

I

(Resoluciones, recomendaciones y dictámenes)

DICTÁMENES

COMITÉ ECONÓMICO Y SOCIAL EUROPEO

526º PLENO DEL CESE DE 31 DE MAYO Y 1 DE JUNIO DE 2017

Dictamen del Comité Económico y Social Europeo sobre la «Inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad»

(Dictamen de iniciativa)

(2017/C 288/01)

Ponente: **Cateljne MULLER**

Decisión del Pleno	22.9.2016
Fundamento jurídico	Artículo 29, apartado 2, del Reglamento interno Dictamen de iniciativa
Sección competente	Mercado Único, Producción y Consumo
Aprobado en sección	4.5.2017
Aprobación en el pleno	31.5.2017
Pleno n.º	526
Resultado de la votación (a favor/en contra/abstenciones)	159/3/14

1. Conclusiones y recomendaciones

1.1. La inteligencia artificial (IA) experimenta actualmente una serie de importantes avances y ofrece aplicaciones a un ritmo muy rápido en todos los ámbitos de la sociedad. El volumen del mercado de la IA se sitúa en torno a los 664 millones de dólares estadounidenses, y se espera que aumente a 38 800 millones USD en 2025. Dada la influencia tanto positiva como negativa de la IA en la sociedad, el CESE se ha comprometido a seguir la evolución en este ámbito, no solo desde el punto de vista técnico sino también desde la perspectiva de la ética, la sociedad y la seguridad.

1.2. El CESE, como representante de la sociedad civil europea, estructurará, centralizará e impulsará en los próximos tiempos un debate social sobre la IA en el que intervengan todas las partes interesadas pertinentes: responsables políticos, industria, interlocutores sociales, consumidores, ONG, instituciones de enseñanza y salud, así como expertos y académicos de diversas disciplinas (entre ellas, IA, seguridad, ética, economía, ciencias laborales y jurídicas, ciencias del comportamiento, psicología, filosofía).

1.3. Si bien es importante, el debate sobre la superinteligencia predomina en estos momentos y eclipsa el debate sobre el impacto de las aplicaciones actuales de la IA. Por tanto, parte de la labor y los objetivos de este proceso consistirá en aumentar y ampliar los conocimientos sobre la IA para contribuir a crear un debate equilibrado y bien fundado, sin alarmismos ni relativizaciones extremas. El CESE se compromete a promover el desarrollo de una IA que beneficie a la humanidad. Ahora bien, en este proceso, una importante labor y finalidad consiste también en reconocer, definir y controlar las disrupciones en el desarrollo de la IA, a fin de poder abordarlas adecuadamente y a su debido tiempo. Eso conducirá a un mayor grado de implicación y aceptación de la sociedad que nos permitirá continuar desarrollando y utilizando la IA de forma sostenible.

1.4. El impacto de la IA es transfronterizo, por lo que también sería necesario establecer marcos políticos supranacionales. El CESE recomienda que **la UE desempeñe un papel de liderazgo mundial** en la adopción de marcos políticos uniformes y universales para la IA, de conformidad con los valores europeos y los derechos fundamentales. El CESE puede y quiere aportar su contribución a este fin.

1.5. El CESE señala en estos momentos **once áreas** donde la IA plantea desafíos sociales: ética; seguridad; privacidad; transparencia y rendición de cuentas; trabajo; educación y desarrollo de capacidades; (des)igualdad e inclusión; legislación y reglamentación; gobernanza y democracia; guerra; y superinteligencia. El CESE formula las siguientes recomendaciones.

1.6. El CESE defiende un enfoque de la IA basado en el **control humano (*human-in-command*)**, con un marco de condiciones que regule el desarrollo responsable, seguro y útil de la IA de manera que las máquinas continúen siendo máquinas y los humanos conserven en todo momento el dominio sobre ellas.

1.7. El CESE pide que se elabore un **código deontológico** para el desarrollo, despliegue y utilización de la IA, de modo que durante todo su proceso de funcionamiento los sistemas de IA sean compatibles con los principios de la dignidad humana, la integridad, la libertad, la privacidad, la diversidad cultural y de género y los derechos humanos fundamentales.

1.8. El CESE aboga por el desarrollo de un **sistema de normalización** para la verificación, validación y control de los sistemas de IA, basado en un amplio espectro de normas en materia de seguridad, transparencia, inteligibilidad, rendición de cuentas y valores éticos.

