibilita por parte tanto del público informado como de la ciudadanía en general. Al respecto, destaca el alegato favorable promulgado por la Comisión Europea, para quien “La interconexión de objetos físicos dará lugar a un auténtico cambio en el modelo de sociedad”, puesto que según su opinión permitirá “[...] responder a muchos de los retos actuales, como el envejecimiento de la población, la deforestación o la emisión de CO2 gracias, sobre todo, al desarrollo de sistemas de vigilancia sanitaria, de conexión de árboles y de automóviles” (European Commission, 2009).

Sin embargo, el proceso de transformación digital que genera la implementación y desarrollo del IoT en los distintos ámbitos de actividad produce consecuencias inintencionadas —o intencionadas— altamente corrosivas para los usuarios del ecosistema ciberfísico y la sociedad en general. Como afirma Cathy O’Neil (2016, pp. 16, 20-21), una de las consecuencias más negativas de este mundo cada vez más hiperconectado es la exacerbación de las desigualdades en todas sus dimensiones, especialmente su peculiar tendencia a beneficiar a los privilegiados y castigar a los desfavorecidos. Pero, sobre todo, porque su expansión y acrítica aceptación está produciendo una sutil colonización sistémica del mundo de la vida, generando distorsión y crisis en sus diferentes componentes estructurales —cultura, sociedad y personalidad— tal y como advirtió Jürgen Habermas (1987). Especialmente en ese saber moral<sup>4</sup> que intuitivamente hacen servir los participantes en aquellos procesos de diálogo tendentes al entendimiento intersubjetivo sobre algo en el mundo, que ahora parece concretarse a través de cursos de acción socialmente atomizados y modelos matemáticos de procesamiento de la información disponible. Un

2 Para un estudio sobre las implicaciones económicas y sociales de la transformación digital, ver Zamagni (2018).

3 El IoT está generando alta expectativa de lograr algún tipo de beneficio, mejora, diferenciación o ventaja competitiva (Raskino, 2011).

4 Para un estudio sobre el saber moral y su función social, ver García-Marzá (2016, 2005).

saber moral cuya función social es, precisamente, servir como elemento “[...] protector ante la extrema vulnerabilidad que nos caracteriza como personas” (García-Marzá, 2016, p. 882).

Estas y otras cuestiones sobre el proceso de digitalización del mundo de la vida —especialmente el social, donde los participantes se proveen de normas legítimas para coordinar las acciones interpersonales y asegurar la solidaridad gracias al diálogo— exigen de una reflexión profunda sobre las implicaciones éticas de la reducción del ámbito moral a regulaciones empíricas cuantificables. Precisamente, el objetivo de este estudio será ahondar críticamente en el proceso de *datafización* de la ética que se está pergeñando actualmente para mostrar los retos, límites y consecuencias que le subyacen. Para ello, en primer lugar, se mostrará en qué consiste un enfoque *datafizado* de ética como el que se está proponiendo y diseñando actualmente desde diferentes ámbitos. En segundo lugar, se mostrará cómo esta tendencia a la *datafización* del saber moral puede producir alteraciones y anomalías importantes en los diferentes componentes estructurales del mundo de la vida. Finalmente, se mostrarán algunos límites del enfoque *datafizado* de ética desde una propuesta ético-crítica.

## 2. Etificación, la hiperdatafización de lo moral

La *datafización*, la posibilidad de monitorizar y medir de forma precisa y en tiempo real las preferencias, opiniones y comportamientos de los consumidores y usuarios del sistema, es la consecuencia más visible de la tendencia actual a la digitalización de los procesos productivos, asistenciales, paliativos, relacionales, decisorios y comunicativos implicados en el correcto desarrollo de las diferentes esferas de actividad humana. Esta surge gracias a la convergencia de tres ámbitos tecnológicos —el Big Data, la Inteligencia Artificial y el Internet de las Cosas<sup>5</sup>— y ha logrado expresar rápidamente el enorme potencial estratégico y predictivo que atesora en ámbitos como el industrial, el *retail* y el sanitario.

El concepto *datafication* es un neologismo acuñado por Kenneth Neil Cukier y Victor Mayer-Schöenberger en *Big Data. A Revolution that will transform how we live, work, and think* (2013, pp. 73-97) con el propósito de etiquetar los procesos de transformación de la acción social en datos

<sup>5</sup> La emergencia de estos tres ámbitos tecnológicos ha sido posible gracias a la complementación de las *Key Enabling Technologies* (KET) un grupo de tecnologías compuesto por la micro y nanoelectrónica, nanotecnología, la biotecnología industrial, los materiales avanzados, la fotónica y las tecnologías avanzadas de fabricación.

cuantificados en línea que permiten a las empresas y agencias gubernamentales llevar a cabo un seguimiento y análisis predictivo en tiempo real (van Dijck, 2014, p. 198). Como argumenta José van Dijck “[...] exploran los montones de metadatos que crecen exponencialmente a través de redes sociales y plataformas de comunicación, como Facebook, Twitter, LinkedIn, Tumblr, iTunes, Skype, WhatsApp, YouTube y servicios gratuitos de correo electrónico como gmail y hotmail, con el fin de rastrear información sobre el comportamiento humano” (van Dijck, 2014, p. 198). De este modo, actualmente los gobiernos, organismos y empresas entienden la *dataficación* —o *datificación*— como “[...] un medio legítimo para acceder, comprender y monitorizar el comportamiento de las personas que se está convirtiendo en un principio rector, no solo entre los tecnópatas, sino también entre los académicos que ven la dataficación como una oportunidad de investigación revolucionaria para investigar la conducta humana” (van Dijck, 2014, p. 198).

