

Inteligencia artificial, poder y sociedad: aprendizajes de dos jornadas

Asier Gallastegi Octubre 2026

Índice

[Índice](#)

[Introducción](#)

[Notas de Asier Gallastegi de ambos encuentros](#)

Inteligencia Artificial: sesgos y retos sociales en la sociedad vasca 19 de octubre de 2026 · Tabakalera, Donostia / San Sebastián Jornada organizada conjuntamente por Orai NLP Teknologiak (Elhuyar) e Ikuspegi (EHU), e impulsada por KSIgune.

1. Sesgos humanos y sesgos de la inteligencia artificial: el atractivo físico como factor de discriminación

- Tres grandes ámbitos de investigación
- Los sesgos cognitivos que trasladamos a las máquinas
- El sesgo del atractivo y el lookismo algorítmico
- Una investigación sobre los filtros de belleza
- La interacción con otros sesgos sociales
- El atractivo en la generación de imágenes
- El impacto sobre los sistemas de reconocimiento de género
- Modelos cada vez más realistas, pero no necesariamente menos sesgados
- ¿Por qué aprenden estos sesgos?
- La homogeneización de nuestra representación de la realidad

2. Cómo cambian los sesgos según las tareas que pedimos a la inteligencia artificial

- La inteligencia artificial no es neutra: también importa qué le pedimos
- Una misma situación puede generar respuestas muy diferentes
- Cómo se construye una investigación sobre sesgos
- Comparación entre 16 modelos
- El tipo de tarea modifica los resultados

3. Sesgos y retos sociales de la inteligencia artificial en la sociedad vasca

- Los sesgos en el ámbito de la salud
- Servicios Sociales de Barcelona: detenerse a pensar dónde y para qué utilizar IA
- El reto no es únicamente tecnológico, sino organizativo
- La inteligencia artificial y los cambios en el periodismo
- Desinformación y centralización de los relatos
- Sobre los sesgos técnicos y las garantías jurídicas
- El contexto como elemento fundamental
- El riesgo de confiar demasiado en las recomendaciones automatizadas
- Educación, alfabetización mediática y pensamiento crítico
- Retos compartidos por los diferentes ámbitos

4. Tendencias tecnológicas, evaluación de sesgos y desarrollo de modelos propios

Sabemos que hay sesgos, pero todavía nos faltan herramientas para comprenderlos

¿Qué significa eliminar los sesgos?

Sesgos que aparecen en las traducciones

Auditorías y nuevas formas de evaluación

¿Es posible construir una inteligencia artificial sin sesgos?

Modelos abiertos, regulación y autonomía tecnológica

El euskera y la necesidad de desarrollar tecnología propia

La investigación y el apoyo institucional

Notas del Seminario Internacional «Tecnología, poder y ciudadanía global: ¿Qué futuro nos depara la Inteligencia Artificial?» 20 de octubre de 2026 · Encuentro impulsado por Alboan y la Universidad de Deusto

1. Inteligencia artificial, poder y regulación: un sistema jurídico que parece haber quedado obsoleto

El sistema jurídico como ficción compartida

La externalización de la cognición humana

La tensión entre regulación y competitividad

Europa, Estados Unidos y China: distintos escenarios de poder

Datos de entrenamiento, agentes y transformación del trabajo

Fiscalidad, cotizaciones y sostenibilidad del sistema de bienestar

El riesgo de perder capacidades profesionales

Una sociedad preparada para regalar sus datos

2. Inteligencia artificial y sostenibilidad ambiental: hacia algoritmos verdes

La huella ambiental de la inteligencia artificial

La nube es una infraestructura física

No todas las interacciones tienen el mismo impacto

Entrenamiento e inferencia: dos momentos diferentes

Aprovechar el calor y repensar las infraestructuras

Diseñar algoritmos que contaminen menos

La inteligencia artificial también puede ayudar al planeta

¿Qué podemos hacer como personas usuarias?

3. Minerales críticos, extractivismo y resistencias desde los territorios

Introducción

Los días 19 y 20 de octubre de 2026 participé en dos jornadas que abordaron, desde perspectivas complementarias, los **retos sociales, políticos, económicos y ambientales de la inteligencia artificial**.

