

# **“Hacia una inteligencia artificial por y para la sociedad”**

**Nuria Oliver**

**Llició inaugural del  
curs acadèmic 2019-2020  
del sistema universitari  
català**

**6 de setembre del 2019**



# Hacia una inteligencia artificial por y para la sociedad

Molt Honorable President, Rector Magnífic de la UPF, Excel·lentíssima Alcaldessa, presidenta del Consell Social, Honorable Consellera, il·lustríssims regidors i regidores, rectores i rectors magnífics, autoritats, membres de la comunitat universitària, estudiants, amics i amigues, senyores i senyors,

Els agraeixo la seva presència en aquest solemne acte d'inauguració del curs acadèmic 2019-2020, i especialment agraeixo al rector Jaume Casals la seva amable invitació a pronunciar aquesta conferència inaugural. Gràcies per la confiança dipositada en mi.

Reconozco que mi tarea no es fácil. En primer lugar, dada la altísima calidad de los discursos de mis predecesores. En segundo lugar, porque el comienzo de curso suele ser duro, y los ánimos no acostumbran a estar muy boyantes. El inicio de curso marca el regreso a la rutina después del descanso estival, las clases, los días cada vez más cortos, los madrugones, el estrés... Quizás debería haber propuesto pronunciar la conferencia de clausura de curso, cuando ya se respira el verano, las vacaciones –llenas de posibilidades– son inminentes y hay un sentimiento generalizado de alegría... Espero que esta conferencia magistral les inspire y marque el comienzo de un curso apasionante.

Me hace mucha ilusión pronunciar esta conferencia en la UPF, una universidad hacia la que siento un cariño especial y que considero casi una segunda casa. Hace doce años regresé de Estados Unidos aquí, a Barcelona, y desde entonces he colaborado en distintas ocasiones con la UPF, impartiendo clases en programas de postgrado, actuando como miembro de tribunal de tesis de doctorado y, más intensamente, formando parte del Comité Científico Asesor Externo de la Ingeniería en Tecnologías de la Información y las Comunicaciones (ETIC). A través de este trabajo asesor he podido conocer con detalle tanto el programa docente como el de investigación e innovación de la ETIC, y he de reconocer que me han impresionado. Destaca el espíritu moderno y abierto de esta universidad, su internacionalización y la excelente calidad científica y humana de sus profesores, a muchos de los cuales considero amigos.

Además, hace exactamente un año celebramos con éxito el congreso internacional ACM MobileHCI 2018, que tuvo lugar en este mismo auditorio. Fui coorganizadora, junto al profesor Josep Blat, de esta casa, y la profesora Lynne Baillie, de la Universidad Heriot-Watt, en Edimburgo. Era la primera vez que el MobileHCI, el principal congreso del ACM dedicado a la interacción persona-móvil, se celebraba en la península Ibérica, coincidiendo con el vigésimo aniversario de su creación.

Desde hace 25 años he centrado mi actividad investigadora en el modelado computacional del comportamiento humano utilizando técnicas de inteligencia artificial. He creado habitaciones, ropas, coches y móviles inteligentes a través de decenas de proyectos, que han tenido impacto no solo científico, sino también social. Además, algo que quizás no es común en este país, he podido desarrollar una carrera científica en un contexto empresarial.

He sido no solo testigo sino también partícipe activo del progreso tecnológico, de la presencia cada vez más relevante y ubicua de la tecnología en nuestras vidas y de la codependencia que hemos desarrollado con la misma.

Durante mi carrera, la inspiración y el motor de mi trabajo ha sido siempre una pregunta con una clara aplicación social. En todos mis proyectos, la persona –a nivel individual o colectivo– ha sido y es el elemento central: tecnología por y para la sociedad. Tecnología que nos entiende, como paso previo para que pueda ayudarnos. Sin embargo, hoy, más que nunca, siento preocupación por el impacto que

dicha tecnología está teniendo en nuestras vidas, impacto no necesariamente siempre positivo. Me pregunto si nos encontramos ante una posible **crisis social de base tecnológica**.