1.9. El CESE aboga por una **infraestructura de IA europea** de fuente abierta (*open source*), que incluya entornos de aprendizaje respetuosos de la vida privada, entornos de ensayo en condiciones reales (*real life*) y conjuntos de datos de alta calidad para el desarrollo y la formación de sistemas de IA. El CESE destaca la ventaja (competitiva) que puede obtener la UE en el mercado mundial mediante el desarrollo y la promoción de «sistemas de IA de responsabilidad europea», provistos de un sistema europeo de certificación y etiquetado de la IA.

1.10. La UE, los gobiernos nacionales y los interlocutores sociales deberían determinar conjuntamente cuáles son los **sectores del mercado laboral** que se verán influidos por la IA, en qué medida y en qué plazo, y buscar soluciones para mitigar las consecuencias para el empleo, la naturaleza del trabajo, los sistemas sociales y la (des)igualdad. Además debería invertirse en los sectores del mercado laboral en los que la IA tiene poca o ninguna influencia.

1.11. El CESE recomienda a todas las partes interesadas que realicen un esfuerzo colectivo para apostar por los **sistemas complementarios de IA** y su **creación conjunta** en el lugar de trabajo, como equipos persona-máquina donde la IA sirva para complementar y mejorar la actuación humana. Las partes interesadas también deben invertir en **aprendizaje formal e informal, enseñanza y formación** a fin de que todos puedan trabajar con la IA, pero también para desarrollar capacidades que no pueda suplir la IA.

1.12. Es necesario llevar a cabo una **evaluación concreta de la legislación y la reglamentación de la UE** en las seis áreas de interés definidas por la Unidad de Prospectiva Científica (STOA) que son susceptibles de revisión o adaptación. El CESE puede y quiere desempeñar un papel en este proceso. El CESE está en contra de la introducción de cualquier tipo de personalidad jurídica para los robots o la IA, puesto que socavaría los efectos correctores preventivos de la legislación en materia de responsabilidad, generaría un riesgo moral tanto en el desarrollo como en la utilización de la IA y daría lugar a un posible uso indebido.

1.13. El desarrollo de las **aplicaciones de IA que benefician a la sociedad**, promueven la integración y la mejora de la vida de los ciudadanos debería promoverse y fomentarse activamente, tanto a nivel público como privado. La Comisión Europea debe financiar en sus programas la investigación sobre el impacto social de la IA, así como las innovaciones financiadas por la UE en materia de IA.

1.14. El CESE apoya el llamamiento de Human Rights Watch y otras organizaciones para que se **prohíban las armas autónomas**. En este sentido, celebra el acuerdo anunciado por las Naciones Unidas, pero opina que debería extenderse también a las aplicaciones de la IA en el ámbito de la ciber guerra.

2. Inteligencia artificial

2.1. No existe una definición establecida y unánimemente aceptada de la IA. La IA es un concepto que engloba muchas otras (sub)áreas como la informática cognitiva (*cognitive computing*: algoritmos capaces de razonamiento y comprensión de nivel superior —humano—), el aprendizaje automático (*machine learning*: algoritmos capaces de enseñarse a sí mismos tareas), la inteligencia aumentada (*augmented intelligence*: colaboración entre humanos y máquinas) o la robótica con IA (IA integrada en robots). Sin embargo, el objetivo fundamental de la investigación y el desarrollo en materia de IA es la automatización de comportamientos inteligentes como razonar, recabar información, planificar, aprender, comunicar, manipular, observar e incluso crear, soñar y percibir.

2.2. En términos generales, se puede distinguir entre IA débil (*narrow AI*) e IA fuerte (*general AI*). La IA débil es capaz de realizar tareas específicas. La IA fuerte es capaz de realizar las mismas tareas intelectuales que un ser humano.

2.3. En el ámbito de la IA débil se ha progresado considerablemente en los últimos tiempos, gracias sobre todo al crecimiento de la capacidad de procesamiento (*computer processing power*), la disponibilidad de grandes cantidades de datos y el desarrollo del aprendizaje automático (*machine learning* o ML). El aprendizaje automático incluye algoritmos capaces de enseñarse a sí mismos tareas específicas sin estar programados para ello. El método se basa en el procesamiento de «datos de entrenamiento» que sirven de base al algoritmo para aprender a reconocer patrones y formular normas. El aprendizaje profundo (*deep learning* o DL), una forma de ML, utiliza estructuras de redes neuronales (*neural networks*) basadas a grandes rasgos en el cerebro humano que aprenden mediante el ensayo y la respuesta. El resultado de estos avances es que los sistemas de IA (por medio de algoritmos) ya pueden aprender por sí mismos, y ser autónomos y adaptativos.

