

## Ética discursiva e inteligencia artificial. ¿Favorece la inteligencia artificial la razón pública?\*

Discourse ethics and artificial intelligence. Does artificial intelligence favor public reason?

JESÚS CONILL SANCHO\*\*

**Resumen.** Este artículo muestra que la ética del discurso en versión de la *ethica cordis* contribuye a mantener una actitud crítica de discernimiento ante las tecnologías de la inteligencia artificial. Propone distinguir entre opinión pública y razón pública, para destacar el sentido crítico del uso público de la razón en la línea de Kant, Rawls, Habermas y Cortina. También se propone afrontar las dificultades para ejercer la razón pública en la era digital: impacto de la inteligencia artificial en la comunicación, hiperconectividad, dataficción, “panóptico digital” y la espiral del silencio, que ponen en peligro el uso de la razón y la intimidad personal.

**Palabras clave:** Ética, Inteligencia, Razón, Crítica, Opinión pública, Dataficción

**Abstract.** This article shows that discourse ethics in version of *ethica cordis* contributes to maintain a critical attitude of discernment towards artificial intelligence technologies. It proposes to distinguish between public opinion and public reason, to highlight the critical sense of public use of reason in the line of Kant, Rawls, Habermas and Cortina. It also aims to address the difficulties in exercising public reason in the digital age: impact of artificial intelligence on communication, hyperconnectivity, datafication, “digital panopticon”, and the spiral of silence, that endanger the use of reason and personal intimacy.

**Keywords:** Ethics, Intelligence, Reason, Criticism, Public Opinion, Datafication

---

Recibido: 24/ 03/2023. Aceptado: 04/06/2023.

\* Este estudio se inserta en el Proyecto de Investigación Científica y Desarrollo “Ética cordial y Democracia ante los retos de la Inteligencia Artificial” PID2019-109078RB-C22 financiado por MCIN/ AEI /10.13039/501100011033 y en las actividades del grupo de investigación de excelencia PROMETEO/2018/121 y del “Programa Prometeo 2022 para grupos de investigación de excelencia, CIPROM/2021/072” de la Generalitat Valenciana. Agradezco en especial las sugerentes intervenciones de César Ortega, Javier Gracia, Juan Carlos Siurana y Pedro Pérez-Zafrilla, así como las de Domingo García-Marzá, Ramón Feenstra y Patrici Calvo.

\*\* Catedrático de Filosofía Moral y Política (Universidad de Valencia). Sus libros más recientes son: *Intimidad corporal y persona humana. De Nietzsche a Ortega y Zubiri* (2019) y *Nietzsche frente a Habermas. Genealogías de la razón* (2021), ambos en la editorial Tecnos. Contacto: Jesus.Conill@uv.es

## 1. Ética discursiva en su versión cordial

La ética discursiva propuesta por Karl Otto Apel (1985; Siurana, 2003) y Jürgen Habermas (1985; García-Marzá, 1992; Ortega, 2021) se inscribe en un proyecto ilustrado de modernidad crítica, en el que el *télos* insito al lenguaje y la comunicación orienta la acción humana. Articulada en dos partes, una de fundamentación y otra de aplicación, la ética discursiva se constituye a través de una transformación hermenéutico-comunicativa de la ética kantiana en una ética de la responsabilidad solidaria, capaz de afrontar las consecuencias planetarias del desarrollo científico-técnico en el medio de la intersubjetividad con unos mínimos de justicia (Cortina, 1985, 1986, 2007 y 2021).

Este proyecto mantiene el sentido eleuteronómico y emancipador de la razón moral moderna, más allá del uso instrumental y estratégico de la razón. La autonomía personal, expresada reflexivamente a través de las pretensiones de validez en los actos de habla, y la capacidad argumentativa en la interacción comunicativa abren un nuevo mundo, en el que pueda escucharse la voz de todos los posibles afectados. Libertad e intersubjetividad conforman la nueva perspectiva en la que nos reconocemos y respetamos mutuamente como personas en la modalidad de interlocutores válidos, pues “somos lo que somos gracias a nuestra relación con otros” (Mead, 1972: 307).

En la versión de la ética discursiva que ofrece la *ethica cordis*, la idea kantiana de persona moral se transforma en la de interlocutor válido, un ser dotado de “competencia comunicativa”, es decir, capaz de defender en la interacción comunicativa la exigencia de atenerse a las pretensiones racio-cordiales de validez argumentativa mediante el diálogo (Cortina 1985 y 2007; Martínez-Navarro, 2022). Se trata del giro pragmático-lingüístico (comunicativo) de la filosofía trascendental, que se presenta en forma de pragmática trascendental y/o universal, en la que se mantiene un criterio de “validez” y de “incondicionalidad”, irreductible a las vigencias que se imponen fácticamente.

Esta “ética discursiva”, también denominada “dialógica” y “comunicativa”, tiene como principio moral el ideal dialógico en el medio de la comunicación. En su fundamentación, pues, no basta la razón instrumental y estratégica, sino que recurre a una nueva figura de la razón práctica, que no es la aristotélica ni tampoco estrictamente la kantiana, aunque tenga en cuenta sus ineludibles aportaciones, sino a una razón práctica comunicativa, cuya “fuerza vinculante” se presupone en el uso del lenguaje y se descubre mediante reflexión reconstructiva, incluso a través de una peculiar genealogía de la razón (Habermas, 2019; Conill, 2021).

La razón comunicativa, como toda forma de razón y de saber, está regida por un peculiar interés. Pero, a diferencia de la razón técnica que está orientada por el interés de dominio, la razón comunicativa busca el entendimiento mutuo y, en la medida en que pone en marcha su dinamismo crítico, se complementa con el interés por la emancipación o liberación humana. Sin tener en cuenta la función de estos intereses no se comprendería el sentido de las diversas formas de la razón en la interacción humana.

La interacción humana de la que parte la ética discursiva es la del *factum* argumentativo, en el que pueden expresarse y explicitarse intersubjetivamente las presupuestas pretensiones de validez (sentido, inteligibilidad, verdad, corrección, veracidad, autenticidad) que al menos están siempre latentes de modo virtual como condición de posibilidad en los juegos lingüísti-

cos de la comunicación humana (Apel, 2017). Estas pretensiones de validez son condiciones universales de la constitución y validez del discurso humano. De lo contrario, no es posible coordinar la acción humana intersubjetivamente y orientarla hacia el entendimiento mutuo, sino que la acción estaría regida únicamente por el interés técnico y estratégico de dominio.