La disponibilidad de cantidades ingentes de **datos** (Big Data) sobre el comportamiento humano, en gran parte resultado de la adopción masiva de los dispositivos móviles (ya hay muchos más móviles en el mundo que seres humanos) y de la digitalización del mundo físico –impulsado, entre otras cosas, por el desarrollo del Internet de las cosas– están transformando el mundo en que vivimos.

Hoy contamos con cantidades masivas de **datos** que podemos utilizar para crear algoritmos de **inteligencia artificial** que permiten a investigadores, empresas, gobiernos y otros actores del sector público abordar problemas complejos.

Decisiones con impacto tanto individual como colectivo que previamente eran tomadas por humanos –con frecuencia expertos– cada vez más son tomadas hoy en día por sistemas de **inteligencia artificial**, incluyendo decisiones relativas a la contratación de personas, las sentencias judiciales, la concesión de créditos y préstamos, los tratamientos y diagnósticos médicos o la compra-venta de acciones en bolsa.

Las decisiones algorítmicas basadas en datos tienen el potencial de mejorar nuestra toma de decisiones. De hecho, la historia ha demostrado que las decisiones humanas no son perfectas: están sujetas a conflictos de interés, a corrupción y a sesgos cognitivos, lo que ha resultado en procesos y resultados injustos y/o ineficientes. Por tanto, el interés hacia el uso de algoritmos puede interpretarse como el resultado de una demanda de mayor objetividad en la toma de decisiones. El potencial es inmenso.

Precisamente desde hace diez años, una de las áreas de investigación que creé primero en Telefónica, luego en Vodafone y a la que está dedicada la ONG Data-Pop Alliance –donde soy investigadora jefa de Datos– es el área que llamamos **ciencia de datos e inteligencia artificial para el bien social**. El objetivo es investigar el valor que tiene el Big Data, datos –agregados y anonimizados– para ayudarnos a tomar mejores decisiones que impactan en la vida de millones de personas en áreas alineadas con los **17 Objetivos de Desarrollo Sostenible**, como son la salud pública, la inclusión financiera, la educación, la respuesta ante desastres naturales, las migraciones y el cambio climático.

Sin embargo, la toma de decisiones basadas en algoritmos creados con datos no está exenta de limitaciones, que describo a continuación. Las siguientes palabras de Platón, pronunciadas hace 2400 años, son en este sentido sorprendentemente vigentes: “Una buena decisión está basada en conocimiento, no en números (datos)”

Cuando estas decisiones afectan a miles o millones de personas, surgen dilemas importantes. Por ejemplo, ¿quiere esto decir que las decisiones automáticas escapan a nuestro control? ¿Qué niveles de seguridad poseen estos sistemas para protegerse de los ciberataques o de un uso malicioso de los mismos? ¿Cómo podemos garantizar que dichas decisiones y/o actuaciones no tienen consecuencias negativas para las personas? ¿Quién es responsable de dichas decisiones? ¿Qué sucederá cuando un algoritmo nos conozca a cada uno de nosotros mejor que nosotros mismos y pueda aprovechar dicho conocimiento para manipular de manera subliminal nuestro comportamiento?

Por ello, me gustaría destacar **cinco pilares** que deberíamos abordar para conseguir que esta nueva manera de tomar decisiones, basada en la **IA** y los **datos**, tenga impacto positivo en la sociedad, y que resumo con el acrónimo inglés **FATEN**:

1. **F de *fairness* o justicia**: Las decisiones basadas en algoritmos pueden discriminar, porque los datos utilizados para crear dichos algoritmos pueden contener sesgos que den lugar a decisiones discriminatorias, por el uso de un determinado algoritmo; o por el mal uso de ciertos modelos en diferentes contextos.