La primera, **«Inteligencia Artificial: sesgos y retos sociales en la sociedad vasca»**, se celebró el 19 de octubre en Tabakalera (Donostia), organizada por **Orai NLP Teknologiak (Elhuyar) e Ikuspegi (EHU), e impulsada por KSIgune**. La jornada se centró en los sesgos que reproducen los modelos de inteligencia artificial, sus consecuencias en diferentes ámbitos profesionales y sociales, y la necesidad de desarrollar capacidades tecnológicas propias, especialmente en relación con el euskera.

La segunda, **«Tecnología, poder y ciudadanía global: ¿Qué futuro nos depara la Inteligencia Artificial?»**, tuvo lugar el 20 de octubre, impulsada por **Alboan y la Universidad de Deusto**. En este encuentro se abordaron las relaciones de poder que acompañan al desarrollo tecnológico, las dificultades de regulación, la transformación del trabajo, los impactos ambientales y las consecuencias del extractivismo sobre diferentes territorios.

Ambas jornadas comparten una reflexión de fondo: la inteligencia artificial no es únicamente una cuestión tecnológica. Es una transformación social y política que está modificando nuestras relaciones con el conocimiento, el trabajo, las instituciones y el entorno. Comprender sus posibilidades exige observar también las relaciones de poder, las desigualdades y las dependencias que genera.

Notas de Asier Gallastegi de ambos encuentros

He partido de mis notas recogidas en directo y luego he pedido a una IA generativa que limpie el texto. Tras este ajuste he revisado cada texto y entresacado aquellas ideas que he podido llevar al post resumen y que en estos textos resalto en negrita.

Inteligencia Artificial: sesgos y retos sociales en la sociedad vasca 19 de octubre de 2026 · Tabakalera, Donostia / San Sebastián Jornada organizada conjuntamente por Orai NLP Teknologiak (Elhuyar) e Ikuspegi (EHU), e impulsada por KSIgune.

Estas notas recogen algunas de las ideas, investigaciones y conversaciones compartidas durante el encuentro. El recorrido permite comprender cómo los sistemas de inteligencia artificial reproducen y amplifican sesgos sociales, qué efectos pueden tener sobre ámbitos como la salud, los servicios sociales, el

periodismo o la educación, y qué retos aparecen para la investigación, la regulación y el desarrollo tecnológico en euskera.

1. Sesgos humanos y sesgos de la inteligencia artificial: el atractivo físico como factor de discriminación

Presentación vinculada a ELLIS Alicante, Instituto de Inteligencia Artificial Centrada en la Humanidad

Tres grandes ámbitos de investigación

El instituto organiza su trabajo en torno a tres grandes áreas: la inteligencia artificial que nos ayuda a entendernos, la inteligencia artificial con la que interactuamos y el desarrollo de sistemas de inteligencia artificial en los que podamos confiar.

Esta tercera dimensión aborda cuestiones como los límites éticos, los sesgos y el impacto social de estas tecnologías.

Una de las principales preocupaciones es que los grandes modelos multimodales no aprenden únicamente lo que sabemos. **También aprenden cómo pensamos, cómo juzgamos y cómo interpretamos la realidad.** Existe una evidencia creciente de que estos modelos reproducen sesgos relacionados con la edad, el sexo, el género y otras características personales y sociales.

Los sesgos cognitivos que trasladamos a las máquinas

Las personas tenemos **sesgos cognitivos: patrones sistemáticos que nos llevan a desviarnos de lo que consideraríamos una decisión estrictamente racional.**

Se recuperó como referencia el trabajo de Daniel Kahneman y su explicación de los **dos sistemas de pensamiento, uno más rápido, intuitivo y automático, y otro más lento y deliberativo.**

Se mencionó la existencia de alrededor de 200 sesgos cognitivos que influyen en nuestras decisiones cotidianas. Estos mecanismos se utilizan desde hace tiempo en ámbitos como el marketing, los casinos y las estrategias comerciales, pero también **tienen consecuencias en procesos electorales, decisiones médicas o resoluciones judiciales.**

La pregunta que orienta esta línea de investigación es si los modelos fundacionales están aprendiendo también nuestros sesgos cognitivos.

El sesgo del atractivo y el lookismo algorítmico

Uno de los fenómenos estudiados es el sesgo del atractivo físico, también denominado lookismo algorítmico.