2.4. Los objetivos principales de la investigación y el desarrollo en IA residen desde hace tiempo en el razonamiento, la adquisición de conocimientos, la planificación, la comunicación y la percepción (visual, auditiva y sensorial). De ahí se derivan numerosas aplicaciones de la IA: asistentes virtuales, automóviles sin conductor, publicación automática de noticias, reconocimiento de voz, programas de traducción, programas de síntesis de voz, negociaciones financieras automatizadas, investigación electrónica jurídica (*e-discovery*), etc.

2.5. El CESE señala que el número de aplicaciones e inversiones en IA ha crecido exponencialmente en los últimos tiempos. El volumen del mercado de la IA se sitúa en la actualidad en torno a unos 664 millones de dólares estadounidenses, y se espera que aumente a 38 800 millones USD en 2025.

3. Oportunidades y amenazas de la IA

3.1. Casi nadie pone en tela de juicio que la IA ofrece grandes ventajas sociales: entre los ejemplos están la agricultura sostenible, la seguridad vial, la seguridad del sistema financiero, procesos productivos más respetuosos con el medio ambiente, mejor sanidad, mayor seguridad en el trabajo, una enseñanza más personalizada, una justicia de mayor calidad y una sociedad más segura. Quizá hasta pueda contribuir a la erradicación de la enfermedad y la pobreza. La IA también puede contribuir de manera importante al crecimiento de la industria y a la mejora de la competitividad de la UE.

3.2. Sin embargo, como cualquier otra tecnología disruptiva, la IA entraña riesgos y retos políticos complejos en áreas como la seguridad, la controlabilidad, los aspectos socioeconómicos, la ética, la privacidad, la fiabilidad, etc.

3.3. Estamos en un momento crucial para definir las condiciones en las que se enmarcarán el desarrollo presente y futuro de la IA y su utilización. Las ventajas de la IA solo se pueden obtener de manera sostenible si se abordan adecuadamente los retos que conlleva. Para ello es preciso tomar decisiones políticas.

a) Ética

3.4. El desarrollo de la IA plantea numerosas cuestiones éticas. ¿Cómo influye la IA autónoma (autodidacta) sobre nuestra integridad, autonomía, dignidad, independencia, igualdad, seguridad y libertad de elección personal? ¿Cómo aseguramos que se respeten y garanticen nuestros valores, normas y derechos humanos fundamentales?

3.5. Cabe señalar a este respecto que, actualmente, el desarrollo de la IA se produce en un entorno homogéneo compuesto principalmente de varones jóvenes blancos, lo cual deja una impronta (consciente o no) de disparidad cultural y de género, entre otras cosas porque los sistemas de IA aprenden sobre la base de datos de formación. Esos datos deben ser correctos, y además de buena calidad, variados, suficientemente profundos y ecuánimes. Existe una tendencia generalizada a creer que los datos son por definición objetivos, lo cual no es cierto. Los datos son fáciles de manipular, pueden estar sesgados, reflejar prejuicios y preferencias culturales, de género o de otra naturaleza, y contener errores.

3.6. Los sistemas de IA que se diseñan actualmente no contienen valores éticos. Incorporarlos a dichos sistemas y al entorno donde se utilizan es tarea de los humanos. El desarrollo, la activación y el uso de los sistemas de IA (tanto públicos como comerciales) deben supeditarse a nuestros valores, normas, libertades y derechos humanos fundamentales. Por consiguiente, el CESE aboga por el desarrollo de un código deontológico uniforme y universal para el desarrollo, despliegue y utilización de la IA.

b) *Seguridad*

3.7. El uso de la IA en el mundo real plantea evidentes problemas de seguridad. Puede hacerse una distinción entre la seguridad interna y la externa.

— *Seguridad interna*: ¿el sistema de IA es lo bastante sólido para funcionar (seguir funcionando) correctamente? ¿El algoritmo está bien programado? ¿No tiene posibilidad de fallar, es resistente a los ciberataques? ¿Es eficaz? ¿Es fiable?