El potencial estratégico y predictivo que se le presupone a la aplicación de procesos de *dataficación* en las distintas esferas de relacionalidad humana y el tipo de datos que le interesan —conductuales— ha dado pie a un preocupante intento por *dataficar* o *datificar* lo moral —*etificación*— en ámbitos como la política o la economía (Morozov, 2013; Bostrom, 2014; Matsumoto, 2018; Malle, 2016; Jentzsch et al., 2019). Compuesto por la palabra «ética» por un lado y «cuantificación» en sentido amplio<sup>6</sup> por otro, la *etificación* (*ethification* en inglés) es la acción de recopilar, registrar, cuantificar, procesar y gestionar datos masivos sobre las opiniones, preferencias y conductas de la ciudadanía hiperconectada para, desde el criterio del mayor bien para la mayor cantidad de gente posible, establecer mediante sistemas inteligentes qué es moralmente válido<sup>7</sup>. De este modo, el sentido sobre las diferentes cosas de este mundo, como diría Habermas (1987), ya no depende de aquellos procesos de diálogo donde los afectados por una norma, acción y decisión se ponen de acuerdo sobre las consecuencias derivadas de su aplicación fáctica, sino de la agregación y entrelazamiento de todas las preferencias, opiniones, intereses y hábitos individuales expresadas por la ciudadanía hiperconectada.

6 Aquí el término «cuantificación» encierra un sentido mucho más amplio que la mera expresión numérica de una magnitud. Este tiene que ver con los distintos procesos que subyacen a la transformación digital —dataficación, hiperconectivización y algoritmización— cuyo principal objetivo es procesar datos masivos para convertirlos en conocimiento aplicable a través de modelos matemáticos.

7 Para observar esta tendencia a través de propuestas concretas de aplicación, véase Whitters (2018); Sáez (2018); Motta et al. (2016), entre otros, y experimentos de AI aplicada tales como *Moral Machine* (MediaLab, 2019).

La reflexión ética, por consiguiente, ve alterada su legítima pretensión de convertirse en un acuerdo intersubjetivamente válido y recíprocamente criticable. Por un lado, porque la *etificación* elimina toda posibilidad de concretar en la práctica una acción comunicativa, tendente al entendimiento sobre lo justo y verdadero, al reducir el diálogo a una mera expresión de opiniones subjetivas vertidas por los diferentes elementos que participan de la singularidad de una sociedad hiperconectada concreta o general. Y, por otro, porque el acuerdo es externo al diálogo, establecido *a posteriori* por modelos matemáticos diseñados para tal fin, y cuya concreción responde a una agregación de opiniones subjetivas y conductas convencionales.

Además, la ética como disciplina, en tanto que saber práctico con pretensiones de universalidad que intenta orientar de manera racional las acciones y decisiones en todos los ámbitos de actividad humana (Cortina, 1986, 1990, 2013; Cortina y Martínez, 1996), pierde su sentido interno diluida entre saberes y métodos propios de otros ámbitos de conocimiento cuyo fin no es prescribir la realidad comportamental del ser humano, sino describirla para predecirla y, de ese modo, controlarla o dominarla. La *etificación*, en tanto que *datafificación* de la ética como disciplina, pretende hacer de esta una ciencia predictiva basada en el procesamiento y cruce de datos masivos sobre la conducta humana a través de modelos matemáticos capaces de establecer de forma precisa y en tiempo real qué es lo justo y deseable para una sociedad y prever los comportamientos de las personas en sus diferentes espacios de acción<sup>8</sup>. Como manifestó Chris Anderson durante los primeros momentos de este mundo cada vez más *datafificado*, cuando sólo se vislumbraba una pequeña parte de su potencialidad actual, existen claras evidencias de que todos los modelos científicos sobre la conducta humana están obsoletos y, por tanto, representan teorías erróneas. De ahí la necesidad, según su opinión, de empezar a remplazarlos por propuestas *datafificadoras* del comportamiento que ofrezcan mayor claridad, exactitud y predictibilidad:

Este es un mundo en el que cantidades masivas de datos y matemáticas aplicadas reemplazan cualquier otra herramienta que pueda utilizarse. Fuera con cada teoría del comportamiento humano, desde la lingüística hasta la sociología. Olvida la taxonomía, la ontología y la psicología. ¿Quién sabe por qué la gente hace lo que hace? El punto es que lo hacen, y podemos rastrearlo y medirlo con una fidelidad sin precedentes. Con suficientes datos, los números hablan por sí mismos (Anderson, 2008).

8 Resultaría de enorme interés para este trabajo poder abordar el funcionamiento y potencialidad de la Inteligencia Artificial para mostrar si cabe la posibilidad de pensar en nuevas formas de toma de decisiones éticas, sin embargo, la complejidad del tema y las normales restricciones en cuanto a la extensión de las publicaciones científicas obligan a que tales cuestiones sean abordadas en otros trabajos. Véase al respecto Calvo (2019a, 2019b).