Se trata del **tratamiento diferencial de las personas en función de su apariencia física. El atractivo condiciona los juicios que hacemos sobre otras personas y puede influir en cómo las trata la inteligencia artificial.**

Detrás de este fenómeno aparece una asociación ampliamente documentada desde los años setenta: lo que es bello es también percibido como bueno. **Tendemos a atribuir a las personas físicamente atractivas otras características positivas. Las consideramos más competentes, inteligentes, fiables o adecuadas para una promoción profesional.**

Es un atajo cognitivo. A partir de un rasgo visible, como la belleza, inferimos otras cualidades para las que no tenemos información suficiente.

Este mecanismo puede tener efectos en las oportunidades laborales, los salarios, las calificaciones educativas y las decisiones judiciales. Durante la ponencia se mencionaron diferencias salariales que, en determinados estudios y condiciones, pueden alcanzar el 20 %.

La cuestión es hasta qué punto estos patrones de juicio humano están siendo incorporados y reproducidos por los modelos de inteligencia artificial.

Una investigación sobre los filtros de belleza

Se presentó un estudio que analizaba cómo cambian nuestras percepciones cuando observamos fotografías de personas a las que se han aplicado filtros de belleza mediante inteligencia artificial.

La investigación trabajó con imágenes de 462 personas. Se compararon las fotografías originales con las mismas imágenes modificadas mediante filtros de embellecimiento.

Las personas participantes evaluaban características como el atractivo, la felicidad o la confiabilidad.

Los resultados mostraron que las imágenes modificadas no solo eran percibidas como más atractivas. También obtenían puntuaciones superiores en otras características positivas. En algunos casos, las diferencias alcanzaban entre dos y tres puntos.

El estudio, titulado What Is Beautiful Is Still Good: The Attractiveness Halo Effect in the Era of Beauty Filters, muestra que el efecto halo asociado a la belleza continúa funcionando en un entorno en el que la apariencia puede modificarse digitalmente de forma masiva.

Las preguntas se repetían varias veces, alterando también el orden de las respuestas, para obtener medias y comparar los resultados.

Además se diseñaron 91 escenarios que incluían profesiones estereotipadas, atributos personales, condiciones económicas y localizaciones geográficas, entre otras variables.

Los resultados fueron significativos:

- En el 84 % de los escenarios, los modelos ofrecieron respuestas diferentes en función de si la fotografía tenía aplicado un filtro de belleza.
- En el 93 % de los casos analizados, las imágenes embellecidas se asociaban con características más positivas.
- Entre los modelos estudiados, DeepSeek aparecía como uno de los más afectados por este sesgo.

Los resultados muestran que los modelos no son neutros ante la apariencia física. El atractivo modifica sus respuestas incluso cuando las preguntas se refieren a características que no pueden inferirse de manera fiable a partir de una fotografía.

La interacción con otros sesgos sociales

El sesgo del atractivo no opera de manera aislada.

La investigación también permitió observar su relación con otras variables sociales, especialmente el género, la edad y los estereotipos raciales.

El género aparecía como una de las variables más estereotipadas. En las profesiones tradicionalmente asociadas a hombres o mujeres, el efecto era todavía mayor.

También se identificaron diferencias en la atribución de estudios o competencias según el origen racial percibido. En determinados escenarios, las personas negras tenían menos posibilidades de ser identificadas como personas con estudios.

Los filtros de belleza, además, podían modificar la percepción de edad y la racialización percibidas de las personas. Las imágenes embellecidas tendían a parecer más jóvenes y, en algunos casos, presentaban rasgos interpretados como más "blancos".

Esto significa que un mismo filtro puede alterar simultáneamente diferentes características percibidas y activar varios sesgos. No estamos únicamente ante un sesgo estético, sino ante una interacción entre apariencia, género, edad, origen racial y posición social.

El atractivo en la generación de imágenes

La investigación también exploró cómo funcionan estos sesgos cuando los modelos no analizan una fotografía existente, sino que generan nuevas caras a partir de descripciones.

Por ejemplo, cuando pedimos una imagen fotorrealista de una persona inteligente, pueden aparecer hombres con gafas y libros. Cuando solicitamos una persona competente, las imágenes suelen mostrar a alguien limpio, arreglado y con una apariencia considerada atractiva.