Si la interacción humana sólo se rigiera por la razón instrumental y estratégica, la moral sólo se compondría de imperativos hipotéticos y no contaríamos con ninguna incondicionalidad como base de la exigencia de un posible imperativo categórico. En la versión cordial de la ética discursiva se refuerza el valor de la acción comunicativa, cuya racionalidad tiene primacía sobre la técnica y estratégica, porque en ella cabe descubrir un “momento de *incondicionalidad*”.

La acción comunicativa es aquella en que los actores se coordinan orientados por las pretensiones de validez. Pues, aunque incluso la acción lingüística puede utilizarse de modo instrumental y estratégico, el *télos* inherente al lenguaje humano es el entendimiento mutuo, de tal manera que el modo primordial y “originario” de usar el lenguaje es el que pretende lograr el acuerdo. Por tanto, según Adela Cortina, hay un uso del lenguaje más valioso que otro, hay una jerarquía axiológica entre los usos del lenguaje, porque “el *télos* del lenguaje, inherente a la razón comunicativa, muestra la primacía de ésta frente a la estratégica” y por eso en el orden del discurso y de la auténtica comunicación ha de prevalecer la “fuerza del mejor argumento” en busca de un consenso racional, anticipado contrafácticamente como idea regulativa (Cortina, 1985, 1990 y 2007).

En la comprensión pragmático-lingüística del uso de la razón, en el intercambio discursivo de las razones prácticas, la fuerza sin violencia del mejor argumento sólo puede entrar en juego si los participantes adoptan la “perspectiva del nosotros”, desde la que juzguen “*imparcialmente*”, sin perder de vista la dimensión de la “validez del deber” [*Sollgeltung*], qué importa en interés de todos los posibles afectados (Habermas, 2019: 784). El “nosotros” inclusivo no se restringe a una comunidad particular, sino que se extiende contrafácticamente a todas las personas; la razón práctica en la ética discursiva se orienta por esta idea de justicia comunicativa, adoptando una perspectiva del nosotros que trasciende todos los límites sociales y comunidades locales (Habermas, 2019: 786).

La ética discursiva sigue siendo una ética universalista como la kantiana, pero ahora el principio de universalización incorpora el consecuencialismo de la satisfacción de los intereses de los afectados, reformulando así el imperativo categórico mediante una razón dialógica. La “voluntad racional”, es decir, lo que “todos podrían querer”, como criterio para legitimar las normas morales, se traslada al orden del diálogo entre todos los afectados por la norma, al menos de modo virtual: “todos los seres capaces de comunicación lingüística deben ser reconocidos como personas, puesto que en todas sus acciones y expresiones son interlocutores virtuales” (Apel, 1985: II, 380).

En esta última formulación, el concepto de persona moral se sitúa en el orden dialógico, en la interacción comunicativa constituida por el reconocimiento recíproco entre interlocutores virtuales con competencia comunicativa y, por tanto, capaces de plantear pretensiones de validez y defenderlas discursivamente. Lo decisivo desde la perspectiva de la ética discursiva es mantener la distinción crítica entre “vigencia” (facticidad) y “validez”.

Pero la validez procedimental requiere el presupuesto de alguna valoración fundamental (dignidad humana, acuerdo racional), alguna elección de valor; en último término, es nece-

saría una opción personal, pues “la *realización práctica de la razón* a través de la voluntad (buena) siempre necesita un compromiso que no puede demostrarse” (Apel, 1985: II, 392), pero que puede someterse a una constante consideración crítica como expresión de la libertad racio-cordial.

La ética discursiva, pues, fomentará una “actitud humana”, configuradora de un *êthos* en la vida personal y compartida socialmente, cultivando la voluntad de verdad (el interés por la verdad) y de diálogo (la deliberación) sobre las necesidades e intereses de los afectados. El procedimiento de formación discursiva de la voluntad tiene en cuenta la libertad de las personas y su interrelación intersubjetiva en la vida social, orientada por la comunidad ideal de comunicación. Aquí la universalización adquiere un carácter dialógico a través del reconocimiento recíproco en la praxis comunicativa, en una argumentación polifónica donde se vive la tensión entre la incondicionalidad de las pretensiones de validez y la facticidad de los contextos concretos. Por eso, prosiguiendo el espíritu ilustrado la ética discursiva aboga por una razón comunicativa en forma de “publicidad razonante” en la vida social y sus instituciones (incluso de la democracia mediante la formación deliberativa de la voluntad) (Habermas, 2019: II, 765-766).

Se plantea entonces la cuestión de los procesos de aprendizaje correspondientes a la moral, pues en principio somos capaces de aprender también con respecto a la dimensión práctico-moral, de modo especial a través de la función educativa de los procesos comunicativos (Gozálvez, 2012; Gracia, 2018). Pues igual que en la dimensión biológica tenemos un triple cerebro a partir del más primitivo, que no podemos erradicar, también (¡o más!) tenemos en el orden sociocultural formas tribales de pensamiento que nos es imposible erradicar y difícil superar. No obstante, según Habermas, es constatable que la razón práctica ha dejado también huellas de procesos de aprendizaje y de progreso moral; que la razón comunicativa tiene un potencial cognitivo, es decir, que se puede aprender en el orden moral a través de la adopción recíproca de las perspectivas de los otros en busca del interés universalizable, al menos reconociendo al otro como persona igualmente libre (Habermas, 2019: II, 789 y ss.). Se presenta aquí la autocomprensión de la libertad racional del hombre primordialmente en virtud de su capacidad sociocognitiva de comunicación, porque “nadie puede ser autónomo [libre] por sí solo” (Habermas, 2019: II, 806).

## **2. ¿Razón pública en vez de opinión pública?**

La ética de la razón cordial se basa en la ética discursiva, pero renovada y hasta transformada a través de la integración de nuevos componentes como las emociones y los sentimientos, los valores, las virtudes, la voluntad racio-cordial, la deliberación, la inteligencia sentiente (reforzada por las neurociencias) y el poder de la libertad en virtud de la intimidad corporal en el medio de la vida social (Conill, 2019). Uno de los escenarios más relevantes de esta nueva versión de la ética discursiva es la compleja intersubjetividad comunicativa (con todos sus ingredientes) a través del ejercicio de la razón pública. Así es como los ciudadanos ilustrados y razonables ponen de manifiesto su pluralismo y conviven con discrepancias, pero cooperando unidos por una razón cordial.