La *etificación*, empero, parece más bien la consecuencia directa de la aplicación de imperativos sistémicos sobre los distintos componentes estructurales del mundo de la vida: la cultura (mundo objetivo), la personalidad (mundo subjetivo) y la sociedad (mundo social). Contrariamente a la visión deformada que defienden tecnópatas y académicos oportunistas que ven en ello una oportunidad para imponer su ideología, su análisis crítico sugiere que se trata, tal y como advirtió Habermas (1984, p. 88), de “[...] un progreso cuasi-autónomo de la ciencia y de la técnica, del que de hecho depende la otra variable más importante del sistema, es decir, el progreso económico”, que produce anomalías importantes en el mundo de la vida —pérdida de sentido, déficit de normas y psicopatologías— poniendo en jaque tanto la necesaria autonomía de sus elementos estructurales como el correcto desarrollo de sus propios procesos internos de racionalización<sup>9</sup> (Habermas, 1987, pp. 279-280).

### 3. Etificación, la colonización algorítmica del mundo de la vida

En el segundo volumen de la *Teoría de la acción comunicativa* (1987), Habermas parte de la necesidad de superar la perspectiva dual de la teoría social. Tanto la *perspectiva de la integración social*, vinculada con la expresión mundo de la vida, como la *perspectiva de la integración del sistema*, vinculada con la expresión sistema, han aportado mucho a la Teoría Social (Habermas, 1987, pp. 213-214). Sin embargo, por separado ambas presentan graves limitaciones y deficiencias. Por ello, Habermas construye su teoría desde la necesidad de encontrar un modo de “[...] articular de forma satisfactoria esas dos estrategias conceptuales caracterizadas por las expresiones «sistema» y «mundo de la vida»” (Habermas, 1987, p. 215).

Desde esta perspectiva, Habermas presenta una propuesta radical y crítica de sociedad estructurada alrededor de dos niveles de aplicación autónomos pero complementarios y necesarios: el sistema y el mundo de la vida. Por un lado, el sistema está conformado por tres subsistemas —economía, estado y familia, que se proyectan y desarrollan a través de procesos de racionalización con arreglo a fines. Por otro lado, el mundo de la vida está conformado por tres esferas de actividad humana —cultura, sociedad y personalidad— que dan cuenta del mundo objetivo —en tanto que totalidad de las entidades sobre las que son posibles enunciados verdaderos—, mundo social —en tanto que totalidad de las relaciones interpersonales legítimamente reguladas— y mundo

9 Para un estudio pormenorizado de la teoría habermasiana, ver García-Marzá (1992).

subjetivo —en tanto que totalidad de las propias experiencias/vivencias— a través de procesos internos de racionalización mediados por la comunicación.

Llamo *cultura* al acervo de saber, en que los participantes en la comunicación se abastecen de interpretaciones para entenderse sobre algo en el mundo. Llamo *sociedad* a las ordenaciones legítimas a través de las cuales los participantes en la interacción regulan sus pertenencias a grupos sociales, asegurando con ello la solidaridad. Y por *personalidad* entiendo las competencias que convierten a un sujeto en capaz de lenguaje y de acción, esto es, que lo capacitan para tomar parte en procesos de entendimiento y para afirmar en ellos su propia identidad. El campo semántico de los contenidos simbólicos, el espacio social y el tiempo histórico constituyen las *dimensiones* que las acciones comunicativas comprenden. El entretrejimiento de interacciones de que resulta la red de la práctica comunicativa cotidiana constituye el medio a través del que se reproducen la cultura, la sociedad y la persona. Tales procesos de reproducción sólo se refieren a las estructuras simbólicas del mundo de la vida. De ellos hemos de distinguir el mantenimiento del sustrato material del mundo de la vida. (Habermas, 1987, pp. 196-197)

Definido por Habermas como “[...] un acervo de patrones de interpretación transmitidos culturalmente y organizados lingüísticamente” (1987, p. 176), el mundo de la vida constituye “[...] un depósito de autoevidencias o de convicciones incuestionadas, de las que los participantes en la comunicación hacen uso en los procesos cooperativos de interpretación” (1987, p. 176), que se desarrolla a través “[...] de las estructuras de la intersubjetividad lingüística y se mantiene a través del mismo medio en que los sujetos capaces de lenguaje y de acción se entienden entre sí acerca de algo en el mundo” (1987, p. 182). Es decir, este subsistema representa para Habermas el ámbito trascendental donde los interlocutores válidos se aprovisionan de los elementos necesarios para, recíprocamente, reclamar su lugar en el mundo y exigir el reconocimiento de su competencia y capacidad para criticar o confirmar la validez de las pretensiones, abordar sus discrepancias y llegar a acuerdos sobre el sentido de algo en el mundo a través del lenguaje; donde, en definitiva “[...] hablante y oyente se entienden desde, y a partir de, el mundo de la vida que les es común, sobre algo en el mundo objetivo, en el mundo social y en el mundo subjetivo” (1987, p. 179).

Por ello, en tanto que al mundo de la vida le es constitutivo el entendimiento *como tal* (Habermas, 1987, p. 179), es a este al que le corresponde definir la consistencia del sistema social en su conjunto. De ahí que para Habermas los mecanismos sistémicos tengan la necesidad de institucionalizarse, de anclar sus raíces en el mundo de la vida para no quedar desdibujados tras déficits de sentido y legitimidad que ponen en jaque su supervivencia y desarrollo (Habermas, 1987, p. 217). No obstante, la cada vez mayor complejidad del

mundo de la vida ha generado un proceso de colonización mediante imperativos sistémicos como el poder y el dinero que produce anomalías importantes que ponen en jaque su autonomía, desarrollo y subsistencia, como anomia, psicopatologías o alienación.