Por el contrario, atributos como la incompetencia pueden asociarse con personas desordenadas o con rasgos físicos determinados, como unas orejas grandes.

Para estudiar este fenómeno se pidió a los modelos que generaran imágenes frontales de personas combinando atributos, género y categorías raciales.

Se utilizaron características como atractivo, inteligencia, felicidad, sociabilidad y confiabilidad, junto con sus correspondientes atributos opuestos.

Los resultados mostraron una relación importante entre los atributos positivos y el atractivo físico de las caras generadas. **Cuando se pedía una persona inteligente, confiable o sociable, los modelos tendían a crear rostros que respondían a determinados cánones de belleza.**

De esta manera, la inteligencia artificial no solo reproduce estereotipos cuando interpreta imágenes; también los refuerza cuando crea nuevas representaciones visuales.

El impacto sobre los sistemas de reconocimiento de género

Otra parte del estudio analizó si el atractivo físico influía en la capacidad de los modelos para clasificar el género de las personas a partir de fotografías.

Se observaron diferencias importantes.

Algunos modelos mostraban patrones de discriminación especialmente fuertes. **Por ejemplo, determinadas imágenes de mujeres consideradas poco atractivas o**

asociadas con atributos negativos eran clasificadas como masculinas con mayor frecuencia.

Los sistemas funcionaban mucho mejor con las caras de los hombres. En determinados grupos de imágenes femeninas, los porcentajes de acierto llegaban a ser inferiores al 10 %.

Esta situación se relacionó con la idea de invisibilidad algorítmica: determinadas personas o grupos sociales quedan peor representados o son reconocidos con menor precisión por los sistemas tecnológicos.

También aparecieron diferencias entre modelos, relacionadas con los datos y procesos utilizados en su entrenamiento.

Modelos cada vez más realistas, pero no necesariamente menos sesgados

Una de las conclusiones más preocupantes es que algunas nuevas versiones de los modelos tienden a producir imágenes cada vez más jóvenes, homogéneas y ajustadas a determinados cánones estéticos.

En el caso de las mujeres, las imágenes generadas con atributos positivos podían presentar una apariencia más sexualizada y, al mismo tiempo, ser interpretadas como menos adecuadas para determinados trabajos.

La mejora del realismo visual no implica necesariamente una reducción de los estereotipos.

En algunos casos, los modelos parecen avanzar hacia una representación visual más sofisticada mientras refuerzan patrones sociales de discriminación. Las nuevas versiones generan imágenes más realistas, pero también presentan sesgos más acentuados.

Esto obliga a analizar los sesgos desde una perspectiva interseccional, observando cómo se combinan diferentes características en lugar de estudiarlas de manera independiente.

¿Por qué aprenden estos sesgos?

Los modelos han aprendido de nosotros y de los contenidos que hemos producido. **Una parte importante de los datos utilizados para entrenarlos procede de internet, que no es un reflejo fiel del mundo.**

Los contenidos digitales están atravesados por desigualdades, estereotipos y mecanismos de selección. Por ejemplo, una proporción importante de las fotografías que las mujeres publican en redes sociales utiliza filtros de belleza.

Si estos contenidos forman parte de los datos de entrenamiento, los modelos pueden aprender una representación distorsionada de la apariencia real de las personas.

Durante la ponencia también se explicó que el entrenamiento de los grandes modelos suele desarrollarse en diferentes fases.

En una primera fase, los sistemas aprenden patrones generales a partir de grandes colecciones de datos. Posteriormente, se realizan procesos de ajuste con datos más específicos para mejorar su eficacia, introducir salvaguardas y tratar de reducir determinados sesgos.

Estos procesos se han ido sofisticando, pero no eliminan necesariamente las desigualdades.

También se observaron diferencias según el idioma utilizado para interactuar con los modelos multilingües. Un mismo sistema puede presentar sesgos distintos dependiendo de la lengua en la que formulamos las preguntas.

Se mencionó que los modelos chinos y occidentales presentan patrones bastante similares en determinados sesgos relacionados con la belleza.

La homogeneización de nuestra representación de la realidad

El uso masivo de filtros de belleza puede contribuir a homogeneizar la apariencia de las personas.