2. **La A es de *autonomy* o autonomía**, valor central en la ética occidental, según la cual cada persona debería tener la capacidad de decidir sus propios pensamientos y acciones, asegurando por tanto la libre elección, la libertad de pensamiento y acción. Sin embargo, hoy en día podemos construir modelos computacionales de nuestros deseos, necesidades, personalidad y comportamiento, con capacidad para influir en nuestras decisiones y acciones de manera subliminal, como ha quedado patente en procesos electorales recientes en Estados Unidos y en el Reino Unido.

Por ello, deberíamos garantizar que los sistemas inteligentes autónomos adapten su toma de decisiones preservando la autonomía y la dignidad humanas.

- **La A también es de *accountability* o atribución de responsabilidad:** es decir, teniendo claridad respecto a la atribución de responsabilidad de las consecuencias de las decisiones algorítmicas.
- **Y de *augmentation* o aumento de la inteligencia humana,** de manera que los sistemas de inteligencia artificial se utilicen para aumentar o complementar la inteligencia humana, no para reemplazarla.

3. **La T es de *trust* o confianza.** La confianza es un pilar básico en las relaciones entre humanos e instituciones. La tecnología necesita la existencia de un entorno de confianza con sus usuarios, que cada vez más delegan (delegamos) sus vidas a servicios digitales. Sin embargo, el sector tecnológico está experimentando una pérdida de confianza por parte de la sociedad, alimentada por escándalos recientes como los de Facebook / Cambridge Analytica y Huawei.

Para que exista confianza han de cumplirse tres condiciones: (1) la competencia, es decir, la habilidad para realizar con solvencia la tarea comprometida; (2) la fiabilidad, es decir, la competencia sostenida en el tiempo; y (3) la honestidad y transparencia. Por ello:

- **La T también es de *transparency* o transparencia:** La transparencia hace referencia a la cualidad de poder entender un modelo computacional. Un modelo es transparente si una persona puede observarlo y entenderlo con facilidad. Las decisiones algorítmicas pueden ser no transparentes (opacas) por tres motivos: intencionalmente, para proteger la propiedad intelectual de los creadores de dichos algoritmos; por la falta de conocimiento por parte de los usuarios, que les impida entender cómo funcionan los algoritmos y modelos computacionales construidos a partir de los datos; e intrínsecamente, dado que ciertos métodos de aprendizaje por ordenador (por ejemplo, modelos de deep learning o aprendizaje profundo) son extremadamente complejos.

Asimismo, es imprescindible que los sistemas de inteligencia artificial sean transparentes no solo con relación a qué datos captan y analizan sobre el comportamiento humano y para qué propósitos, sino también respecto a en qué situaciones los humanos están interactuando con sistemas artificiales (por ejemplo, chatbots) y no con otros humanos.

4. **La E es de *education* o educación,** es decir, que hace referencia a la inversión en educación a todos los niveles, empezando por la educación obligatoria. En el libro Los nativos digitales no existen escribí

un capítulo llamado “Eruditos digitales” en el que enfatizo la necesidad de enseñar tanto pensamiento computacional desde primero de primaria como de desarrollar el pensamiento crítico, la creatividad y las habilidades de las inteligencias social y emocional que hoy en día no estamos desarrollando pero que cada vez van a resultar más importantes para nuestra salud mental y nuestra coexistencia pacífica y armoniosa, tanto con la tecnología como con otros seres humanos y con nuestro planeta.

Educación también de la ciudadanía, de los profesionales –sobre todo a aquellos cuyas profesiones están siendo transformadas por la tecnología– y de los políticos y los trabajadores del sector público. Una apuesta ambiciosa por la educación en competencias tecnológicas es vital para poder reducir la situación de asimetría en la que nos encontramos inmersos hoy en día: asimetría con respecto al acceso a los datos, y aún más importante con respecto al acceso al conocimiento experto para saber qué hacer con dichos datos.

Comparto en este sentido las palabras de Marie Curie: “Nada en la vida debería temerse, sino entenderse. Ahora es momento de entender más para así temer menos.”

– **La E también es de bEneficence o bEneficiencia**, es decir, que hace referencia a la necesidad de maximizar el impacto positivo del uso de la inteligencia artificial, con **sostenibilidad, diversidad, honestidad y veracidad**.

