



**Santander**  
Fundación

# la colección de ideas

IA & EL PENSAMIENTO CRÍTICO



# INTELIGENCIA ARTIFICIAL & PENSAMIENTO CRÍTICO

El 28 de abril de 2023 se celebró en la sede de la Fundación Banco Santander la sesión "**Inteligencia Artificial y pensamiento crítico**".

Esta sesión se enmarca dentro del foro de debate '**La Colección de Ideas**'. Una iniciativa que la Fundación está llevando a cabo con el objetivo de debatir sobre cuestiones de actualidad en un entorno que permita profundizar en el conocimiento, la reflexión informada y el intercambio de ideas y propuestas.

La primera sesión, que contó con una intervención inicial de Nuria Oliver, Directora de la Fundación ELLIS Alicante, conocida como

el Instituto de inteligencia Artificial centrada en la Humanidad. Asimismo, se consiguió poner en común sobre una misma mesa de diálogo a diferentes expertos<sup>1</sup> para reflexionar sobre los principales retos y oportunidades que surgen en torno al tema propuesto: Inteligencia Artificial y pensamiento crítico.

El presente está elaborado a partir de las intervenciones de la ponente y de los diferentes invitados. Gran parte de las aportaciones de la ponente pueden encontrarse en su libro "[Inteligencia Artificial, naturalmente](#)" publicado por el ONTSI en el Ministerio de Economía y Sociedad Digital.



Nos encontramos en la Cuarta Revolución Industrial, que representa una unión íntima, sin precedentes, entre el mundo físico, el mundo biológico y el mundo digital. Una Revolución impulsada por disciplinas como la biotecnología, la nanotecnología la ingeniería genética y la Inteligencia Artificial (IA).

En concreto, la Inteligencia Artificial existe desde los años 50 e históricamente ha habido dos grandes escuelas de pensamiento en torno a ella.

La primera de ellas, llamada la **escuela Top-Down o simbólico-lógica**. Los métodos en esta escuela se basan en representaciones simbólicas del conocimiento y el uso de las reglas de la lógica para derivar conocimiento nuevo a partir de un cierto conocimiento base. Este enfoque de la IA fue el paradigma dominante desde el nacimiento de la disciplina en la década de los 50 hasta finales de los años 1980. El ejemplo canónico de la escuela de arriba hacia abajo son los sistemas expertos.

La segunda escuela de referencia en el campo de la IA es la **Bottom-Up o la escuela conexionista**. Sus bases están inspiradas en la biología: los seres biológicos inteligentes aprenden de sus interacciones con su entorno, de la experiencia. Por lo tanto, este enfoque de la IA se centra en el desarrollo de métodos

que aprenden a partir de datos en lugar de modelar descripciones simbólicas del entorno. El ejemplo canónico de método en esta escuela de la IA son las redes neuronales. Esta escuela empieza a tomar fuerza en los años 90 y desde entonces ha vivido un crecimiento exponencial, suponiendo un decrecimiento de la escuela clásica o Top Down.

Este crecimiento se ha debido a la confluencia de tres factores. En primer lugar, la disponibilidad de cantidades ingentes de datos, datos que son fruto de la ubicuidad de los dispositivos y de todas las huellas digitales que vamos dejando con los servicios digitales, además de la digitalización del mundo físico a través de sensores.

En segundo lugar, la disponibilidad de grandes cantidades de computación a bajo coste, que nos permiten procesar todos esos datos. Y en tercer lugar el desarrollo de modelos de aprendizaje, que están basados en las redes neuronales de los años 40 y 50, pero muchísimo más complejas, que se conocen como redes neuronales profundas, y que



entrenadas con estas grandísimas cantidades de datos y capacidades de computación están consiguiendo hitos que eran inimaginables hace 15 o 20 años.

Actualmente, vivimos rodeados de algoritmos de IA, que inciden en la toma de decisiones diarias tanto a nivel consumo como a nivel relacional. Por ello, la IA ha dejado de ser un tema estrictamente tecnológico para pasar a ser un tema político, que tiene que ser abordado desde diferentes dimensiones: La dimensión tecnológica, la dimensión legal y regulatoria, la dimensión ética, la dimensión social incluyendo la necesaria inversión en educación y la dimensión económica y laboral.

Estamos hablando de ciencia y tecnología, pero también de derechos, de economía, de democracia, de igualdad, de inclusión, de ciudadanía, de paz y de poder.