Cuando millones de personas modifican sus fotografías siguiendo criterios estéticos similares, también están generando nuevas referencias sobre qué se considera atractivo, deseable o normal.

Estas imágenes modificadas pasan a formar parte del ecosistema digital del que aprenden los sistemas de inteligencia artificial.

Se produce así una dinámica de retroalimentación: los modelos aprenden de nuestras representaciones, generan nuevas imágenes que refuerzan determinados cánones y estas imágenes vuelven a influir sobre nuestras percepciones y prácticas sociales.

Esto lleva a una reflexión más amplia sobre la gobernanza de la inteligencia artificial.

No importa únicamente que una imagen generada se parezca a la realidad. También importa qué tipo de sociedad y qué modelos de persona está contribuyendo a construir.

Si las grandes tecnológicas operan dentro de una lógica económica que promueve determinados modelos de consumo, ¿hasta qué punto las imágenes generadas contribuyen a reproducirlos? ¿Podemos compensar los sesgos sociales mediante ajustes algorítmicos sin cuestionar las dinámicas económicas, culturales y políticas que los producen?

Una idea importante atraviesa esta conversación: no podemos ceder a estas máquinas la responsabilidad de tomar decisiones que afectan a las personas.

2. Cómo cambian los sesgos según las tareas que pedimos a la inteligencia artificial

Muitze Zulaika, Orai NLP · Lía González, Ikuspegi

La inteligencia artificial no es neutra: también importa qué le pedimos

Esta investigación parte de una idea conocida: los modelos de inteligencia artificial tienen sesgos y sus respuestas no son neutras. Sin embargo, introduce una pregunta adicional: **¿cómo cambian esos sesgos en función de las tareas que les pedimos?**

Solemos referirnos a la inteligencia artificial como una caja negra. Introducimos una pregunta o instrucción y recibimos una respuesta, sin conocer completamente los procesos internos que la han generado.

La investigación se centró en analizar las respuestas de los modelos ante diferentes tareas textuales. **Se trataba de observar qué cambios se producen cuando les pedimos responder una pregunta, continuar una historia, elaborar un resumen o escribir un poema. De alguna manera, la tarea que planteamos ilumina determinados aspectos de la realidad y deja otros en la sombra.**

Una misma situación puede generar respuestas muy diferentes

Para estudiar este fenómeno, se plantearon situaciones en las que aparecían personas pertenecientes a determinados grupos sociales. A partir de un mismo contexto, los modelos tenían que realizar diferentes tareas: proponer un titular, continuar una historia o emitir un juicio moral.

Se observó que una misma información podría generar respuestas sin sesgos en una tarea y respuestas estereotipadas en otras.

Uno de los criterios para identificar los sesgos era observar si el modelo atribuía determinadas características o comportamientos a una persona por su pertenencia a un grupo social.

Esto resulta especialmente relevante cuando hablamos de colectivos históricamente discriminados, como el pueblo gitano o las personas LGTBI.

Los sesgos pueden aparecer mediante estereotipos negativos, pero también mediante valoraciones aparentemente positivas. Por ejemplo, cuando una respuesta destaca una acción no por lo que la persona ha hecho, sino por pertenecer a un determinado colectivo. Una respuesta que evita el sesgo puede ser simplemente reconocer que no existe información suficiente para emitir un juicio.

Cómo se construye una investigación sobre sesgos

El proceso de investigación incluía cuatro momentos.

1. Primero, diseñar las tareas y situaciones que se presentaban a los modelos.
2. Después, definir qué se consideraría una respuesta sesgada y cómo se evaluaría.
3. En tercer lugar, explorar la posibilidad de automatizar esta evaluación mediante inteligencia artificial.
4. Finalmente, analizar los resultados para identificar patrones y diferencias.

Como punto de partida se utilizó una base de datos estadounidense, BBQ (Bias Benchmark for Question Answering), adaptándola a nuestro contexto social y cultural. Se incorporaron nuevas categorías, entre ellas estereotipos relacionados con el feminismo. Por ejemplo, se planteaban situaciones en las que aparecían diferentes personas y se preguntaba cuál de ellas era una militante feminista.