— *Seguridad externa*: ¿el sistema de IA es seguro para su uso en la sociedad? ¿Funciona siempre con seguridad, no solo en condiciones normales sino también en situaciones desconocidas, críticas o imprevisibles? ¿Cómo influye la capacidad de autoaprendizaje en su seguridad, especialmente cuando continúa aprendiendo después de su puesta en servicio?

3.8. El CESE considera que los sistemas de la IA solo puedan entrar en funcionamiento si cumplen una serie de requisitos específicos de seguridad interna y externa. Estos requisitos deben determinarlos conjuntamente los responsables políticos, los especialistas en IA y seguridad, las empresas y las organizaciones de la sociedad civil.

c) *Transparencia, inteligibilidad, controlabilidad, rendición de cuentas*

3.9. La aceptación de la IA, así como su desarrollo y aplicación sostenibles, dependen de la posibilidad de entender su funcionamiento, sus acciones y sus decisiones, controlarlos y responder por ellos, especialmente *a posteriori*.

3.10. Las acciones y decisiones de los sistemas de IA (por medio de algoritmos inteligentes) influyen cada vez más en la vida de las personas. Cabe citar como ejemplos la utilización de la IA en la función policial basada en la inteligencia, para la evaluación de solicitudes de créditos hipotecarios o para procedimientos de autorización en el sector de los seguros. La inteligibilidad, controlabilidad y rendición de cuentas del proceso de toma de decisiones de un sistema de IA es crucial en este sentido.

3.11. Actualmente ya hay muchos sistemas de IA que resultan poco comprensibles para los usuarios. Lo mismo se aplica, cada vez en mayor medida, a los desarrolladores de esos sistemas. Las redes neuronales (*neural networks*), sobre todo, son a menudo cajas negras donde ocurren procesos (de decisión) que ya no se pueden descifrar y para los que no existen mecanismos explicativos.

3.12. El CESE aboga por unos sistemas de IA transparentes, comprensibles y controlables cuyo funcionamiento sea objeto de rendición de cuentas también *a posteriori*. Además, debería establecerse qué procesos de toma de decisiones pueden confiarse a los sistemas de IA y cuáles no y cuándo es aconsejable o necesaria la intervención humana.

d) *Privacidad*

3.13. La privacidad de los sistemas de IA es motivo de preocupación. Muchos productos (de consumo) ya llevan IA incorporada: aparatos domésticos, juguetes, coches o dispositivos de control de la salud. Todos ellos transmiten datos (a menudo personales) a las plataformas de los fabricantes en la nube. En vista del auge del comercio de datos, que significa que los datos que se generan no van a parar solo al fabricante, sino que se venden a terceros, un motivo de preocupación es si la privacidad está suficientemente garantizada.

3.14. Por otra parte, la IA es capaz de influir en las decisiones humanas a través del análisis de grandes cantidades de datos (a menudo personales) en muchos terrenos (desde decisiones comerciales a elecciones y referendos). Los menores constituyen un grupo particularmente vulnerable. El CESE manifiesta su preocupación por las aplicaciones de IA dirigidas explícitamente a influir en el comportamiento y los deseos de los niños.

3.15. Debe evitarse que la aplicación de la IA al tratamiento de los datos personales restrinja la libertad real o percibida de las personas. El Reglamento general de protección de datos (RGPD) contiene importantes disposiciones sobre la protección de la privacidad de los datos personales suministrados digitalmente. Es preciso controlar adecuadamente si en la práctica está suficientemente garantizado el derecho de las personas al consentimiento informado y la libertad de elección en la transmisión de datos, y también al acceso, la adaptación y el control de los datos emitidos, habida cuenta de la evolución de la IA.

e) **Normas e infraestructuras**

3.16. Es necesario desarrollar nuevos procedimientos normativos basados en una amplia variedad de requisitos para verificar y validar los sistemas de IA, a fin de evaluar y controlar su seguridad, transparencia, inteligibilidad, rendición de cuentas y responsabilidad ética.