Desde la dimensión social, es importante resaltar que la IA se plantea como una herramienta esencial para afrontar los retos a los que nos enfrentamos como sociedad, como el cambio climático, el reto demográfico o la desigualdad de género. En el informe *Data Revolution Report*, se expone como la IA puede, por un lado, ayudarnos a medir mejor el nivel de contribución a los Objetivos de Desarrollo Sostenible, identificando debilidades y áreas de mejora; y, por otro lado, permitir y acelerar el logro de la Agenda 2030.

Asimismo, en el artículo *"Artificial Intelligence for Social Good: the Way Forward"* de Nuria Oliver, se detalla cómo la IA puede contribuir a cada uno de los ODS, haciendo un llamamiento a la necesidad de invertir en una IA que contribuya al progreso, sin dejar a nadie atrás y sin destruir el planeta en el proceso.

En este sentido, es importante hacer referencia a los debates éticos que se derivan de la IA, mitigando los retos que plantea para que su impacto sea positivo.





Uno de los principales retos que plantea es la violación **computacional de la privacidad**. Es decir, poder atribuir características personales a partir de datos no personales utilizando técnicas de IA. Un ejemplo de ello se demostró en un estudio de la Universidad de Stanford liderado por Yilun Wang y Michael Kosinski. Las conclusiones de este estudio probaron que la IA puede determinar si alguien es heterosexual u homosexual a través de sus rasgos faciales con mayor precisión que una persona. Algo que hizo saltar las alarmas de diferentes organizaciones en defensa de los derechos LGTBI por la invasión en la intimidad que supone la posibilidad de detectar con una máquina la orientación sexual.

Otro gran reto que plantean los sistemas de IA es el de los sesgos y la discriminación algorítmica. Los algoritmos han sido entrenados con datos que reflejan una realidad sesgada. Según un estudio de la Universidad de Virginia, un algoritmo de IA aprendió a vincular mujeres con imágenes de cocinas, ya que, en internet, debido a los estereotipos, aparecen más imágenes de mujeres que hombres fotografiados en cocinas. Esto no solo replica, si


no que amplifica los estereotipos. Algo que también sucede a nivel racial, generando posibles discriminaciones a la hora de acceder a un seguro médico o evaluar el riesgo crediticio de las personas, ya que los algoritmos funcionan mejor con caras de hombres blancos.

Además, con la IA generativa se acrecienta el problema. Estos modelos de IA no solo detectan patrones, si no que generan contenido, tanto en formato texto, como conocido ChatGPT o de imagen como Dall-E2.

Otro gran reto que plantea la IA es la asimetría. Europa se encuentra en una situación de debilidad y vulnerabilidad en investigación, innovación, uso y despliegue de IA frente a Norte América y China.

Un problema que aparece bien reflejado en el libro “Europa frente a EE.UU. y China. Prevenir el declive en la era de la inteligencia artificial” coescrito por Andrés Pedreño, uno de los asistentes a la sesión.

Esta asimetría aparece reflejada en el acceso a los datos, que se pueden utilizar para manipular el comportamiento humano; pero



**“No contamos con una sociedad preparada para poder contribuir y entender los retos a los que nos enfrentamos”**  
**Nuria Oliver**



también en lo que respecta a las capacidades. No contamos con una sociedad preparada para poder contribuir y entender los retos tecnológicos, especialmente vinculado a la Inteligencia Artificial, a los que nos enfrentamos.

Estados Unidos y China cuentan con una gran estrategia digital, mientras que Europa no cuenta con una estrategia sólida en este ámbito. Es más, en Europa no tenemos gigantes tecnológicos y a las startups les cuesta un trabajo enorme escalar.

La **falta de transparencia** de los sistemas actuales de IA también supone uno de los grandes retos y se puede manifestar de 3 formas:

Opacidad intencional: el objetivo de los diseñadores de estos programas es la protección de la propiedad intelectual de los inventores.

Opacidad de conocimiento: la gran mayoría de las personas carecen de las habilidades técnicas para entender cómo funcionan los modelos computacionales construidos en base a los algoritmos.

Opacidad intrínseca: surge por la complejidad de ciertos métodos de aprendizaje por ordenador, como los modelos de *deep learning*.

La **falta de veracidad** de estos sistemas también es uno de los principales retos que plantea la IA. Los modelos generativos pueden crear imágenes, textos, videos o música y son realmente indistinguibles de un contenido veraz. Además, a través de la generación de contenido irreal tienen la capacidad de definir la opinión pública en temas de gran relevancia a nivel social.