En el contexto moderno lo que llamamos “razón pública” quedó bien determinado en principio por la noción kantiana del “uso público de la razón” (Kant, 1968, 1983, 1985, 1981), que ha sido actualizado de modo significativo por Rawls (1971/1972, 1978, 1993, 1996, 2001, 1999: 573-611), Habermas (1981, 1987, 2019, 2022) e incluso por Sen (2010: caps. 15-17). La razón pública, que es expresión de la libertad efectiva y de la capacidad crítica, prosigue ahora el impulso ilustrado en su versión comunicativa.

Ahora bien, la razón pública no debería confundirse con la opinión pública, que tiene un carácter ambiguo. Por una parte, ésta constituye una presión social, ahora tecnológicamente reforzada; pero también lleva incorporado un potencial sentido crítico. Ya en la filosofía clásica antigua, Aristóteles distingue entre opinión y elección; ésta va conformando el *êthos*, no así la opinión. No bastan el apetito, el impulso, el deseo y la opinión, sino que hace falta cultivar el *orthós lógos*, la *recta ratio*, para orientar la acción como es debido —la buena *prâxis*— y configurar una vida virtuosa, potencialmente feliz. En los términos de una filosofía crítica (moderna y contemporánea), la voz de la razón comunicativa ha de conformar la “voluntad general” —la voluntad común—, al llevar ínsita —incoada— la pretensión de universalidad frente a los dogmatismos. No se reduce a la mera opinión, ni a la suma de opiniones, sino que constituye una vía crítica al estar orientada por una comunidad ideal de comunicación. Se instaura así el tribunal de la razón, que tiene vigor en la esfera pública universal, mostrando la efectividad de la razón en la historia, pues lo racional tiene que hacerse efectivo. Kant y Hegel, “*Moralität*” y “*Sittlichkeit*” han de articularse en la vida social mediante la razón pública<sup>1</sup>.

Por su parte, la versión rawlsiana de la idea de razón pública considera su ejercicio como un deber de civilidad, pero se restringe a lo política y jurídicamente razonable (Conill, 2022). En cambio, la versión habermasiana amplía el sentido de la razón pública al conjunto de la sociedad civil, a todos los ámbitos de la cultura, como espacios de comunicación social abiertos a la deliberación pública. La visión ampliada de razón pública incorpora a todos los ciudadanos y se entiende como razonabilidad abierta, más allá de lo estrictamente político y jurídico, a los contenidos de las diversas concepciones del mundo en una sociedad postsecular (Habermas, 1981, 2009, 2019; Conill, 2021: caps. 1-3).

Habermas confía avanzar en la racionalización de la vida social y política mediante la aplicación del principio de publicidad (Habermas, 2009: 15)<sup>2</sup>. Lo explica ya en *Strukturwandel der Öffentlichkeit* y así lo reafirma posteriormente desarrollando un concepto normativo de publicidad y de política deliberativa (Habermas, 2022). No obstante, a pesar de esta aportación de un concepto normativo de publicidad, su uso ampliado comporta una cierta ambigüedad, ya que los promotores de la publicidad son las libres asociaciones que forman la red de comunicación a partir del entrelazamiento de las diversas manifestaciones de la opinión pública (*Öffentlichkeiten*). Tales asociaciones autónomas están especializadas en la creación y difusión de las convicciones prácticas, por tanto, en descubrir temas de relevancia social general, en contribuir a solucionar problemas, en interpretar valores y en producir

1 Agradezco a Juan Carlos Siurana sus intervenciones sobre la necesidad de aclarar este punto, que concilia las aportaciones de Kant y Hegel en la ética discursiva a través de la intersubjetividad.

2 Habermas utiliza en estos contextos indistintamente los términos “*Publizität*” y “*Öffentlichkeit*”.

buenas razones o desvalorizar otras (Habermas, 2009: 62-63); son las instituciones de la libertad crítica.

Precisamente el modelo deliberativo de democracia explicará la fuerza legitimadora del proceder democrático con la ayuda del carácter racional de la formación de la opinión y de la voluntad (Habermas, 2009: 87). Este modelo parece ser un ejemplo especialmente convincente y eficaz para afrontar el profundo foso entre lo normativo y lo empírico en la ciencia política, intentando responder al problema de cómo ajustar —hacer compatible— un concepto normativo como el de “política deliberativa” con nuestra imagen realista de la sociedad mediática (Habermas, 2009: 87-88). Pues uno de los elementos del marco institucional de las democracias modernas y que constituyen el núcleo normativo de los Estados de derecho democráticos es el de una opinión pública [Öffentlichkeit] independiente que en tanto que esfera de libre formación de la opinión y de la voluntad une el Estado y la Sociedad Civil (Habermas, 2009: 89).

Pero Habermas, aunque es consciente de las crecientes dificultades que está generando la revolución de los nuevos medios de comunicación, no renuncia a su propósito de que el paradigma deliberativo no fracase en el intento de conectar sus “fuertes ideas normativas” con la actual complejidad social (Habermas, 2009: 92). El modelo deliberativo espera que su apuesta por la racionalización mejore la calidad de las decisiones, dirigiendo su mirada a las funciones cognitivas de la formación de la opinión y de la voluntad, así como a la busca cooperativa de la solución de problemas (Habermas, 2009: 95). Pretende así aprovechar el potencial racional de la deliberación y del discurso en el contexto de la comunicación de masas<sup>3</sup>. Lo que ocurre es que las meras opiniones están sometidas a las transacciones de la oferta y la demanda, pero no están regidas por la meta de encontrar soluciones legítimas a los problemas en litigio (Habermas, 2009: 109). Pues los participantes en la presunta comunicación de masas son espectadores y consumidores, que no hacen uso de la auténtica “razón pública”, que siempre ha de llevar incorporado un sentido crítico y normativo (Marina, 2023).

En realidad, hace falta una publicidad crítica, no sólo nacional, sino europea y global o mundial. Un camino apropiado sería transnacionalizar las publicidades —opiniones públicas— existentes y hacerlas reflexivas, a fin de que la opinión pública masificada no quede sometida a la manipulación y la indoctrinación, sino que sirva a la ilustración (Habermas, 2009: 112 y ss., 137-138, 346).

### 3. El sentido crítico del uso público de la razón

El uso público de la razón es expresión del único derecho innato, que es la libertad; no es sólo una aplicación de las exigencias racionales al mundo político-moral, sino el núcleo mismo de la filosofía crítica, porque la crítica consiste en la posibilidad de que los ciudadanos libres presenten sus objeciones (Cortina, 2021; Conill, 2022). Con ello, el uso público-crítico de la razón se convierte en un presupuesto de la argumentación misma. Ya al comienzo de la *Crítica de la Razón pura* destaca Kant la dimensión comunicativa de la

3 Me parece que es imposible una “comunicación de masas”, porque las masas no se comunican, sino que las relaciones entre ellas son de otro género.

autocomprensión humana, en la *Crítica del Juicio* reconoce el carácter nuclear de la comunicabilidad y, en el II Apéndice de *La paz perpetua*, el *Principio de Publicidad* adquiere carácter jurídico-político (Kant, 1968, 1983, 1985, 1981, García-Marzá, 2012).