3.17. El CESE recomienda que la UE desarrolle su propia infraestructura de IA, integrada por entornos de aprendizaje de fuente abierta (*open source*) y respetuosos de la vida privada y de conjuntos de datos de alta calidad para el desarrollo y la formación de sistemas de IA. La UE puede obtener una ventaja (competitiva) en el mercado mundial mediante el desarrollo y la promoción de «sistemas de IA de responsabilidad europea». El Comité recomienda a este respecto que se examinen las posibilidades de un sistema europeo de certificación y etiquetado de la IA.

f) **Repercusiones en el trabajo, el empleo, las condiciones laborales y los sistemas sociales**

3.18. Las opiniones acerca de la rapidez con que se producirán y cuál será su alcance son muy diversas, pero es evidente que la IA influirá en la disponibilidad de empleo y en la naturaleza y el carácter de muchos puestos de trabajo, y por lo tanto también en los sistemas sociales.

3.19. Brynjolfsson y McAfee, del MIT (Instituto Tecnológico de Massachusetts), llaman a los actuales desarrollos tecnológicos (incluida la IA) la «segunda era de las máquinas». Hay dos grandes diferencias: i) las máquinas «antiguas» servían sobre todo para reemplazar la fuerza muscular, mientras que las nuevas sustituyen la capacidad intelectual y las habilidades cognitivas, lo que tiene repercusiones no solo para los trabajadores poco cualificados (*blue collar workers*) sino también para los trabajadores cualificados (*white collar workers*), y ii) la IA constituye una tecnología de alcance general (*general purpose technology*) que incide en todos los sectores al mismo tiempo.

3.20. La IA puede reportar grandes ventajas cuando se utiliza en trabajos peligrosos, pesados, extenuantes, sucios, desagradables, repetitivos o aburridos. Pero los sistemas de IA también son capaces, y cada vez más, de realizar tareas rutinarias, procesar y analizar datos o desempeñar funciones que requieren un alto grado de planificación o predicción, trabajos que a menudo se encomiendan a trabajadores muy capacitados.

3.21. Sin embargo, la mayoría de los empleos conllevan diversas tareas. Parece poco probable que la IA o un robot puedan asumir todas las tareas de una persona. Sin embargo, casi todo el mundo tendrá que hacer frente a la automatización de algunas de sus funciones. El tiempo libre resultante podrá dedicarse a otras tareas, siempre que las administraciones y los interlocutores sociales atiendan a esa cuestión. Es necesario por tanto tener en cuenta la incidencia que estas situaciones pueden tener en los profesionales y gestores, y promover su participación en el proceso, para que sean dueños y no víctimas de la situación.

3.22. Además, se crearán nuevos puestos de trabajo. Realmente nadie puede prever cuáles serán, cuánto abundarán ni con qué celeridad aparecerán. Empresas como Google y Facebook han logrado crear un valor inconmensurable con un número relativamente reducido de empleados. Por otra parte, los empleos nuevos no son siempre buenos empleos. Existe la preocupación de que la continuidad del desarrollo de la IA deje solo tareas pequeñas y mal remuneradas para un grupo creciente de trabajadores flexibles.

3.23. Las consecuencias de la IA no afectarán solo a la cantidad de trabajo disponible, sino también a la naturaleza del trabajo existente. Los sistemas de IA ofrecen cada vez más opciones para supervisar y controlar a los empleados, lo que pone en tela de juicio su autonomía y privacidad. La organización y distribución del trabajo cada vez se hace más mediante algoritmos, sin intervención humana, lo que influye en la naturaleza y las condiciones del trabajo. El uso de los sistemas de IA conlleva un riesgo de empobrecimiento de las tareas y pérdida de habilidades importantes.

3.24. El hecho es, sin embargo, que la tecnología no es algo que «sobrevenga» sin más. Los poderes públicos y los interlocutores sociales tienen la posibilidad de determinar cómo se desarrollará la IA y cómo se aplicará en el marco de la actividad profesional; hay que aferrarse a esta posibilidad. Por eso es importante no centrarse solo en lo que puede hacer la IA, sino también en lo que pueden hacer las personas (creatividad, empatía, colaboración), lo que queremos que sigan haciendo, y buscar formas de que humanos y máquinas puedan trabajar mejor juntos (complementariedad).