En una de las situaciones, el modelo podía atribuir esa condición a una persona concreta, Nora, a partir de una descripción estereotipada. **Sin embargo, la respuesta adecuada era reconocer que no se podía saber con la información disponible.**

Este tipo de pruebas permite observar cómo los modelos completan los vacíos de información utilizando asociaciones aprendidas.

~~También aparecía una pregunta metodológica interesante: cuando incorporamos estereotipos a los datos de entrenamiento para enseñar a los modelos a identificarlos y evitarlos, ¿podemos estar contribuyendo al mismo tiempo a reproducirlos?~~

Comparación entre 16 modelos

La investigación analizó 16 modelos de inteligencia artificial, entre ellos Latxa, Kimu, Llama, Llama adaptado al euskera, Gemma, Mistral Small y Qwen.

Los resultados permitieron identificar algunas tendencias.

En términos generales, los modelos más grandes presentaban menos sesgos. También se observaron menores niveles de sesgo en determinados modelos cerrados frente a los abiertos.

Sin embargo, el tamaño no era el único factor. La familia del modelo y sus procesos de entrenamiento también influían de manera importante. Por ejemplo, Qwen presentaba menos sesgos que Llama en algunas de las comparaciones realizadas.

Los resultados de los modelos adaptados al euskera mostraron que este proceso no tiene un efecto uniforme.

La adaptación de Gemma para desarrollar Kimu había reducido determinados sesgos, mientras que la adaptación de Llama para desarrollar Latxa había incrementado algunos de ellos.

Por tanto, incorporar una nueva lengua a un modelo no consiste únicamente en mejorar su capacidad lingüística. También necesitamos evaluar qué ocurre con sus sesgos durante el proceso de adaptación.

El tipo de tarea modifica los resultados

Uno de los aprendizajes más interesantes fue observar cómo cambian los sesgos cuando pedimos al modelo que adopte un determinado rol.

- **Cuando se le solicita actuar como una persona concreta o asumir un perfil determinado, las respuestas pueden volverse mucho más estereotipadas.**
- En cambio, cuando se le piden **tareas vinculadas a la educación, los sesgos tienden a reducirse.**
- Las **tareas creativas y las propuestas más abiertas parecen generar mayores oportunidades para que aparezcan estereotipos.**

- También se encontraron diferencias entre categorías. **Los sesgos relacionados con la edad aparecían como algunos de los más intensos**, mientras que los vinculados a la orientación sexual eran menores en las pruebas realizadas.

Estos resultados refuerzan una idea: no podemos evaluar el sesgo de un modelo como si fuera una propiedad fija e independiente de su utilización. El sesgo también depende del contexto, la tarea y el tipo de respuesta que solicitamos.

Como resultado del proyecto se desarrolló una herramienta, disponible en GitHub, para automatizar parte de estas evaluaciones y facilitar la comparación entre modelos.

3. Sesgos y retos sociales de la inteligencia artificial en la sociedad vasca

Mesa de conversación con Francesca Tassinari (EHU), Ana Milián Huerta (Servicios Sociales de Barcelona) y Simón Peña (EHU, Gureiker)

Los sesgos en el ámbito de la salud

Francesca Tassinari planteó una reflexión sobre el significado del término sesgo y su utilización en el ámbito de la salud. Habitualmente entendemos el sesgo como una desviación sistemática respecto a un razonamiento lógico o racional. Es un concepto que solemos asociar con consecuencias negativas. **Sin embargo, cabe preguntarse si todos los sesgos tienen necesariamente efectos negativos o si, en determinadas circunstancias, algunos pueden parecer beneficiosos.**

La inteligencia artificial se está incorporando progresivamente a ámbitos como la prevención, el diagnóstico, el cribado, los tratamientos y la asistencia sanitaria.

En estos contextos pueden aparecer formas de sesgo que conduzcan a una sobreasistencia o a una prestación excesiva de determinados servicios.

Aunque inicialmente pudiera interpretarse como una atención más cuidadosa, no siempre es una solución adecuada. **La sobreasistencia también puede generar problemas desde la perspectiva del equilibrio de los servicios y la sostenibilidad del gasto público.**

La cuestión es comprender los efectos reales de los sesgos, más allá de clasificarlos automáticamente como positivos o negativos.