El ámbito de los ciudadanos que pueden presentar sus objeciones a las propuestas y argumentaciones de la razón común humana no se circunscribe a una parte de la humanidad, la que comparte las mismas bases culturales, como si fueran los únicos capaces de comprender tales propuestas y argumentaciones, sino que el derecho a presentar objeciones es un derecho de cualquier ser humano, de modo que no se puede impedir a nadie que lo ejerza, ya que es un derecho de la humanidad, es decir, un derecho de rango cosmopolita (Cortina, 2021). Así lo confirma con claridad el siguiente texto de la *Crítica de la razón pura*: “También forma parte de esta libertad el exponer a pública consideración los propios pensamientos y las dudas que no es capaz de resolver uno mismo... [...]. Esto entra ya en el derecho originario de la razón humana, la cual no reconoce más juez que la misma razón humana común, donde todos tienen voz” (Kant, 1983: A 751-752/B 779-780).

La referencia a la metáfora del tribunal de la razón y a la esfera de la opinión pública universal apunta a una estrecha relación entre la crítica de la razón y el cosmopolitismo (Bösch, 2007: 480). En esta línea interpretativa, Adela Cortina precisa que, al haber un derecho originario de la razón humana que no reconoce más juez que la misma razón, donde todos tienen voz, es necesaria una república mundial como condición de posibilidad del uso crítico de la razón y de la superación del dogmatismo.

El propósito de Kant con los textos citados de la “Doctrina Trascendental del Método” de la *Crítica de la razón pura* es profundizar en el proceso de Ilustración (*Aufklärung*), que expone claramente en *Beantwortung der Frage: Was ist Aufklärung?* (Kant, 1968: VIII, 33-42) y que exige fomentar un uso público de la razón. De ahí que la filosofía crítica kantiana se haya interpretado como una *paideia* (Munzel, 2012: XXI; Cortina 2021), con el propósito de cultivar la libertad interna. No obstante, todo ello supone procesos sociales de aprendizaje, como muestra Habermas reiteradamente en *Auch eine Geschichte der Philosophie*.

La razón crítica necesita del uso público de la razón, que todos los hombres se consideren “partícipes potenciales de una república mundial” (Andaluz, 2018: 438; Cortina, 2021). Una sociedad de ciudadanos del mundo es un requisito indispensable para llevar adelante el proceso de ilustración de la humanidad. Para que la razón crítica sea viable se requiere tomar en serio el mandato de la razón de avanzar hacia una sociedad cosmopolita (Cortina, 2021). La razón crítica reclama una esfera pública donde todos puedan tener voz, y la perspectiva del participante exige que el horizonte político de los ciudadanos se amplíe “para posibilitar una formación política común de la voluntad por encima de las fronteras nacionales y una acción política común en el nivel transnacional” (Habermas, 2019: II, 800).

### **¿Puede ejercerse la razón pública en el medio tecnológico de la IA?**

Hemos visto que la ética discursiva surge precisamente desde la experiencia comunicativa y confía en la significación de una racionalidad comunicativa en el espacio público.

Pero ¿favorece la IA el desarrollo de la razón comunicativa y la razón pública, o constituye un nuevo obstáculo e incluso una peligrosa amenaza?

La ética discursiva, también denominada “comunicativa” o “del discurso”, se desarrolla a partir del medio de la comunicación y presupone una noción de razón comunicativa operante en el espacio público, a diferencia de la razón técnica, instrumental y estratégica. En ese sentido, una de las pretensiones de la ética discursiva es contribuir a formar la propia conciencia y la voluntad, para que los ciudadanos lleguen a ser capaces de debatir y dialogar personalmente y en público.

No obstante, es palmaria la relevancia mundial del impacto de la IA en todos los órdenes de la vida. Nuestro mundo vital está siendo alterado —colonizado— por las tecnologías de la IA (en los ámbitos de la salud, el trabajo, la educación, la economía, las finanzas, la Administración), y muy especialmente se está transformando la interacción comunicativa. Pero en esta nueva realidad social los imperantes medios tecnológicos están reprimiendo la libertad que quería expresarse a través de la opinión pública y no permiten, o al menos dificultan gravemente, el ejercicio de la razón pública.

Los cambios sociales y culturales han conducido a una situación en que los ciudadanos se encuentran atrapados en entornos digitales y mediáticos que distorsionan la realidad y están dominados por los algoritmos diseñados para las redes sociales, creando lo que se han llamado “cámaras de eco” (que reducen el pluralismo a la insistente repetición de lo mismo), de modo que cada cual sólo escucha lo que ya comparte, impidiendo la comunidad de diálogo entre posiciones diversas. Cuando lo que se necesita es promocionar una ciudadanía madura e ilustrada también a través de los medios digitales y mediáticos, fomentando diálogos auténticamente argumentativos. Sólo si se logra una ciudadanía digital y mediática se podrá avanzar en la mejora de la calidad de la democracia, en la medida en que el universalismo moral y la justicia cordial inspiren la educación en la responsabilidad que exigen las nuevas condiciones sociales y los nuevos medios tecnológicos que han invadido la sociedad y el ejercicio de las profesiones.

### **La hiperconectividad digital no favorece el diálogo y la reflexión**

Una interpretación de la era digital es la que considera que en ella emerge una nueva figura del ser humano como *homo poieticus* (Floridi/Sanders, 2003) en el seno de la cuarta revolución tecnológica de carácter digital. Pero también hay quienes han insistido (Foucault, 1975) en que la era digital hace peligrar la libertad por las sociedades modernas, debido al creciente control por medio de un cierto “panóptico digital”. Todo se convierte en información y peligran la interioridad personal. La nueva tecnología digital posibilita una mayor dominación mediante el control del comportamiento cada vez más mecanizado (vaciado de significación vital propia) y, por tanto, más fácil de cuantificar y hasta de predecir (aprovechando el negocio de los *Big Data*).