Porque no olvidemos que no todo desarrollo tecnológico es progreso. Y a lo que deberíamos aspirar y en lo que deberíamos invertir es en el **progreso**. Por supuesto, para ello hemos de definir qué entendemos por progreso. Desde mi punto de vista, el progreso implica una mejora en la calidad de vida de las personas, del resto de seres vivos y de nuestro planeta. En este sentido, me gustaría destacar la gran iniciativa de esta universidad llamada **Planetary Wellbeing**, que está totalmente alineada con esta visión. Su objetivo es invertir en investigación **transversal** –tan necesaria, rompiendo silos– orientada a contribuir a la consecución de los 17 Objetivos de Desarrollo Sostenible... Transmito mi más sincera y efusiva enhorabuena a la UPF por el liderazgo y por la sabia decisión de invertir en un proyecto cuyo objetivo es el progreso. Ojalá esta iniciativa sirva de inspiración para otros muchos proyectos similares.



**-Y la E también es de equality o igualdad.** El espíritu de solidaridad y la igualdad se están diluyendo con el desarrollo tecnológico.

El desarrollo y crecimiento de Internet y del World Wide Web durante la Tercera y Cuarta Revoluciones Industriales han sido sin duda clave para la **democratización** del acceso al conocimiento. Sin embargo, los principios de universalización del conocimiento y democratización del acceso a la tecnología están siendo cuestionados hoy en día por la situación de **dominancia** extrema de las grandes empresas tecnológicas norteamericanas (Apple, Amazon, Microsoft, Facebook y Alphabet/Google) y chinas (Tencent, Alibaba, Baidu), que se conoce como el fenómeno “winner takes all” (el ganador se lo lleva todo). Juntos, estos gigantes tecnológicos tienen un valor de mercado de más de cinco millones de millones (trillones americanos) de euros y unas cuotas de mercado en Estados Unidos de más de un 90% en las búsquedas de Internet (Google), de un 70% de las redes sociales (Facebook) o de un 50% del comercio electrónico (Amazon).

De hecho, el siglo **xxi** se caracteriza por una polarización en la acumulación de la riqueza. Según un estudio reciente de Crédit Suisse, el 1% más rico del planeta posee la mitad de la riqueza mundial, y las 100 personas más ricas del mundo tienen más que los 4.000 millones de personas más pobres del mundo. Esta acumulación de la riqueza en tan pocas personas ha sido atribuida parcialmente al desarrollo tecnológico y a la Cuarta Revolución Industrial.

Con la Revolución Agraria en el neolítico y durante miles de años, la propiedad de la tierra conllevaba riqueza. A partir de la Primera Revolución Industrial, la riqueza pasó a estar ligada a la propiedad de las fábricas y las máquinas. Hoy en día, podríamos argumentar que los **datos** y la capacidad para sacar partido de los mismos son el activo que más riqueza genera, dando lugar a lo que se conoce como la “economía de los datos”.

No podemos olvidar que de los cinco “países” más poblados del mundo (Facebook, WhatsApp, China, India e Instagram), tres son de Facebook. Países digitales, globales, con menos de quince años de existencia, con miles de millones de ciudadanos y que son gobernados por un presidente no elegido democráticamente. En consecuencia, un elevado porcentaje de los **datos** sobre el comportamiento humano existentes hoy en día son datos privados, captados, analizados y explotados por estas grandes empresas tecnológicas que conocen no solo nuestros hábitos, necesidades, intereses o relaciones sociales,

sino también nuestra orientación sexual o política, nuestros niveles de felicidad, de educación o de salud mental.

Por tanto, si queremos maximizar el impacto positivo del desarrollo tecnológico y, en particular, de la inteligencia artificial en la sociedad, dado que dicha inteligencia necesita de datos para poder aprender, deberíamos plantearnos **nuevos modelos de propiedad, gestión y regulación de los datos**. La Regulación General Europea para la Protección de Datos (RGDP) es un ejemplo en esta dirección. Sin embargo, la complejidad en su implementación y aplicación práctica pone de manifiesto la dificultad para definir e implementar el concepto de “propiedad” cuando hablamos de un bien intangible, distribuido, variado, creciente, dinámico y replicable infinitas veces a coste prácticamente cero.