3.25. La inteligencia aumentada (complementariedad), donde persona y máquina trabajan juntas y se fortalecen mutuamente, es la aplicación más interesante de la IA porque se trata del humano *con* la máquina, no del humano en lugar de la máquina. Sin embargo, la participación en su creación es de gran importancia: hay que involucrar a los trabajadores en el desarrollo de esos sistemas de IA complementarios para garantizar que sean fáciles de usar y que el trabajador conserve suficiente autonomía y control (*human-in-command*) y satisfacción en su trabajo.

g) *Educación y desarrollo de capacidades*

3.26. Las personas necesitan adquirir o actualizar las capacidades digitales para ser capaces de adaptarse a la rápida evolución en materia de IA. La Comisión Europea se ha comprometido a fondo con la alfabetización digital con su Coalición por las capacidades y los empleos digitales (*Digital Skills and Jobs Coalition*). Sin embargo, no todo el mundo puede codificar y hacerse programador o está interesado en ello. Por eso es necesario que la política y los recursos financieros se destinen también al aprendizaje y desarrollo de competencias en áreas donde los sistemas de IA supongan una amenaza (como las que requieren fundamentalmente interacción humana o colaboración entre el humano y la máquina, o las que queremos que sigan en manos de los humanos).

3.27. Apostar por la complementariedad entre el humano y la IA (inteligencia aumentada) requerirá una educación generalizada en el trabajo con los sistemas de IA y su manejo desde una edad temprana, a fin de que las personas conserven la autonomía y el control en su trabajo (control humano). En este sentido, es especialmente importante la educación sobre ética y privacidad, puesto que la IA repercute fuertemente en esos ámbitos.

h) *Accesibilidad, (des)igualdad social, inclusión, reparto*

3.28. La mayoría de los avances en IA y sus correspondientes elementos (plataformas de desarrollo, datos, conocimiento y experiencia) están en manos de cinco grandes empresas tecnológicas (Amazon, Apple, Facebook, Google y Microsoft). Aunque son partidarias del desarrollo libre de la IA, y algunas ofrecen sus plataformas en código abierto, eso no garantiza una plena accesibilidad a los sistemas de IA. La UE, los responsables políticos internacionales y las organizaciones sociales deben desempeñar un importante papel garantizando el acceso a los sistemas de IA para todo el mundo, así como su desarrollo en un entorno abierto.

3.29. Los cambios tecnológicos favorables para el capital, con innovaciones que beneficien sobre todo a sus propietarios, empeoran la posición de la mano de obra frente a la del capital. Los cambios tecnológicos también pueden dar lugar a desigualdades (de ingresos) entre las personas (tanto a nivel local como a nivel regional y mundial). La IA puede intensificar estas tendencias.

3.30. Es importante seguir de cerca estas tendencias para poder abordarlas adecuadamente. Se aboga por un impuesto para las IA, un reparto de los dividendos que generan o que la propiedad de los sistemas de IA se comparta entre empleados y empleadores. También se habla cada vez más de la necesidad de unos ingresos básicos incondicionales.

3.31. En un dictamen previo⁽¹⁾, el CESE señala la posibilidad de un dividendo digital y el reparto equitativo de dividendos, con el fin de lograr efectos positivos sobre el crecimiento. El CESE concede importancia a la búsqueda de este tipo de soluciones, pero opina que es necesario evitar el entorpecimiento del desarrollo de una IA que beneficie a la humanidad. También debe prevenirse el riesgo moral de que la responsabilidad sobre los sistemas de IA se transfiera a entidades que, en la práctica, sean incapaces de responder.

⁽¹⁾ DO C 13 de 15.1.2016, p. 161.

i) **Legislación y reglamentación**

3.32. Las repercusiones de la IA sobre la legislación y la reglamentación existentes son considerables. La Unidad de Prospectiva Científica (STOA) del Parlamento Europeo presentó en junio de 2016 una relación de la legislación y reglamentación europea que se verá afectada por los avances en materia de robótica, sistemas ciberfísicos e IA. Se citan seis áreas de interés —transporte, sistemas de doble uso, libertades civiles, seguridad, salud y energía— donde puede ser necesario revisar o adaptar nada menos que 39 directivas, reglamentos, declaraciones y comunicaciones de la UE, y la Carta de los Derechos Fundamentales de la Unión Europea. Esta evaluación debería abordarse con urgencia y determinación, y el CESE puede y quiere desempeñar un papel en este proceso.