Nos habríamos convertido en “organismos informacionales” (*inforgs*), que comparten el entorno informacional o “infoesfera” con otros seres que son artefactos inteligentes. Habría que reinterpretar la posición del hombre en la nueva realidad, en la que todo ha de estar funcionalizado y, por tanto, peligran la dimensión experiencial. Lo decisivo es

que las condiciones digitales no asfixien la experiencia del significado de la vida. Pues lo que está en juego es la forma de vida, si ésta se reduce primordialmente a producción y consumo de datos. Pues la coerción de la inmediatez y la brevedad de los mensajes está impidiendo la comunicación serena y la experiencia reflexiva, la disposición para meditar o pensar con serenidad y calma. Se necesitan espacios y tiempos de concentración, para encontrarse consigo mismo y orientarse en la vida, no dejarse invadir ni alterar constantemente por la hiperconexión digital. Pues se puede vivir hiperconectado, pero descentrado y desorientado. La calidad de vida no se mide por los instrumentos digitales que alteran continuamente la vida.

En el reino de los *inforgs* y del ciberespacio se suele confundir la comunicación con la mera información y la experiencia significativa con los meros datos; pero la comunicación no se identifica con la información. La sobrecarga de información impide pensar, reflexionar con calma y sosiego, formarse la conciencia moral, porque falta la necesaria concentración por exceso de alteración en el trepidante ritmo de la vida. La hiperconexión descentra y produce desorientación vital. Vivimos con más medios que nunca (Ortega y Gasset, 2005), con cada vez más sofisticados instrumentos digitales, pero sintiendo un deterioro de la auténtica calidad de vida, que no consiste en más bienestar, y sin haber logrado ser más y mejor humanos. ¿Se está produciendo realmente una mejor humanización universal de la vida compartida? ¿O está creciendo la desigualdad entre las personas por una nueva brecha tecnológica? (O’Neil, 2018).

### **Datificación de la esfera pública**

La denominada “dataficación” de la esfera pública no está siendo el mejor camino para ejercer una razón pública que incorpora la crítica humanizadora. En la nueva sociedad de la información algunos están “fascinados” por el creciente poder de las tecnologías de la IA en virtud de la hiperconectividad y los medios de codificación algorítmica (Calvo, 2019a). Pero la posibilidad efectiva de mantenerse permanentemente conectados con todo produce un flujo torrencial de datos, que constituye el nuevo elemento en el que se desarrolla la nueva forma de comunicación entendida como interacción (intercambio) de datos. Como bien se pregunta Patrici Calvo en relación con las democracias actuales, ¿se solucionan así mejor (con más imparcialidad, honestidad, integridad, comprensión y justicia) los problemas a los que nos enfrentamos en la vida real?

En su respuesta, Calvo señala varias deficiencias relevantes, de entre las que entresacamos las siguientes: 1) el carácter problemático de la presunta “objetividad” algorítmica, que se nutre del flujo masivo de datos proporcionados por una sociedad hiperconectada, pero que no puede ofrecer la determinación del “bien común” mediante la regla de la mayoría, ni garantiza la transparencia ni la fiabilidad, sino que introduce una nueva opacidad; 2) la imposible “neutralidad”, debido a que hasta los modelos matemáticos incorporan sesgos (prejuicios) de diverso género, que reproducen los de cada sociedad, dependen de ciertas ideologías y de los diseñadores o programadores; 3) la nueva exclusión social, provocada por el silenciamiento de las personas más desfavorecidas de la sociedad o de los grupos que

carecen del poder suficiente para lograr relevancia pública (Cortina, 2022a, 2022b), porque están sometidos a alguna forma de aporofobia (Cortina, 2017).

Este nuevo mundo, de la interacción comunicativa ha generado un nuevo tipo de economía digital y de ciencia social computacional, a partir de una concepción renovada de la naturaleza humana en los términos siguientes: no somos individuos racionales, sino que “somos producto de nuestras redes sociales” (Moreno, 2021/2), seres influenciables mediante incentivos y presión social y, por tanto, predecibles y controlables.

Por otra parte, mediante la datificación se impone una concepción utilitarista de la vida y de la sociedad, reforzada por los nuevos instrumentos tecnológicos y el tipo de conocimiento que ofrecen, orientado primordialmente por el cálculo. De hecho, los procesos de “datatificación” o “datificación” van invadiendo incluso el ámbito moral produciendo una “etificación” (Calvo, 2019b). Se relega o anula el diálogo y la reflexión crítica regida por las pretensiones de validez, en favor de la agregación de las opiniones, preferencias y hábitos. Por tanto, se desconsidera o elimina la razón comunicativa en favor del cálculo y la matematización de los datos, que sirve para predecir el comportamiento y alimentar la razón instrumental y estratégica. Se elimina así la crítica recíproca entre los que participan en la esfera pública, porque en el fondo se presupone una concepción de la sociedad orientada por un individualismo cuantitativista.

### **Enfoque datificado de la ética versus ética discursiva**

Por consiguiente, el enfoque datificado de la ética constituye un reto actual para la ética discursiva, en la medida que pretende ofrecer una mejor respuesta a qué es lo justo y lo bueno a través del análisis y procesamiento algorítmico de los datos masivos sobre lo moral procedentes de los individuos hiperconectados (Calvo, 2019a, 2019b).

Según Calvo, este nuevo enfoque posibilitado por el uso de las tecnologías de la inteligencia artificial constituye un nuevo intento de colonización sistémica del mundo de la vida, que distorsiona el saber moral e impide el entendimiento mutuo, porque más bien sirve para fomentar la polarización social, como ha destacado Pedro Pérez-Zafrilla (Pérez-Zafrilla, 2021a). Por tanto, desde la ética discursiva de Apel, Habermas, Cortina y García-Marzá, Patrici Calvo se propone criticar y superar esta datificación de lo moral, que en realidad conduce a una nueva versión del utilitarismo; pues la etificación consiste en recopilar, cuantificar, procesar y gestionar datos masivos sobre opiniones, preferencias y conductas de la ciudadanía hiperconectada para, desde el criterio del mayor bien para el mayor número, establecer qué es lo moralmente válido. En cambio, desde la ética discursiva hay que restablecer el sentido del diálogo de los afectados para la consecución de un acuerdo racional, la reflexión crítica ateniéndose a las pretensiones de validez en el medio de la interacción comunicativa. La ética no puede convertirse en una ciencia predictiva de la conducta y estar a merced de la razón instrumental y estratégica. Y la ética discursiva no debe difuminarse en función de la datificación de lo moral mediante los algoritmos, pues supondría un regreso al nivel convencional de lo moral, en el que el peso de las vigencias sociales hace desaparecer el diálogo reflexivo y la deliberación crítica.