Servicios Sociales de Barcelona: detenerse a pensar dónde y para qué utilizar IA

Ana Milián Huerta compartió algunas de las experiencias y reflexiones que se están desarrollando en los Servicios Sociales de Barcelona.

Durante los últimos años se ha extendido una cierta **presión para incorporar la inteligencia artificial a las organizaciones. Parece que todas las instituciones necesitan tener proyectos vinculados a estas tecnologías.**

Desde los Servicios Sociales de Barcelona decidieron detenerse y reflexionar sobre dónde y para qué tenía sentido utilizarlas.

A partir de este trabajo identificaron diferentes ámbitos de aplicación.

- Mejorar la experiencia de la ciudadanía. Una de las posibilidades es utilizar dispositivos de **traducción simultánea para reducir las barreras idiomáticas en la relación con los servicios públicos.** Estas herramientas pueden facilitar la comunicación y mejorar la accesibilidad.
- Apoyar la tramitación administrativa. Se están explorando **asistentes virtuales y chatbots que ayuden a las personas a realizar determinados trámites, por ejemplo, en convocatorias vinculadas a fondos de infancia.** Estas herramientas también pueden facilitar la traducción, identificar incidencias o detectar errores y carencias en la información de las convocatorias.
- Identificar y gestionar riesgos. Otra línea tiene que ver con la **detección de posibles situaciones de vulnerabilidad, tanto en colectivos como en casos individuales. Es un ámbito especialmente delicado por las consecuencias que pueden tener las recomendaciones automatizadas.**
- Mejorar la gestión de programas sociales. La inteligencia artificial puede ayudar a **gestionar grandes volúmenes de información y convertir datos desestructurados en información estructurada.** Este último punto tiene un interés especial para el trabajo social. A menudo resulta difícil conseguir que las personas respondan a formularios cerrados, mientras que es más sencillo que expliquen su situación mediante un relato.

La inteligencia artificial puede facilitar la transformación de esos relatos en información organizada, sin necesidad de que toda la recogida de datos se realice mediante preguntas estructuradas.

Las trabajadoras sociales dedican una parte importante de su tiempo a tareas burocráticas, registros de entrevistas, tramitación de recursos y elaboración de documentos. Se están explorando herramientas que permitan automatizar algunas de estas actividades, como el **registro de entrevistas o la elaboración de propuestas de respuesta.**

En una de las experiencias, un sistema generaba borradores de respuesta a determinadas demandas. **El trabajo permitió comprobar que los sesgos estaban muy presentes, lo que llevó a realizar auditorías y procesos de reentrenamiento.**

El reto no es únicamente tecnológico, sino organizativo

La incorporación de inteligencia artificial a los Servicios Sociales presenta riesgos especialmente importantes.

Muchas de las aplicaciones trabajan con datos personales sensibles, relacionados con la salud, la orientación sexual, los antecedentes penales o las situaciones familiares.

Además, algunas herramientas pueden afectar directamente a los derechos de las personas, especialmente cuando se utilizan para elaborar perfiles, anticipar riesgos o recomendar determinadas actuaciones.

Esto hace que muchos proyectos puedan clasificarse como sistemas de alto riesgo, con las correspondientes obligaciones de evaluación y supervisión.

Antes de incorporarlos, es necesario analizar sus posibles impactos.

Sin embargo, una de las principales dificultades no es estrictamente tecnológica. **El desafío tiene mucho que ver con la capacidad de las organizaciones para evaluar las herramientas, establecer procedimientos, desarrollar competencias profesionales y garantizar una utilización adecuada.**

Barcelona ha avanzado en la elaboración de un **protocolo para el uso de inteligencia artificial**, con la participación de un equipo multidisciplinar.

El trabajo plantea también una tensión entre utilizar soluciones comerciales ya disponibles, como Copilot, y **desarrollar espacios tecnológicos propios y corporativos que permitan una mayor supervisión de los datos y los procesos.**

La utilización rápida de una herramienta comercial puede resultar atractiva, pero también puede introducir sesgos y riesgos difíciles de controlar. De ahí la importancia de capacitar tanto a las personas profesionales como a la ciudadanía.

No se trata únicamente de aprender a utilizar estas tecnologías, sino de comprender sus límites y desarrollar criterios para una integración razonable y ética.