3.33. Existe mucha controversia sobre la cuestión de quién es el responsable de los daños que pueda causar un sistema de IA, sobre todo cuando se trata de sistemas autodidactas que continúan aprendiendo después de su entrada en servicio. El Parlamento Europeo ha formulado algunas recomendaciones relativas a la legislación civil en materia de robótica, incluida la propuesta de examinar la posibilidad de dotar a los robots de una «personalidad jurídica» (*e-personality*) para poder atribuirles la responsabilidad civil por los daños que causen. El CESE se opone a cualquier tipo de estatuto jurídico para los robots o sistemas de IA por el riesgo moral inaceptable que ello conlleva. La legislación en materia de responsabilidad tiene un efecto correctivo y preventivo que podría desaparecer en cuanto el riesgo de responsabilidad civil dejase de recaer sobre el autor por haberse transferido al robot (o sistema de IA). Además, una forma jurídica así sería susceptible de uso y aplicación indebidos. La comparación con la responsabilidad limitada de las sociedades no es válida, puesto que el responsable en última instancia es siempre una persona física. A este respecto, hay que investigar en qué medida responden satisfactoriamente a este problema la legislación, la reglamentación y la jurisprudencia de la UE y de cada Estado miembro en materia de responsabilidad (sobre el producto y el riesgo) y atribución de culpa, y, en su defecto, qué soluciones legales se ofrecen.

3.34. Además, para abordar adecuadamente la legislación y la normativa en materia de IA se requiere un buen entendimiento de lo que puede conseguir la IA y lo que no a corto, medio y largo plazo.

3.35. La IA no conoce fronteras. Por consiguiente, es importante examinar la necesidad de una reglamentación mundial, pues la normativa regional será insuficiente e incluso de efectos no deseados. Visto su sistema de normas sobre seguridad de los productos, o la tentación de proteccionismo en otros continentes, el alto nivel de conocimiento en toda Europa, el sistema europeo de los derechos fundamentales y de valores sociales y diálogo social, el CESE recomienda que la UE adopte un papel de liderazgo estableciendo marcos universales y uniformes para la IA, y que fomente este proceso a nivel mundial.

j) **Gobernanza y democracia**

3.36. Las aplicaciones de IA pueden contribuir a una mejor participación de los ciudadanos en la política pública y a la transparencia en la toma de decisiones. El CESE pide a la UE y a los gobiernos nacionales que utilicen la IA en este contexto.

3.37. El CESE manifiesta su preocupación por el uso específico de los sistemas de IA (en forma de algoritmos inteligentes) para la agregación de informaciones, por ejemplo en las redes sociales, que parece haber entrañado una limitación y fragmentación de la sociedad (por ejemplo, las *filter bubbles* o «burbujas» y las pseudonoticias en Twitter y Facebook durante las elecciones estadounidenses).

3.38. El CESE también muestra su preocupación por los indicios de que supuestamente se utilizaron sistemas de IA para influir en el comportamiento (electoral) de los ciudadanos. A través de los algoritmos inteligentes es posible predecir los comportamientos humanos e influir en ellos. Esto es una amenaza para una democracia abierta y fiable. En la época actual de polarización y desmantelamiento de instituciones internacionales, la precisión y el poder de semejante tecnología de propaganda puede acarrear rápidamente una mayor desestabilización social. Esta es una de las razones por las que se necesitan normas de transparencia y rendición de cuentas de los algoritmos (inteligentes).

k) **Guerra**

3.39. La Convención de las Naciones Unidas sobre armas convencionales decidió consultar en 2017 la opinión de expertos gubernamentales acerca de las implicaciones de las armas autónomas. El CESE participó y apoyó el llamamiento de Human Rights Watch y otras organizaciones para prohibir el armamento autónomo. El CESE considera que esta prohibición debe volver a analizarse y examinarse seriamente. No obstante, eso no basta para abordar adecuadamente el uso de la IA en situaciones de guerra y conflicto. El debate de las Naciones Unidas debe incluir también las aplicaciones de la IA en caso de ciberguerra.

3.40. También debe evitarse que la IA vaya a parar a manos de personas o regímenes que pretendan utilizarla para actividades terroristas.

l) *Superinteligencia*

3.41. Por último se plantea la cuestión de las posibilidades y los riesgos que trae consigo el desarrollo de una superinteligencia. Stephen Hawking opina que el desarrollo de una IA fuerte significaría el fin de la humanidad. Llegado ese momento, la IA evolucionaría a un ritmo que los humanos no podrían seguir, según Hawking. Algunos expertos abogan por lo que llaman un *kill-switch* o *reset-button*, un botón para desactivar o reiniciar los sistemas de IA desbocados o superinteligentes.