## El poder represor de la opinión pública

Una de las concepciones más lúcidas de la opinión pública es la que considera que lo que en ella se expresa son las opiniones y conductas que pueden mostrarse en público sin miedo al aislamiento, porque lo que más tememos es el aislamiento, dado que necesitamos a los otros para vivir a gusto, siendo acogidos por los demás. Como decía Tocqueville, “la gente teme al aislamiento más que al error” (Noelle-Neumann, 1982: 25). Todo lo contrario de lo que hemos visto que proponía Kant en relación con el sentido crítico del “uso público de la razón”: “cada uno de los [ciudadanos libres] tiene que poder exponer sin temor sus objeciones e incluso su veto” (Kant, 1983: A 738-739 B 766-767).

Esta teoría de la espiral del silencio muestra el enorme poder de la autocensura (Cortina, 2022a, 2022b). A diferencia de los procedimientos brutos y agresivos de la represión manifiesta y violenta, el mecanismo más sutil y eficaz para silenciar determinadas propuestas en la vida pública, que tiene su raíz y está entañado en la naturaleza de nuestro ser social, funciona a través de esa compleja realidad que es la *opinión pública* (Pérez-Zafrilla, 2021b).

“Hoy se puede demostrar que, aunque la gente vea claramente que algo no es correcto, se mantendrá callada si la opinión pública (opiniones y conductas que pueden mostrarse en público sin temor al aislamiento) y, por ello, el consenso sobre lo que constituye el buen gusto y la opinión moralmente correcta, se manifiesta en contra” (Noelle-Neumann, 1982: 14). El paréntesis que aclara qué sea la opinión pública es sumamente expresivo: la constituyen las opiniones y conductas que pueden mostrarse en público sin temor al aislamiento.

Desde la noción expuesta de “razón pública” hay que criticar el silencio a que se ven sometidos los disidentes y no claudicar a la presión social de la opinión pública. Porque la tendencia que sentimos a imitar y asimilarnos a los demás no proviene de la pretensión de verdad, sino de otro motivo más fuerte: evitar el aislamiento. Una tendencia tan poderosa, que ha hecho que se comprenda su vigencia social: “Quizá no simpatizamos con la naturaleza social del hombre, pero tenemos que intentar comprenderlo para no ser injustos con la gente que se mueve con la multitud” (Noelle-Neumann, 1982: 14).

Según Elisabeth Noelle-Neumann, la espiral del silencio es un proceso en que las observaciones realizadas en unos u otros contextos incitan a unas gentes a expresar sus opiniones y a otras, a tragárselas, a mantenerse en silencio, hasta que en un proceso en espiral un punto de vista domina la vida pública (Noelle-Neumann, 1982: 22). Pero no domina la vida pública ese punto de vista porque sea el más verdadero, sino que triunfa porque en todas las sociedades, también las oficialmente democráticas y en apariencia tolerantes, funciona la autocensura de aquellas opiniones que no van a ser bien acogidas. Por supuesto, en las totalitarias la autocensura va de suyo, excepto en el caso de disidentes valerosos, que suelen pagar su osadía, pero en todas las sociedades funciona la espiral del silencio, lo cual constituye un grave obstáculo para el auténtico pluralismo, la deliberación y la democracia.

Se podría decir que de igual modo que las democracias en los últimos tiempos no mueren tanto por aparatosos golpes de estado, sino por el paulatino deterioro de las instituciones y porque pierden fuerza unas reglas “morales” que la comunidad aceptaba y respetaba (Levitsky y Ziblatt, 2018; Cortina, 2021: 46), asimismo hay ideas valiosas que desaparecen

de la vida social no porque dejen de ser convincentes tras un debate abierto, sino porque las silencian quienes temen al aislamiento más que al error.

Y actualmente habría que añadir: porque temen al aislamiento y al linchamiento público más que a la mentira. En tiempos de la llamada “posverdad” este riesgo es mayor si cabe que antes, porque la posverdad puede caracterizarse como “una mentira emotiva” (Nicolás, 2019). Según Wikipedia, se trata de un “neologismo que describe la distorsión deliberada de una realidad con el fin de crear y modelar la opinión pública e influir en las actitudes sociales, en la que los hechos objetivos tienen menos influencia que las apelaciones a las emociones y a las creencias personales”.

Las noticias falsas tienen mucho más impacto que las verdaderas, pues al generar más interacción atraen más la atención. Y como los algoritmos registran y valoran la interacción de los seguidores, produciendo entre otros fenómenos el dominio de los influyentes o influenciadores (Siurana, 2021), se pone en peligro la auténtica comunidad de comunicación.

El poder de los medios de comunicación llega actualmente a sustituir la experiencia directa y vivida personalmente, de tal manera que se produce un fenómeno muy llamativo, que ya advirtió en su momento Maquiavelo, aunque ahora haya que aplicarlo a la nueva situación en la que vivimos actualmente por la influencia inmediata de las tecnologías y redes de comunicación masiva: “los hombres, en general, juzgan más por los ojos que por las manos, que a todos es dado ver, pero tocar a pocos. Todos ven lo que parecen, pero pocos palpan lo que eres y esos pocos no se atreven a oponerse a la opinión de la mayoría” (Maquiavelo, 1985: 140 y 141; 1987: 103).

El comportamiento de la mayoría de la gente en su medio social se rige por las opiniones que se percibe que van ganando terreno y se convierten en dominantes. De tal modo que los que confían en la victoria se expresan en público, pero los perdedores tienden a callarse (Noelle-Neumann, 1982: 27, 40 y 44). Y la cuestión se ha agravado por el funcionamiento de las redes sociales, que transmiten rápidamente (“viralizan”) las “sentencias” sobre lo que es aceptable por la opinión pública a través de nuevos mecanismos inquisitoriales como el movimiento *Woke* (un pensamiento rigorista sobre lo que es lícito pensar) o el castigo llamado “cancelación”, que consiste en atacar a determinadas personas con el objeto de destruir su reputación y de provocar su muerte social. Aquí ya no se teme sólo al aislamiento, sino hasta incluso a la pérdida del trabajo profesional y de los propios medios de vida. Sigue siendo verdad, como decía Nietzsche, que “nos las arreglamos mejor con nuestra mala conciencia que con nuestra mala reputación” (Nietzsche, 1986: § 52; Conill, 2016: 806-807). Se está imponiendo por estas nuevas vías el miedo a la mala reputación y a la pérdida de estatus en la vida social. Cada vez dependemos más del beneplácito de los demás, expresado a través de unos medios de comunicación de masas que se convierten en potentes instrumentos de control social y que merman la libertad de las personas.