Asimismo, la capacidad exponencial de generar contenidos pone en jaque la creatividad y la protección de los autores.

Por otro, también es importante atender a la falta de diversidad de la IA. Los algoritmos intentan modelar lo que nos interesa, lo que nos gusta y entonces tienden a encasillarnos en ciertos estereotipos, de manera que nos



conectan con personas o con contenidos que son afines a nuestras creencias.

Esto da lugar a lo que se conoce como filtros burbuja y cámaras de resonancia y que están contribuyendo a la polarización de la sociedad.

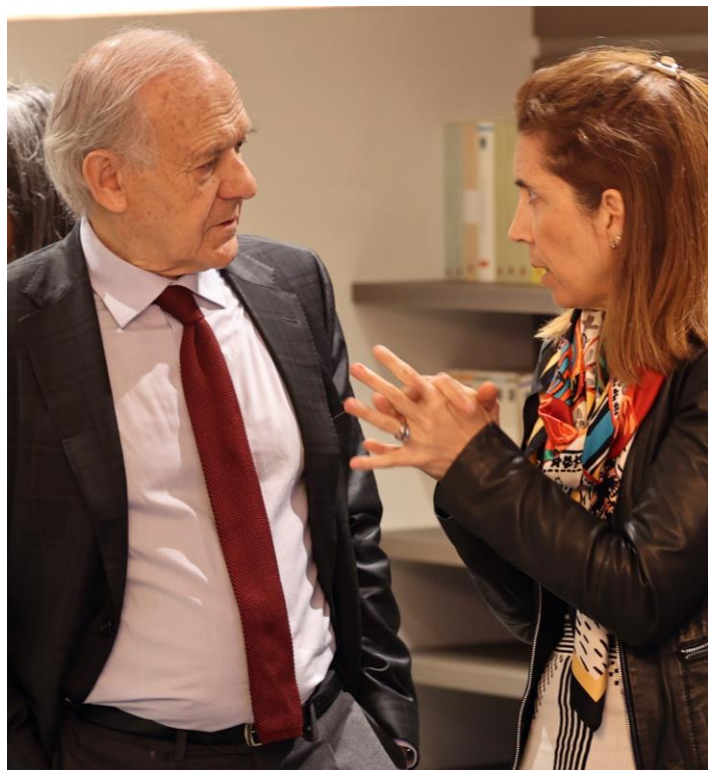
La **vulnerabilidad** de estos sistemas es otro de los retos a los que hay que atender, ya que son sistemas a los que se puede engañar. De hecho, existe una disciplina que es la Inteligencia Artificial Adversarial y cuyo único objetivo es engañar a sistemas de IA. Cada vez más estamos integrando más sistemas de IA en nuestro día a día, y es importante conocer las debilidades que presenta.: Un ejemplo de esto es el casi fallido del Autopilot de Tesla, donde el vehículo autónomo fue incapaz de detectar un camión blanco que estaba en mitad de la carretera, causando la muerte del conductor.

Cuando hablamos de los retos que plantea esta tecnología, no podemos olvidar **su impacto medioambiental**. Si bien es cierto que la IA puede suponer una herramienta clave para proteger el medio ambiente, su uso viene acompañado de un alto consumo de energía, con sus emisiones de CO2 respectivas.

La mayor parte del consumo proviene de los centros de procesamientos de datos. Según la

Escuela Superior de Tecnología de la Universidad de Quebec (Canadá), el incremento exponencial de la computación genera el 4% de los gases de efecto invernadero producidos por la actividad humana. Asimismo, la extracción de metales como el litio, utilizado para fabricar la batería de la mayoría de los dispositivos electrónicos que utilizamos, provoca la degradación del suelo, contamina el agua y tiene efectos negativos en la biodiversidad. Además, en ocasiones, su extracción supone la vulneración a los derechos humanos y la explotación laboral.

Por eso, es cada vez es más importante que el desarrollo tecnológico esté alineado con la responsabilidad humana de garantizar las condiciones básicas para la vida en nuestro planeta. En este contexto, se está desarrollando lo que se conoce como IA Verde y que tiene por objetivo avanzar en la creación de modelos que disminuyan el daño a la naturaleza en su etapa de entrenamiento.