- **Y finalmente la N es de *non-maleficence* o no maleficiencia**, es decir, que hace referencia a la necesidad de minimizar el impacto negativo que pueda derivarse del uso de las decisiones algorítmicas. Para ello, es importante aplicar un principio de **prudencia**, garantizar la **seguridad, fiabilidad y reproducibilidad** de los sistemas, preservando siempre la **privacidad** de las personas.

Será solamente cuando respetemos estos requisitos que seremos capaces de avanzar y conseguir uno de mis sueños: **un modelo de gobernanza democrática basado en los datos y en la inteligencia artificial, por y para las personas.**

## **Impacto económico y laboral**

A pesar de estas limitaciones, la presencia de la IA en nuestras vidas y su impacto económico son innegables.

Según PwC, el mercado mundial de la inteligencia artificial superará los **15 millones de millones de euros** a nivel mundial en 2030. En España, se estima que la adopción de la IA conllevaría un aumento de nuestro PIB en **0,8 puntos** en los próximos años.

Asimismo, el Foro Económico Mundial predice una transformación **profunda del mercado laboral** como resultado de la Cuarta Revolución Industrial y el desarrollo de la IA. En consecuencia, habrá tanto una destrucción como una creación de puestos de trabajo, resultando en una generación neta de **58 millones de puestos de trabajo en 2022**. La clave es que dichos puestos serán de naturaleza muy distinta a los puestos que se verán desplazados por la IA, de manera que es crítico que invirtamos en la formación a los profesionales, sobre todo a aquellos cuyas profesiones se van a ver afectadas por la IA.

## La IA como pilar estratégico

Las grandes potencias mundiales –tanto empresas como gobiernos– han comprendido que el liderazgo en la IA conllevará un liderazgo no solo a nivel económico, sino también político y social. De hecho, en los últimos dos o tres años los gobiernos de una veintena de países –incluyendo a Estados Unidos, China, Canadá, Francia, Singapur, México, Suecia, India, Australia o Finlandia– han elaborado estrategias nacionales de IA. Los líderes mundiales en IA hoy en día son China y Estados Unidos.

A nivel europeo, la Comisión Europea publicó una comunicación sobre la inteligencia artificial en **abril de 2018** y otra en **diciembre de 2018** que incluían un plan de coordinación sobre IA y la exigencia a todos los estados miembros de que publicasen sus estrategias nacionales de IA antes de julio de 2019. En **junio de 2019** el grupo de expertos sobre IA creado por la Comisión Europea –del que soy miembro reserva– publicó un conjunto de guías éticas para la inteligencia artificial y unas recomendaciones sobre inversiones y políticas en IA. La Comisión Europea estima una inversión total –tanto pública como privada– en Europa en I+D dedicada a la IA de **20.000 millones de euros anuales** en el periodo **2021-2027**. Esta inversión es necesaria para reducir la brecha de inversión en IA existente en Europa –de entre **3.000 y 15.000** millones de euros, en comparación a la inversión en China y Norte América.

Desde una perspectiva de los estados miembros, numerosos países europeos han publicado sus estrategias nacionales de IA, acompañadas de ambiciosos compromisos presupuestarios: Alemania ha comprometido **500** millones de euros anuales entre 2019 y 2025, Finlandia **100** millones de euros anuales a partir de 2019 y Francia **500** millones de euros anuales en los próximos cinco años.

Además, algunos gobiernos como el británico han creado sus propias organizaciones de Inteligencia Artificial (Office of AI, Office of Ethics in AI), con carácter transversal para maximizar el impacto de la IA y acelerar su desarrollo y adopción.