### **¿Sobrevivirá la razón pública al poder de la IA?**

Igual que Nathaniel Persily preguntaba si la democracia puede sobrevivir a internet, tenemos que preguntarnos si lo que ha significado la razón pública podrá sobrevivir al poder que emerge del uso de la inteligencia artificial.

La nueva situación va en contra de la tradición de la opinión pública que surge desde el siglo XVIII y que confiaba en que la humanidad había iniciado un proceso de ilustración, en virtud del cual las personas pueden y deben atreverse a servirse de la propia razón: “¡Atrévete a servirte de tu propia razón!” (Kant). Pues de la libertad personal, según Kant, “forma parte [...] el exponer a pública consideración los propios pensamientos” y no reconocer “más juez que la misma razón humana común, donde todos tienen voz” (Kant, 1983: A 751-752 B 779-780).

Existen dos concepciones y dos prácticas de la opinión pública que están funcionando en la esfera pública de las sociedades modernas y contemporáneas. Una es la concepción psico-sociológica: “Lo que importa no es la calidad de los argumentos, sino cuál de los [...] bandos tiene la fuerza suficiente como para amenazar al contrario con el aislamiento, el rechazo y el ostracismo” (Noelle-Neumann, 1982: 288). Según Noelle-Neumann, es la presión social la que tiene realmente fuerza para cambiar los puntos de vista y funciona como control social, porque afecta a todos. Según esta concepción de la opinión pública, incluso los que conocen la realidad no se atreven a contradecir la opinión mayoritaria. La vigencia social se convierte en norma, en tribunal evaluador y juzgador de las opiniones y las conductas, en la medida en que ejerce un poder de control y represión.

En cambio, la concepción considerada “normativa” cree haber logrado una instancia crítica que no se somete a las vigencias sociales y tribales, ni a los mecanismos tecnológicos que las refuerzan, sino que confía en el poder específico de la razón crítica y comunicativa, expresada a través de las pretensiones de validez en un espacio de razones públicas. Esta línea, que se nutre de la herencia kantiana es la que desarrolla la ética discursiva en sus diversas versiones, tanto por parte de Apel, Habermas, Cortina y García-Marzá, como por algunos representantes de las nuevas generaciones de la teoría crítica de la Escuela de Frankfurt. Por ejemplo, atendiendo a la sugerencia de César Ortega en una sesión del grupo de investigación, podría aprovecharse el “principio de justificación” propuesto por Reiner Forst (Ortega, 2021: 377-384) para reforzar el espacio potencialmente crítico de las “razones” frente a las relaciones de dominación, que siguen impidiendo el derecho básico a la justificación, es decir, a ser tenido en cuenta en cada ámbito de la vida social ofreciendo y recibiendo razones. Forst insta esta “exigencia incondicional” de que cada persona sea respetada como alguien que “merece razones” como un principio de reciprocidad universal, que en el tema que nos ocupa podría aplicarse al ámbito del ciberespacio comunicativo creado por las tecnologías de la inteligencia artificial.

La necesidad de una ética en el mundo digital ha conducido ya a formular unos principios éticos, como en la bioética (no maleficencia, beneficencia, autonomía y justicia), pero a los que se añade un principio decisivo en el nuevo contexto, como es el de explicabilidad o de trazabilidad, en el que han insistido especialmente Adela Cortina (2019) y Domingo García-Marzá (García-Marzá/Calvo, 2022)<sup>4</sup>. Según este decisivo principio, los afectados

4 Intervenciones en el Curso virtual *Ética de la ciencia: transparencia y explicabilidad*, Universidad Jaume I de Castellón, Escuela de Doctorado, Facultad de Ciencias de la Salud, 28 de septiembre-1 de octubre de 2021 y en el Coloquio Internacional sobre “Ética discursiva e IA” (*XV Coloquio Latinoamericano sobre Ética del Discurso y IX Coloquio de la Red Internacional de Ética del Discurso*, organizado por la Fundación ICALA, Río Cuarto, Argentina, 11-12 de noviembre de 2021) y en el Congreso de Filosofía en la UJI, 6-9 de Abril de 2022.

tienen que poder controlar el uso de sus datos y conocer los algoritmos que los manejan, pues también los sistemas de IA operan con sesgos, que incluso son más invisibles (O'Neil, 2018), porque vienen a ser cajas negras. Por eso, los afectados por el mundo digital tienen que poder comprender los algoritmos que manejan sus datos, conocer la trazabilidad, quién los construye y con qué criterios y objetivos, es decir, cumpliendo las exigencias de la ética discursiva, aplicadas al poder de estas nuevas tecnologías de la IA.

## Referencias

- Andaluz, A. (2018), “La hospitalidad en el cosmopolitismo kantiano”, en García-Marzá, D., Lozano, J.F., Martínez Navarro, E. y Siurana, J.C. (comps.), *Ética y filosofía política*, Madrid: Tecnos, 435-445.
- Apel, K.-O. (1985), *Transformación de la filosofía*, Madrid: Taurus, 2 vols.
- Apel, K.-O. (2017), *Transzendente Reflexion und Geschichte*, herausgegeben und mit einem Nachwort von Smail Rapic, Berlin: Suhrkamp.
- Bösch, M. (2007). Globale Vernunft. Zum Kosmopolitismus der Kantschen Vernunftkritik. *Kant-Studien*, vol. 98, n° 4, 473-486.
- Calvo, P. (2019a), Democracia algorítmica: consideraciones éticas sobre la datafización de la esfera pública. *Revista del CLAD Reforma y Democracia*, n° 74, 5-30.
- Calvo, P. (2019b), Etificación, la transformación digital de lo moral. *KRITERION*, Belo Horizonte, n° 144, Dez., 671-688.
- Conill, J. (2016), La intimidad corporal y sus bases neurobiológicas. *Pensamiento*, 273, 789-807.
- Conill, J. (2019), *Intimidación corporal y persona humana*, Madrid: Tecnos.
- Conill, J. (2021), *Nietzsche frente a Habermas. Genealogías de la razón*, Madrid: Tecnos.
- Conill, J. (2022), “Razón pública e inteligencia artificial”, en Pereira, G. y Pérez-Zafrilla, P. (eds.), *Actualidad de John Rawls en el siglo XXI*, Granada: Comares, 37-47.
- Cortina, A. (1985), *Razón comunicativa y responsabilidad solidaria*, Salamanca: Sígueme.
- Cortina, A. (1986), *Ética mínima*, Madrid: Tecnos; 2020 [18ª edición, corregida y aumentada].
- Cortina, A. (1990), *Ética sin moral*, Madrid: Tecnos.
- Cortina, A. (2007), *Ética de la razón cordial*, Oviedo: Nobel.
- Cortina, A. (2017), *Aporofobia*, Barcelona: Paidós.
- Cortina, A. (2019). Ética de la inteligencia artificial, *Anales de la Real Academia de Ciencias Morales y Políticas*, 96, 379-394. Sesión del 7 de mayo de 2019.
- Cortina, A. (2021), *Ética cosmopolita*, Barcelona: Paidós.
- Cortina, A. (2022a), La espiral del silencio y la presunta moralización de la sociedad. *Anales de la Real Academia de Ciencias Morales y Políticas*, 99, 419-430. Sesión del 31 de mayo de 2022.
- Cortina, A. (2022b), Autocensura: destruyendo la democracia. *El País*, 8 de junio de 2022.
- Floridi, L./Sanders, J.J. (2003). Internet Ethics: The Constructionist Values of Homo Poieticus. In R. Cavalier (ed.), *The Impact of the Internet on Our Moral Lives*. New York, SUNY, 195-214.