También es importante incorporar al debate cuestiones como el **acceso de los jóvenes a la IA**. Según el informe [El impacto de la tecnología en la adolescencia](#) de Unicef, uno de cada tres adolescentes españoles hace un



uso problemático de las redes sociales. Se estima que más de 600.000 estudiantes de ESO podrían presentar un uso problemático, que está caracterizado no sólo por una elevada frecuencia e intensidad de conexión, sino por un alto grado de interferencia en el día a día de los propios adolescentes y de sus familias.

Asimismo, la edad media en España de acceso a dispositivos móviles con internet es de 10,96 años, mientras que el 98,5% de los adolescentes está registrado en, al menos, una red social.

En este sentido, es necesario construir una sociedad consciente, donde nuestros jóvenes reciban formación en **pensamiento computacional** para que puedan tomar decisiones informadas, que reduzcan su riesgo de exposición y el uso problemático de la tecnología.

El pensamiento computacional se rige por diferentes competencias, que consisten en pensamiento algorítmico, programación, conocimiento y análisis de datos y redes y hardware, que es el sustrato físico de la tecnología.

Por último, como se ha mencionado con anterioridad, la **regulación** es uno de los aspectos clave cuando hablamos de IA. En general, la regulación europea se centra en la tecnología en sí misma y no en los usos de esta. Sin embargo, además de definir lo que es la IA y su alcance, es importante entender las consecuencias de este tipo de tecnología y como mitigar el impacto negativo que deriva de su uso.

La Unión Europea está estudiando un nuevo marco jurídico que pretende reforzar significativamente la normativa sobre el desarrollo y el uso de la IA.







# Conclusiones

- Existe consenso en torno al potencial que tiene la IA para contribuir al progreso de la economía, de las personas y del planeta. Sin embargo, es importante atender a los retos que plantea para maximizar su impacto positivo. Este artículo reciente de The Conversation puede ser de interés y relevancia: [Inteligencia artificial: riesgos reales frente a amenazas hipotéticas](#).
- La tecnología y su futuro están en nuestras manos y es importante atender a cuestiones como la violación computacional de la privacidad, los sesgos y la discriminación algorítmica, la falta de diversidad, transparencia y veracidad, la vulnerabilidad de los sistemas, la asimetría en comparación con potencias como China y EEUU, su gran impacto ambiental y el uso problemático de la tecnología por parte de la juventud.
- La IA ha dejado de ser un tema estrictamente tecnológico para pasar a ser un tema político, que tiene que ser abordado desde diferentes dimensiones: La dimensión tecnológica, la dimensión legal y regulatoria, la dimensión ética, la dimensión social incluyendo la necesaria inversión en educación y la dimensión económica y laboral.
- La **regulación** es uno de los aspectos clave cuando hablamos de IA. Además, de definir lo que es la IA y su alcance, es importante que atienda a las consecuencias de este tipo de tecnología y establezca criterios para mitigar el impacto negativo que deriva de su uso.

# Anexo I.

## Asistentes

### Ponente:

#### **Nuria Oliver**

Directora de la Fundación ELLIS Alicante, dedicada a la investigación en "Inteligencia artificial centrada en el ser humano". Es ingeniera de Telecomunicaciones por la Universidad Politécnica de Madrid y doctora en Inteligencia Artificial por el Instituto Tecnológico de Massachusetts (MIT).

### Moderación:

#### **José Ignacio Torreblanca**

Profesor titular de la Universidad Nacional de Educación a Distancia (UNED) Y Doctor Miembro del Instituto Juan March de Estudios e Investigaciones. Ha sido becario del Programa Fulbright, Unión Europea- Estados Unidos, Profesor en la George Washington University (D.C.) así como investigador posdoctoral en el Instituto Universitario Europeo de Florencia (EUI). También ha sido Investigador Principal, Área Europa, en el Real Instituto Elcano de Estudios Internacionales (2004-2007) y miembro del Consejo Editorial y Director de Opinión del Diario ELPAIS (2016-2018). Desde 2008 es director de la oficina en Madrid del European Council on Foreign Relations (ECFR).