3.42. El CESE defiende el control humano, con un marco de condiciones que regule el desarrollo y uso responsable y seguro de la IA, de manera que las máquinas continúen siendo máquinas, y los humanos conserven en todo momento el dominio sobre ellas. Por tanto, el debate sobre la superinteligencia es importante, pero en este momento eclipsa en exceso a la cuestión de las repercusiones de las aplicaciones actuales de la IA.

4. Una IA al servicio de la humanidad

4.1. Algunos grandes operadores comerciales ya han desplegado iniciativas para fomentar el desarrollo abierto, seguro y socialmente responsable de la IA (como OpenAI). Sin embargo, los responsables políticos no pueden dejarlo en manos de las empresas, sino que deben asumir su papel. Se necesitan medidas específicas y apoyo a la investigación de los retos sociales que plantea la IA, así como al desarrollo de sistemas de IA seguros y sólidos.

4.2. Hay programas europeos, entre ellos Horizonte 2020, que resultan sumamente apropiados para afrontar este reto. El CESE ha constatado que la financiación del pilar «Retos de la sociedad» de Horizonte 2020 queda por debajo de la que se obtiene para los otros dos, «Ciencia excelente» y «Liderazgo industrial», y está disminuyendo. El CESE insiste en que la investigación sobre los extensos desafíos y las aplicaciones sociales de la IA ocupe un lugar importante en el pilar «Retos de la sociedad».

4.3. También es necesario abordar los posibles efectos intersectoriales de la IA. Además de la financiación del desarrollo de innovaciones disruptivas de IA, hay que financiar también la investigación y las maneras de abordar el impacto social de las innovaciones.

4.4. La investigación y el desarrollo de una IA que esté al servicio de la humanidad requiere también una amplia variedad de formaciones de libre acceso y datos procedentes de entornos de prueba en condiciones reales que cumplan los requisitos cualitativos. Por el momento, la infraestructura de IA y muchos datos de calidad solo están disponibles para ciertos organismos privados, y existen obstáculos para probar la IA en el espacio público, lo que dificulta su aplicación en otras áreas. El desarrollo de datos de calidad de libre acceso y de una infraestructura de IA europea es crucial para lograr una IA segura, sólida y útil.

5. Supervisión y respuesta

5.1. Todavía no es posible hacerse una idea global de las amplias repercusiones sociales de la IA. Sin embargo, nadie duda de su importancia. Las novedades en torno a la IA se suceden actualmente a un ritmo muy rápido, lo que requiere una supervisión exhaustiva partiendo de una perspectiva amplia para reaccionar adecuada y oportunamente a los cambios importantes y disruptivos, tanto técnicos como sociales, inherentes a la IA o relacionados con ella (*gamechangers* o «cambios críticos»).

5.2. Serían cambios críticos técnicos, por ejemplo, los avances destacados o significativos en el desarrollo de las capacidades de la IA que pudieran presagiar la llegada de la IA fuerte. Cambios críticos sociales serían los que entrañaran una pérdida considerable de empleo sin perspectivas de reemplazo, situaciones de inseguridad, fallos del sistema, cambios internacionales inesperados, etc.

5.3. Los responsables políticos, la industria, los interlocutores sociales, los consumidores, las ONG, las instituciones de enseñanza y salud, así como los académicos y especialistas de diversas disciplinas (entre ellas, IA —aplicada—, ética, seguridad, economía, ciencias laborales y jurídicas, ciencias del comportamiento, psicología, filosofía) tienen que aunar esfuerzos para seguir de cerca las novedades en materia de IA, identificar estos cambios críticos y mantenerse al día, a fin de poder tomar las medidas adecuadas en el momento oportuno, en forma de políticas, legislación y normativa, autorregulación y diálogo social.

5.4. El CESE, como representante de la sociedad civil europea, estructurará, centralizará e impulsará en los próximos tiempos el debate de todas las partes interesadas en torno a la IA.

Bruselas, 31 de mayo de 2017.

Presidente
del Comité Económico y Social Europeo
Georges DASSIS