Para Habermas, en las sociedades arcaicas el sistema y el mundo de la vida se hallaban entrelazados. No había diferencias significativas entre ambas esferas (Habermas, 1987, pp. 219-220). La modernidad, empero, produjo su gradual «desacoplamiento», y, con este, la colonización sistémica de los componentes estructurales del mundo de la vida: mundo objetivo, mundo social y mundo subjetivo. Este hecho se debe a que, fruto de la evolución de la sociedad feudal del Medievo a la sociedad económica de la Modernidad, la mayor racionalización de cada una de las esferas y sus componentes estructurales generó un incremento gradual de su complejidad que produjo su desacoplamiento y progresivo distanciamiento, y, con ello, la emergencia de un proceso de colonización sistémica del mundo de la vida, cuya estructura es comunicativa, que produjo crisis.

Esto se debe a que, por un lado, el proceso de racionalización de la acción de cada una de las partes adoptó formas diferentes. Mientras la evolución del sistema se centra en la *reproducción material* y se halla determinada por una acción racional con respecto a fines —estratégica— el desarrollo del mundo de la vida se centra en la *reproducción simbólica* y se encuentra mediado por una acción racional con arreglo a sentido —comunicativa. Por otro lado, la racionalización del mundo de la vida exige un aumento de la acción comunicativa, aquella donde los participantes persiguen de común acuerdo sus respectivos planes de acción sobre la base de una definición común de la situación (Habermas, 1987, pp. 180-181), y, con ello, un incremento proporcional de las demandas en el lenguaje que agotan sus capacidades para la reproducción simbólica e integración social. Este déficit es aprovechado por los componentes estructurales del sistema para desplazar mediante imperativos sistémicos los procesos de integración social del mundo de la vida, que ahora son controlados y marginados por procesos de integración sistémica que mediatizan, instrumentalizan y cosifican el mundo de la vida a través del poder y el dinero. Como argumenta Habermas,

Si esta tendencia evolutiva hacia el desacoplamiento de sistema y mundo de la vida se la proyecta sobre el plano de una historia sistémica de las formas de entendimiento, queda de manifiesto la incontenible ironía del proceso histórico universal de Ilustración: la racionalización del mundo de la vida hace posible un aumento de la complejidad sistémica, complejidad que se hipertrofia hasta el punto de que los imperativos sistémicos, ya sin freno alguno, desbordan la capacidad de absorción del mundo de la vida, el cual queda instrumentalizado por ellos. (1987, p. 219)

En este sentido, Habermas destaca cuatro mecanismos de integración sistémica —diferenciación segmentaria, estratificación, medio de control sistémico, organización estatal— cuya emergencia dentro de un proceso evolutivo restringe la comunicación generando anomalías en un mundo de la vida cuya estructura es comunicativa, tales como pérdida de sentido en el mundo objetivo, anomia en el mundo social y psicopatologías en el mundo subjetivo (Habermas, 1987, pp. 234-235). Son los efectos más visibles del proceso de colonización sistémica del mundo de la vida del sistema capitalista, un fenómeno que estrangula los cursos de entendimiento a través de la violencia ejercida mediante la coerción del poder y el dinero sobre los espacios comunicativos.

Cuando se introduce la violencia (Gewalt) como alternativa al mecanismo de coordinación de la acción que representa el entendimiento y el poder (Macht) como producto de la acción orientada al entendimiento, se obtiene, además, la ventaja de no perder de vista las formas de ejercicio indirecto de la violencia que hoy predominan. Me refiero a esa violencia patógena que inadvertidamente penetra en los poros de la práctica comunicativa cotidiana y puede desplegar en ella su latente eficacia en la medida en que el mundo de la vida queda entregado a los imperativos de subsistemas funcionales autonomizados y cosificados por las sendas de una racionalización unilateral. (Habermas 1987, pp. 458-459)

Actualmente, este proceso de colonización sistémico del mundo de la vida viene determinado por la cada vez mayor tecnificación y cosificación de sus componentes estructurales mediante la *dataficación* de la producción simbólica, la imposición de modelos decisionistas —que afrontan la resolución de conflictos de forma dilemática— frente a pragmatistas —que se enfrentan a la resolución de conflictos de manera problemática—, así como la disolución de los cursos comunicativos de acción necesarios para criticar y validar tanto el conocimiento como la acción, entre otras cosas. Como advierte Habermas, “El cálculo llevado al extremo, que caracteriza a estos medios de decisión, deja en estado de pureza a lo que son decisiones, es decir, las limpia de todos aquellos elementos que aún podían considerarse accesibles a algún análisis de tipo vinculante” (1984, p. 136). Todo ello provoca perturbaciones en el ámbito de la reproducción cultural, la reproducción social y la socialización que en el mundo social produce pérdida de legitimación, carencia de normas vinculantes y escasez de motivaciones para la acción que ponen en jaque su racionalidad, solidaridad y autonomía<sup>10</sup> (Habermas, 1987, p. 203). El resultado de todo ello es, como argumenta Habermas,

10 Para un estudio pormenorizado de la teoría habermasiana, ver García-Marzá (1992).

[...] una perspectiva en la que la evolución del sistema social parece estar determinada por la lógica del progreso científico y técnico. La legalidad immanente de este progreso es la que parece producir las coacciones materiales concretas a las que ha de ajustarse una política orientada a satisfacer necesidades funcionales. Y cuando esta apariencia se ha impuesto con eficacia, entonces el recurso propagandístico al papel de la ciencia y de la técnica puede explicar y legitimar por qué en las sociedades modernas ha perdido sus funciones una formación democrática de la voluntad política en relación con las cuestiones prácticas y puede ser sustituida por decisiones plebiscitarias relativas a los equipos alternativos de administradores. (Habermas, 1984, p. 88)

Uno de los elementos que muestra más claramente la interferencia tecnológica en el mundo de la vida es sin duda la *etificación*, el actual proceso de *datafificación* del ámbito moral para la concreción del saber objetivado y matemáticamente válido. Esta tendencia no sólo interfiere en el correcto desarrollo de los procesos comunicativos capaces de concretar qué normas, valores, sentimientos, virtudes y pautas son moralmente válidas, desplaza el posible consenso entre afectados por un análisis cuantitativo de la información subjetiva disponible. También convierte el saber moral en científico-técnico con capacidad predictiva, eliminando con ello su interés emancipatorio y su carácter prescriptivo de la realidad social. Sin embargo, varias cuestiones ponen en duda la validez y viabilidad de un enfoque *datafificado* de ética.