### Invitados:

#### **Borja Baselga**

Licenciado en Ciencias Económicas y Empresariales por CUNEF, ha realizado estudios de postgrado en el I.E. Business School, en INSEAD y en Harvard Business School. Durante los últimos 33 años ha trabajado en el Grupo Santander ocupando diferentes puestos de responsabilidad en las áreas de Banca Corporativa, Banca de Empresas y Banca de Particulares. Ha sido Director de Responsabilidad Social Corporativa del Grupo Santander, Presidente de la Asociación Española del Pacto Mundial de Naciones Unidas y ha formado parte de UNEP Fi, la iniciativa de Naciones Unidas para el medioambiente, así como del grupo de alto nivel de Naciones Unidas para los microcréditos. Desde 2009 es Director Gerente de la Fundación Banco Santander y es patrono de numerosas Fundaciones dedicadas a la investigación, el arte y la acción social.

#### **Olga Grau**

Directora de comunicación con stakeholders del Banco Santander. Cursó sus estudios de periodismo en la Universidad Autónoma de Barcelona y es PDD por el IESE. Grau desarrolló su carrera profesional en varios medios, como Dossier Econòmic, Onda-Cero o El Periódico. También formó parte de las delegaciones catalanas de Expansión y El Mundo. Antes de su actual puesto en el Banco Santander, formó parte del Grupo Zeta, donde ocupaba la jefatura de Economía y era adjunta a la dirección.

# Anexo I.

## Asistentes

### Invitados:

#### **Lucía Ortiz de Zárate**

Investigadora en Ética y Gobernanza de la Inteligencia Artificial (IA) en la Universidad Autónoma de Madrid (UAM) e investigadora asociada en el Círculo de Bellas Artes de Madrid. Forma parte de la segunda edición de la lista «35 under 35» Future Leaders elaborada por el Foro CIDOB-Santander y del grupo de investigación GHECO. Visiting fellow 2023/2024 en el Programa de Ciencia, Tecnología y Sociedad de la Harvard Kennedy School.

#### **Andrés Pedreño**

Catedrático de Economía Aplicada, Rector de la Universidad de Alicante (1993-2000), CEO de Universia-Banco Santander (2000-2004), Fundador de Torre Juana OST Campus tecnológico especializado en Inteligencia Artificial, Director del Observatorio para el Análisis y Desarrollo Económico de Internet en España (ADEI) ADEI, apoyado entre otros por Google en España. Fundador de AlicanTEC asociación sin ánimo de lucro para el desarrollo de la economía digital y los sectores de futuro. Director del Observatorio de IA, Miembro del Comité de Expertos del Gobierno para la Elaboración del Libro Blanco sobre la Inteligencia Artificial y el Big Data (2017-2019); coautor del libro "Europa frente a EEUU y China. Prevenir el declive en la era de la Inteligencia Artificial".

#### **Tomás de la Quadra-Salcedo**

Abogado y doctor en Derecho de la Universidad Complutense de Madrid. Ha sido presidente del Consejo de Estado, ministro de Justicia del Gobierno español y ministro de Administración Territorial. Profesor emérito de la Universidad Carlos III de Madrid. Se ha desempeñado como catedrático de Derecho Administrativo de la Facultad de Ciencias Sociales y Jurídicas de la Universidad Carlos III de Madrid y director de la Escuela Doctoral de esa misma Facultad. Profesor invitado en distintas universidades españolas e iberoamericanas. Coordinador de la Carta de Derechos Digitales.

#### **Borja Sémpser**

Político y escritor. Fue portavoz del Partido Popular en el Parlamento Vasco y, de 2009 a 2020, presidente del Partido Popular de Guipúzcoa. Desde 2020 hasta 2023 ha sido director asociado del área de Relaciones Institucionales de la consultora EY. Actualmente, es el Portavoz del Partido Popular para las elecciones de 2023.

#### **Paloma Villa**

Head of Digital Public Policy en Telefónica. Es licenciada en Ciencias Políticas y especializada en Relaciones Internacionales por la Universidad Complutense de Madrid. Cursó estudios en European Studies por la RWTH-Aachen y un posgrado en 'Managing a Digital Telco' en la Universidad Politécnica de Barcelona. Desarrolló su carrera profesional en diferentes instituciones europeas, como la Comisión Europea o el Parlamento Europeo, donde trabajó como asesora del grupo socialista. Desde 2012 trabaja para Telefónica desarrollando posicionamiento público y representando a la compañía en foros internacionales, como la OCDE, en materia de telecomunicaciones y políticas digitales.



## Colección de Ideas 1 - 2023

**Producción:** Fundación Banco Santander

**Fotografía:** Bonus Studio – Productora Creativa Madrid



# la colección de ideas