A nivel estatal, he sido miembro portavoz de un grupo de nueve expertos multidisciplinares para la elaboración del Libro Blanco de Inteligencia Artificial. Debido al cambio de gobierno, la estrategia estatal todavía no ha sido publicada, pero en junio de este año se publicó la estrategia de I+D+i y se ha creado un grupo interminis-

terial sobre IA cuyo cometido es publicar dicha estrategia en los próximos meses. Desde un punto de vista presupuestario, y tomando como referencia estrategias de países de nuestro entorno, sería razonable esperar un compromiso para el desarrollo de la IA a nivel estatal de **250 millones de euros anuales** durante los próximos cinco a siete años.

La estrategia catalana, llamada CATALONIA.IA, fue publicada en julio de este año. Destaca por su alineamiento con la visión europea para el desarrollo de la IA, su ambición para convertir a Catalunya en un polo internacional de IA y su foco en fomentar el desarrollo de una inteligencia artificial centrada en las personas.

Porque el impacto de la IA no estará necesariamente distribuido de manera homogénea o justa en la sociedad como consecuencia de las limitaciones descritas anteriormente. Por ello, deberíamos abordar con **urgencia** la elaboración también de una **estrategia estatal** de IA para asegurarnos de que su desarrollo e implantación tengan un impacto social positivo: una inteligencia artificial por y para la sociedad.

Tal estrategia de IA debería incluir un claro plan de acción en seis dimensiones fundamentales, relacionadas entre sí, muchas de las cuales están captadas también en la estrategia catalana:

- **Tecnológica**, para asegurar que investigamos, inventamos y desarrollamos tecnología puntera en IA, incluyendo la disponibilidad de las infraestructuras y datos necesarios para hacerlo.
- **Legal/regulatoria**, para actualizar nuestro marco legal de manera que contemple el uso ubicuo de sistemas de IA, siempre centrados en las personas.
- **Ética**, yendo más allá de la garantía de los derechos humanos fundamentales e incluyendo los principios que he destacado anteriormente: la justicia/no-discriminación, la transparencia, la benevolencia, la sostenibilidad, la diversidad y la no-malevolencia, entre otros.
- **Capital humano**, invirtiendo en educación a todos los niveles: educación obligatoria (incluyendo una asignatura troncal de Pensamiento Computacional, así como reforzando la creatividad y las inteligencias social y emocional de los estudiantes); educación a

profesionales –sobre todo aquellos cuyas profesiones se van a ver afectadas por la IA–; educación a la ciudadanía y educación a la clase política y miembros de la administración pública.

- **Social**, incorporando la IA en las administraciones públicas para, por ejemplo, mejorar la eficiencia de los procesos internos, proporcionar un servicio personalizado y permanente a los ciudadanos y optimizar la toma de decisiones basadas en la evidencia.
- **Económica y laboral**, asegurándose de que nuestra economía consiga crecer de manera sostenible e igualitaria gracias al desarrollo de la IA.

Solamente con una apuesta ambiciosa, holística, plurianual, consensuada por una mayoría política a través de un pacto de Estado, avalada por la ciudadanía e impulsada por un organismo transversal dentro del Gobierno, conseguiremos alcanzar el inmenso impacto positivo que la inteligencia artificial puede tener en la sociedad.

Podemos jugar un papel relevante en Europa, actuando como puente con Latinoamérica y África. Para ello, deberíamos reforzar nuestra **excelencia científica** en esta área, **invirtiendo** significativamente en la formación, la atracción y la retención del **talento** –con especial énfasis en el femenino, dado el escasísimo porcentaje de mujeres expertas en IA hoy en día–, en la actualización de nuestro sistema educativo y de nuestras **administraciones públicas**, así como en el fomento de la **creatividad**, el **emprendimiento** y la **innovación** en IA y áreas relacionadas con ella; elevando no solo el crecimiento empresarial y económico, sino sobre todo **nuestra calidad de vida** y la de nuestro **planeta**. Por ello, espero y deseo que no dejemos escapar esta oportunidad. Está en nuestras manos.

Gràcies, i feliç curs 2019-2020!







**Universitat  
Pompeu Fabra**  
*Barcelona*