- Foucault, M. (1975), *Vigilar y castigar*, México: Siglo XXI.
- García-Marza, D. (1992), *Ética de la justicia*, Madrid: Tecnos.
- García-Marzá, D. (2012), Kant's Principle of Publicity, *Kant-Studien*, 103 (1), 96-113.
- García-Marzá, D. / Calvo, P. (2022), Democracia algorítmica: ¿un nuevo cambio estructural de la opinión pública?, *Isegoría*, 67, <https://doi.org/10.3989/isegoria.2022.67.17>
- Gozálvez, Vicent (2012), *Ciudadanía mediática. Una mirada educativa*, Madrid: Dykinson.
- Gracia, J. (2018), *El desafío ético de la educación*, Madrid: Dykinson.
- Habermas, J. (1981). *Historia y crítica de la opinión pública*. Barcelona: Gustavo Gili, 2ª edición. Prólogo de Toni Domènech.
- Habermas, J. (1985), *Conciencia moral y acción comunicativa*, Barcelona: Península.
- Habermas, J. (1987). *Teoría de la acción comunicativa*. Madrid: Taurus.
- Habermas, J. (2009), *Philosophische Texte*, Band 4, Studienausgabe, Frankfurt: Suhrkamp.
- Habermas, J. (2006), *Entre naturalismo y religión*, Barcelona: Paidós.
- Habermas, J. (2019), *Auch eine Geschichte der Philosophie*, Berlin: Suhrkamp, 2 vols.
- Habermas, J. (2022), *Ein neuer Strukturwandel der Öffentlichkeit und deliberative Politik*, Berlin: Suhrkamp.
- Kant, I. (1968). *Was ist Aufklärung?* Kants Werke. Akademie Textausgabe, VIII, Berlin: Walter de Gruyter.
- Kant, I. (1981). *Crítica del Juicio*. Madrid: Espasa-Calpe.
- Kant, I. (1983). *Crítica de la razón pura*. Madrid: Alfaguara, 2ª edición.
- Kant, I. (1985). *La paz perpetua*. Madrid: Tecnos.
- Levitsky, S. y Zibblatt, D. (2018). *Cómo mueren las democracias*. Buenos Aires: Ariel.
- Maquiavelo, N. (1985), *El Príncipe*, Madrid: Cátedra.
- Maquiavelo, N. (1987), *Discursos sobre la primera década de Tito Livio*, I. 25, Madrid: Alianza.
- Marina, J.A. (2023), “¿Es de fiar la opinión pública?”  
<https://www.joseantoniomarina.net/categoria-blog/destacado/es-de-fiar-la-opinion-publica/>  
<https://ethic.es/2023/01/es-de-fiar-la-opinion-publica-democracia/>
- Martínez-Navarro, E. (2022), La aportación de Rawls a la Ética de la razón cordial desde su teoría de los sentimientos morales, en Pereira, G. y Pérez-Zafrilla, P. (eds.), *Actualidad de John Rawls en el siglo XXI*, Granada: Comares, 75-92.
- Mead, G.H. (1972), *Espíritu, persona y sociedad*, Buenos Aires: Paidós.
- Moreno, J.A. (2021/2), “Una distopía digital: Alex Portland”, *Acontecimiento*, 139, 3-5.
- Munzel, F. (2012), *Kant's Conception of Pedagogy*, Evanston: Northwestern University Press.
- Nicolás, J.A. (2019), Posverdad: cartografía de un fenómeno complejo. *Diálogo Filosófico*, 105, 302-340.
- Nietzsche, F. (1986), *El Gay saber*, Madrid: Espasa-Calpe.
- Noelle-Neumann, E. (1982), *La espiral del silencio. Opinión pública: nuestra piel social*. Barcelona: Paidós.
- O'Neil, C. (2018), *Armas de destrucción matemática*, Capitán Swing.
- Ortega, C. (2021), *Habermas ante el siglo XXI*, Madrid: Tecnos.
- Ortega y Gasset, J. (2005), La rebelión de las masas, *Obras completas*, tomo IV, Madrid: Taurus.

- Pérez-Zafrilla, P. (2021a), Polarización artificial: cómo los discursos expresivos inflaman la percepción de polarización política en internet. *RECERCA*, 26 (2), 1-23.
- Pérez-Zafrilla, P. (2021b), “Bases neuroéticas de la corrección política. Una aproximación desde la teoría de la espiral del silencio de Elisabeth Noelle-Neumann”, en López-Orellana, R. y Suárez-Ruiz, J. (eds.), *Filosofía posdarwiniana*, Londres: College Publications, 471-499.
- Rawls, J. (1971/1972). *A Theory of Justice*. Oxford University Press.
- Rawls, J. (1978). *Teoría de la justicia*. México: Fondo de Cultura Económica.
- Rawls, J. (1993). *Political Liberalism*. New York: Columbia University Press.
- Rawls, J. (1996). *El liberalismo político*. Barcelona: Crítica.
- Rawls, J. (2001). *El derecho de gentes*. Barcelona: Paidós.
- Rawls, J. (1999). *Collected Papers*. Harvard University Press.
- Sen, A. (2010), *La idea de la justicia*, Madrid: Taurus.
- Siurana, J.C. (2003), *Una brújula para la vida moral*, Granada: Comares.
- Siurana, J.C. (2021), *Ética para influencers*, Madrid: Plaza y Valdés.