#### **4. Etificación, la reducción de lo moral a regularidades empíricas**

La viabilidad de un enfoque *datafificado* de fundamentación de lo moral como el que se está pergeñando desde algunos sectores y/o teóricos vinculados con la transformación digital, que pretende cuantificar mediante algoritmos las preferencias, opiniones y hábitos del colectivo digitalmente hiperconectado para elucidar qué acciones, normas y decisiones podrían contar con el beneplácito de la mayoría, ofrece serias dudas al respecto. Al respecto, destacan varias cuestiones.

Primero, los algoritmos identifican con más o menos éxito las *señales lingüísticas* de la sociedad hiperconectada para reconstruir el conjunto de *identidad de significados* de una comunidad lingüística. Sin embargo, en ningún caso pueden establecer la validez intersubjetiva de la cual emana la fuerza motivadora de la acción. Como afirma Habermas siguiendo a Ludwig Wittgenstein principalmente, el significado de una regla no puede explicarse con arreglo a hechos empíricamente observados (1987, pp. 28-29). En primer lugar, porque esos hechos también pueden ser explicados por la regla; y, en segundo lugar, porque la comprensión de una acción simbólica está ligada a la capacidad y disposición que muestran al menos dos sujetos a seguir esa regla. Por tanto,

[...] la identidad de significados se remonta a la capacidad de seguir reglas intersubjetivamente válidas junto con, a lo menos, otro sujeto más; ambos han de disponer, así de la competencia de ajustar su comportamiento a reglas, como de la de enjuiciar críticamente tal comportamiento. Un sujeto aislado y solitario, que sólo dispusiera además de una sola de esas dos competencias, ni podría formar el concepto de regla, ni emplear símbolos con identidad de significado. (Habermas, 1987, p. 32)

Segundo, la *etificación* confunde la *identidad de un significado* con la *identidad de un objeto*. No es lo mismo. La *identidad de un objeto* se establece mediante un acuerdo entre sujetos sobre las características convencionales que describen ese objeto. De este modo, esta se reduce a regularidades empíricas (Habermas, 1987, p. 30). Sin embargo, la *identidad de significado* no puede reducirse a regularidades empíricas. Esta depende sobre todo de la validez intersubjetiva, lo que exige *crítica recíproca e instrucción mutua tendente al entendimiento* sobre su significado y validez de, al menos, dos interlocutores para asegurar la identidad de la regla que la subyace y la posibilita; es decir, de la “[...] circunstancia de que a) sujetos que orientan su comportamiento por reglas se desvían de ellas, y de que b) pueden criticar ese su comportamiento desviante como violación de las reglas” (Habermas, 1987, p. 31). La *etificación*, empero, confunde ambos objetivos identitarios y reduce las reglas a regularidades empíricas.

Tercero, la *etificación* se desarrolla desde la perspectiva de un observador omnisciente y neutro que no interactúa críticamente y, por tanto, su capacidad para seguir una regla es impuesta por la situación contextual, ajena o externa a su propia voluntad<sup>11</sup>. No sigue la regla porque la considera válida y moralmente vinculante, sino porque la descripción de los hechos empíricos merced al estudio de los datos masivos sobre las preferencias, opiniones y hábitos de las personas convencionalmente así lo exige. El algoritmo que *datafica* lo moral no

11 El funcionamiento del Internet de las Cosas y los algoritmos que los gobiernan muestran que reglas obtenidas por las máquinas algorítmicas provienen de los grandes volúmenes de datos que se generan en la esfera pública vinculados con los intereses, deseos y actos de los sujetos. Se puede decir, por tanto, que esta nueva forma de *relacionalidad* es predictiva y descriptivamente emocional, ya que se anticipa al deseo subjetivo del usuario/ciudadano. Sin embargo, este artículo pretende defender la idea de que ella no es suficiente para ser constitutiva y prescriptivamente moral, puesto que el consenso sobre lo justo no está relacionado con la cuantificación de las preferencias, opiniones e intereses de la ciudadanía digitalizada. El consenso moral se alcanza mediante un diálogo crítico tendente al entendimiento sobre diferentes cosas de este mundo donde las preferencias, opiniones e intereses de todos los afectados presentes y futuros deben tanto ser tenidas en cuenta como estar expuestas a revisión por parte de los demás. La idea es, precisamente, mostrar que el análisis algorítmico es puramente convencional. Éste no critica lo que es. Recopila, cuantifica y analiza la realidad vigente para, posteriormente, tomar decisiones supuestamente justificadas amparándose en un criterio utilitarista de fundamentación de lo moral. Pero, desde mi punto de vista, esto ni es suficiente ni tiene que ver con lo moral. La ética es siempre postconvencional y crítica. Es decir, su pretensión es ir más allá de los intereses, opiniones y preferencias personales y colectivas para adentrarse en el terreno de la validez intersubjetiva.

participa de ningún diálogo reflexivo tendente al entendimiento ni sus acciones se pueden explicar por la regla que se constata y/o elucubra de su análisis de la realidad.

Cuarto, la *etificación* desplaza el interés *práctico* y *emancipatorio* del conocimiento que subyace a las ciencias hermenéuticas y críticas —legitimador y felicitante— por el interés *técnico* característico de las ciencias empírico-analíticas —estratégico y predictivo. La consecuencia lógica de todo ello es un desarrollo social determinado por el progreso científico-técnico y legitimado por la presunta potencialidad estratégica y predictiva que atesora. Es decir, la *etificación* es otro intento más de reducir la dominación práctica de la historia al control técnico de los procesos objetivados (Habermas, 1984, p. 179). En palabras de Habermas:

La legalidad inmanente de este progreso es la que parece producir las coacciones materiales concretas a las que ha de ajustarse una política orientada a satisfacer necesidades funcionales. Y cuando esta apariencia se ha impuesto con eficacia, entonces el recurso propagandístico al papel de la ciencia y de la técnica puede explicar y legitimar por qué en las sociedades modernas ha perdido sus funciones una formación democrática de la voluntad política en relación con las cuestiones prácticas y puede ser sustituida por decisiones plebiscitarias relativas a los equipos alternativos de administradores. (Habermas, 1984, p. 88)

Quinto, la *etificación* elimina la posibilidad de *crítica recíproca* y *mutua instrucción* tendente a un acuerdo intersubjetivo sobre el significado y la validez tanto del conocimiento como de la acción. Desde una ética discursiva de fundamentación de lo moral como la propuesta por Habermas y Karl-Otto Apel, sólo es posible calificar una acción, norma o decisión como moralmente válida cuando sus consecuencias son aceptadas sin coacción por todos sus afectados dentro de un discurso práctico con ciertas reglas lógicas y un principio moral procedimental. Una propuesta ética basada en la *dataficación*, empero, sustituye el diálogo crítico e instructivo entre interlocutores válidos afectados por la norma, acción o decisión tomada o a tomar por conocimiento analítico-empírico sobre las preferencias, opiniones y hábitos de la ciudadanía, esté o no afectada. De esta forma, renuncia a la legítima pretensión de universalidad de la ética discursiva para centrar su interés en la toma de decisiones prácticas. Pero, ante la complejidad y el riesgo derivado de las nuevas realidades que emergen del proceso de transformación digital, [...] ya no nos sirve un saber instrumental o estratégico, ni siquiera las certezas de un saber científico directamente vinculado con las mismas tecnologías” (García-Marzá, 2016, p. 882). Como argumenta Domingo García-Marzá (2016, pp. 882-883) las nuevas realidades derivadas de las revoluciones tecnológicas, como la actual transformación digital, aumentan

la vulnerabilidad de las personas y, por tanto, incrementa proporcionalmente la exigencia del saber moral, cuya función social es precisamente minimizar la vulnerabilidad y sus consecuencias aportando recursos que permiten tanto orientar la praxis para evitar la emergencia de conflictos como coordinar la acción tendente a la satisfacción de objetivos comunes y la resolución de la conflictividad subyacente. Y todo ello exige decisiones adecuadas, pero sobre todo reflexiones previas que generen el sentido necesario para poder calificar esas decisiones como correctas.

Sexto, la *etificación* basa el discernimiento sobre lo que es moralmente válido en el criterio utilitarista del mayor bien para la mayor cantidad de gente posible. En tanto que evita el diálogo tendente al posible acuerdo sobre diferentes cosas de este mundo, aplica un cálculo utilitarista sobre las preferencias y las posibles consecuencias de la acción, norma o decisión basado en aquello que la mayoría entiende que es justo o correcto. El problema, empero, es que las opiniones, preferencias y pautas de la mayoría de los individuos de una sociedad pueden ser muy injustas y, por tanto, inmorales. Que la mayoría de los individuos de una sociedad esté de acuerdo con la pena de muerte, las desigualdades por razón de sexo, pensamiento o etnia y la insolidaridad con los más desfavorecidos, por ejemplo, no legitima tales opiniones. La *etificación*, si la legislación lo permite, sí. Pero, como se ha comentado, sólo la aceptación por parte de todos los afectados de las consecuencias derivadas de estas ideas, ya sean normativas o prácticas, puede legitimarlas.

Séptimo, la *etificación* se basa en el *atomismo social* como fuente de estudio, análisis y comprensión del ámbito moral. Planteado por Carl Menger en *Grundsätze der Volkswirtschaftslehre* (1871), comprobado por Kenneth J. Arrow (1951) mediante el *teorema de imposibilidad*, y basado en el falso axioma del *homo oeconomicus*, el enfoque de atomismo social entiende que la sociedad es un conjunto de personas autointeresadas unidas por la posibilidad de satisfacer preferencias individuales a través de relaciones estratégicas. Desde esta perspectiva, no hay posibilidad de encontrar consenso sobre objetivos comunes, y mucho menos capacidad de coordinar la acción colectiva para proyectarlos en la práctica. Por ello, el estudio de lo moral debe quedar reducido al análisis cuantitativo de los intereses, preferencias, expectativas y comportamientos individuales de los sujetos en relación. Sin embargo, como demuestra la *teoría de juegos neuronales*<sup>12</sup> la cooperación no sólo es posible, sino también deseada por la mayoría de las personas. Así pues,

12 Para una revisión bibliográfica de los estudios de *teoría de juegos neuronales* sobre la cooperación, ver Calvo (2018).

un planteamiento de estudio sobre lo moral basado en el atomismo social resulta inapropiado por arbitrario, reduccionista y falaz, puesto que da la espalda a los acuerdos contrafácticos que afloran en las sociedades cordialmente maduras y constituyen la base del sentido, normatividad y motivación moral.

En definitiva, como dice Habermas, “El dinero y el poder no pueden ni comprar ni imponer solidaridad y sentido. En una palabra, el resultado del proceso de desengaño [...] es una nueva conciencia, en la que el proyecto que es el Estado social se torna en cierto modo reflexivo, y que se dirige a domesticar no solamente la economía capitalista sino también al propio Estado” (1989, p. 428). Este proceso de desengaño se muestra hoy como un necesario curso comunicativo de clarificación y legitimación de la *dataficación* de la dimensión moral de los distintos ámbitos de actividad humana.

## 5. Conclusiones

El actual proceso de transformación digital y los enormes beneficios estratégicos y pronósticos alcanzados en ámbitos de actividad como el industrial o el *retail* ha generado diferentes propuestas teórico-prácticas para su implementación y desarrollo en ámbitos no económicos como el político, el educativo, el asistencial y, lo que es más preocupante, el moral.

En cuanto a la transformación digital del ámbito moral, su pretensión es concretar y posibilitar en la práctica un enfoque de ética que, sustentado sobre modelos matemáticos capaces de dilucidar imparcial y objetivamente qué es lo justo y correcto mediante la cuantificación y procesamiento de datos masivos relativos a las preferencias, opiniones y hábitos de los individuos digitalmente hiperconectados, permita tomar decisiones racionalmente objetivas —basadas en datos comportamentales cuantificables, comparables y verificables— y moralmente justas —basadas en un criterio de mayorías: el mejor bien para la mayor cantidad de personas hiperconectadas posibles. Para ello, aprovechando la cada vez mayor hiperconectividad de la sociedad, el enfoque propone la recopilación de datos masivos sobre cuestiones morales, así como el diseño e implementación de algoritmos de comportamiento capaces de procesar estos datos para tomar decisiones óptimas en los distintos ámbitos de actividad.

Como se ha criticado, este enfoque puede producir perturbaciones importantes en el mundo de la vida, en ese conjunto de evidencias y convicciones aporoblemáticas a partir de las cuales los interlocutores válidos establecen procesos cooperativos de interpretación sobre algo en el mundo objetivo, social y subjetivo. Especialmente en el mundo social, donde los procesos de diálogo entre afectados tendentes al entendimiento intersubjetivo sobre qué

es lo justo son sustituidos por procesos de recopilación y procesamiento de datos masivos entre algoritmos de IA tendentes a dilucidar el mejor curso de acción posible.

Pero, además, se ha intentado mostrar cómo este enfoque de ética basado en la hiperconectividad, la dataficación y la algoritmización muestra carencias, contradicciones y límites que hacen de ella una propuesta inasumible. Este construye erróneamente la *identidad de significados* a través de la captación de señales lingüísticas; confunde la *identidad de un significado* con la *identidad de un objeto*; se despliega desde la perspectiva de un observador omnisciente y neutro que sigue reglas impuestas por la situación contextual; se desarrolla desde el interés estratégico y predictivo característico del saber científico-técnico; subyuga la reflexión ético-crítica a la elección racional del curso de acción óptimo; abraza un criterio de fundamentación de lo moral basado en el mayor bien para la mayor cantidad de gente posible; y se sustenta metodológicamente en el *atomismo social* como fuente de estudio y análisis. Todo ello muestra las carencias, límites y contradicciones que hacen inviable su aplicación e implementación como propuesta alternativa de ética.

No obstante, la pretensión de este trabajo ha sido mostrar los límites y consecuencias de todo este precario pero influyente proceso de digitalización del ámbito moral desde una postura ético-crítica y dialógica con el principal objetivo de abrir espacios de diálogo interdisciplinares con los diferentes ámbitos implicados y/o afectados para revisar, mejorar, orientar o descartar el proceso.

## Bibliografía

- ANDERSON, CH. "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete", In: Wired, 2008. <https://www.wired.com/2008/06/pb-theory/>. Accessed on Nov. 20, 2018.
- ARROW, K. J. "Social choice and individual values". New York: John Wiley & Sons, 1951.
- BOSTROM, N. "Superintelligence: Paths, Dangers, Strategies". Oxford: Oxford University Press, 2014.
- CALVO, P. "Bioética de las Cosas: sobre la algoritmización de la deliberación moral en la práctica clínica". *Filosofía Unisinos*, Vol. 20, Nr. 2, pp. 5-30, 2019, Mat-Aug.
- \_\_\_\_\_. "Moral neurolearning by machine: artificial values, intelligences and neural networks". In: CALVO, P., GRACIA, J. (ed). *Moral Neuroeducation for a Democratic and Pluralistic Society*. Cham: Springer, 2019a.
- \_\_\_\_\_. "The Cordial Economy – Ethics, Recognition and Reciprocity". Cham: Springer, 2018.

- EUROPEAN COMMISSION. “Internet of Things in 2020, Roadmap for future”. Bruselas: EPoSS, 2008. [https://docbox.etsi.org/erm/Open/CERP%2020080609-10/Internet-of-Things\\_in\\_2020\\_EC-EPoSS\\_Workshop\\_Report\\_2008\\_v1-1.pdf](https://docbox.etsi.org/erm/Open/CERP%2020080609-10/Internet-of-Things_in_2020_EC-EPoSS_Workshop_Report_2008_v1-1.pdf). Accessed on Nov. 20, 2018.
- \_\_\_\_\_. “COM (2009) 278 final. Internet of Things — An action plan for Europe”. Bruselas: European Commission, 2009. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52009DC0278&from=ES>. Accessed on June 20, 2018.
- \_\_\_\_\_. “COM (2009) 512 final. Preparing for our future: Developing a common strategy for key enabling technologies in the EU”. Bruselas: European commission, 2009. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52009DC0512&from=FR>. Accessed on June 20, 2018.
- CORTINA, A. “Ética mínima”. Madrid: Tecnos, 1986.
- \_\_\_\_\_. “Ética sin moral”. Madrid: Tecnos, 1990.
- \_\_\_\_\_. “¿Para qué sirve realmente... la ética?”. Barcelona: Paidós, 2013.
- CORTINA, A., MARTÍNEZ, E. “Ética”. Madrid: Akal, 1996.
- DIJCK, J.V. “Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology”. *Surveillance & Society*, Vol. 12, Nr. 2, 2014, pp. 197-208.
- GARCÍA-MARZÁ, D. “Neuroética aplicada: las consecuencias prácticas del neuropositivismo”. *Pensamiento*, Vol. 72, Nr. 273, 2016, pp. 881-900.
- \_\_\_\_\_. “¿Sentimientos virtuosos? El papel de los sentimientos en la vida moral”. *Diálogo Filosófico*, Nr. 62, 2005, pp. 241-256.
- \_\_\_\_\_. “Ética de la justicia: J. Habermas y la ética discursiva”. Madrid: Tecnos, 1992.
- HABERMAS, J. “Teoría de la acción comunicativa: complementos y estudios previos”. Madrid: Cátedra, 1989.
- \_\_\_\_\_. “Teoría de la acción comunicativa II: Crítica de la razón funcionalista”. Madrid: Taurus, 1987.
- \_\_\_\_\_. “Ciencia y técnica como “ideología””. Madrid: Tecnos, 1984.
- JENTZSCH, S., SCHRAMOWSKI, P., ROTHKOPF, C., KERSTING, K. “The Moral Choice Machine: Semantics Derived Automatically from Language Corpora Contain Human-like Moral Choices”. In: *Proceedings of the 2nd AAAI/ACM Conference on AI, Ethics, and Society*. Palo Alto (California): Association for the Advancement of Artificial Intelligence, 2019. [https://ml-research.github.io/papers/jentzsch2019aies\\_moralChoiceMachine.pdf](https://ml-research.github.io/papers/jentzsch2019aies_moralChoiceMachine.pdf). Accessed on Feb. 9, 2019.
- MALLE, B.F. “Integrating robot ethics and machine morality: the study and design of moral competence in robots”. *Ethics InfTechnol*, Nr. 18, 2016, pp. 243-256.
- MATSUMOTO, T. “The Day AI Becomes God. The Singularity will Save Humanity”. Cambridge (NZ): Media Tectonics, 2018.
- MENGER, C. “Grundsätze der Volkswirtschaftslehre”. Viena: Wilhelm Braumuller, 1871.
- MEDIALAB. “Moral Machine”. Cambridge, MA.: Massachusetts Institute of Technology, 2019. <http://moralmachine.mit.edu>. Accessed on Feb. 3, 2019.
- MOROZOV, E. “The Real Privacy Problem”. *MIT Technology Review*, 2013. <https://www.technologyreview.com/s/520426/the-real-privacy-problem/>. Accessed on Nov. 20, 2018.

MOTTA, L.C., OLIVEIRA, L.N., SILVA, E., SIQUEIRA R. “Toma de decisiones en (bio)ética clínica: enfoques contemporáneos”. *Revista de Bioética*, Vol. 24, Nr. 2, 2016, pp. 304-314.

NATIONAL INTELLIGENCE COUNCIL. “Disruptive Technologies Global Trends 2025. Six Technologies with Potential Impacts on US Interests Out to 2025”. <https://fas.org/irp/nic/disruptive.pdf> (2 de julio de 2018). Accessed on Nov. 20, 2018.

NEIL-CUKIER, K., MAYER-SCHÖENBERGER, V. “Big Data. A Revolution that will transform how we live, work, and think”. Boston: Aemon Dolan Book, 2013.

O’NEAL, C. “Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy”. New York: Broadway Books, 2016.

RASKINO, M. “Strategic Information Management for Competitive Advantage”. In: Gartner, 2011. <https://www.gartner.com/doc/1990916/competitive-advantage-business-transformation-innovation>. Accessed on June 11, 2018.

SÁEZ, C. “Augmented Democracy”. *CCCBLAB. Cultural Research and Innovation*, 2018. Consultado en: <http://lab.cccb.org/en/democracia-aumentada/>. Accessed on June 29, 2018.

WHITERS, P. “Robot take over: Machine to run for Mayor in Japan pledging “fair opportunities for all””. In: *Express. For de Daily and Sunday Express*. 2018. <https://www.express.co.uk/news/world/947448/robots-japan-tokyo-mayor-artificial-intelligence-ai-news>. Accessed on June 20, 2018.

ZAMAGNI, S. “Civil Economy. A New Approach to the Market in the Age of the Fourth Industrial Revolution”. *Recerca. Revista de Pensament i Anàlisi*, Nr. 23, 2018, pp. 1-18.