La inteligencia artificial y los cambios en el periodismo

Simón Peña planteó los impactos de la inteligencia artificial sobre el periodismo y, en general, sobre las industrias de creación de contenidos.

En los últimos años estamos viviendo una transformación disruptiva que afecta tanto a los puestos de trabajo como al sentido de algunas profesiones. En el caso del periodismo, la inteligencia artificial cuestiona una de sus funciones tradicionales: la mediación entre la realidad y las personas.

Los medios de comunicación han desempeñado históricamente un papel en la selección, interpretación y contextualización de la información.

La aparición de sistemas capaces de generar textos, sintetizar acontecimientos y ofrecer respuestas directas está modificando esa función.

Esto puede producir una descapitalización de la profesión, tanto por la pérdida de puestos de trabajo como por el debilitamiento de algunas capacidades vinculadas a la producción y mediación de conocimiento.

Desinformación y centralización de los relatos

El segundo gran impacto está relacionado con la desinformación y con las formas en que los algoritmos condicionan nuestra comprensión de la realidad.

La información que recibimos puede depender de nuestras búsquedas anteriores y de los datos que las plataformas conocen sobre nosotros.

Pero, además, **los modelos de inteligencia artificial están entrenados con grandes cantidades de información que no se distribuye de manera equilibrada entre territorios, culturas e idiomas. Hay lugares y comunidades que producen o publican menos contenidos digitales. Como consecuencia, están peor representados en los modelos.**

Esto puede reforzar una centralización de los relatos. La información local o descentralizada corre el riesgo de desaparecer o quedar desplazada por contenidos procedentes de los grandes centros de producción informativa.

Aparecen así nuevos desiertos informativos: territorios sobre los que existe poca información propia y cuya realidad termina siendo explicada desde perspectivas externas, generalistas o globales.

Las consecuencias pueden afectar a múltiples ámbitos, incluidos los procesos electorales y la construcción de la opinión pública.

Sobre los sesgos técnicos y las garantías jurídicas

En la conversación posterior se planteó la necesidad de **diferenciar los sesgos que reproducen prejuicios sociales de otros sesgos derivados del diseño técnico, los datos o los contextos de utilización.**

La protección de datos ha sido uno de los primeros ámbitos desde los que se han establecido límites a los sistemas automatizados. Se mencionó el artículo 22 del Reglamento General de Protección de Datos, relacionado con las decisiones individuales automatizadas.

Sin embargo, se señaló que todavía existen vacíos y dificultades en el tratamiento normativo de los sesgos. También se hizo referencia al debate sobre el paquete normativo Digital Omnibus y a la necesidad de regular cuestiones como la calidad de los datos.

Una de las preguntas planteadas fue si disponer de muchos datos garantiza necesariamente mejores resultados. Y la respuesta fue que no siempre. **Una gran cantidad de información puede seguir reproduciendo desigualdades, especialmente cuando los datos no representan adecuadamente a determinados colectivos.**

También existen riesgos cuando trasladamos un sistema entrenado en un contexto social a otro diferente.

Por ejemplo, utilizar modelos contruidos con estereotipos estadounidenses o europeos para evaluar las necesidades de personas de otros orígenes puede conducir a interpretaciones equivocadas.

En el ámbito de la salud, esta cuestión adquiere especial relevancia cuando se trabaja con personas migrantes o colectivos vulnerables.

Otra dificultad es la falta de datos suficientemente desagregados por género y otras variables sociales. Esta carencia limita la posibilidad de identificar desigualdades y tomar decisiones ajustadas a las realidades concretas.

El contexto como elemento fundamental

En distintos momentos de la conversación apareció una reflexión sobre la tendencia a centrarnos en el objeto de análisis y olvidar el contexto.

Los modelos de inteligencia artificial se alimentan de información que ya incorpora las desigualdades de nuestra sociedad.

Se citó como ejemplo la Wikipedia, donde una mayoría muy significativa de los perfiles biográficos corresponde a hombres. Se recoge una proporción del 81 %.

Cuando utilizamos estos datos para entrenar sistemas, podemos estar trasladando estructuras sociales anteriores a nuevas tecnologías. De alguna manera, colonizamos nuevos espacios digitales con viejos valores.

Los medios de comunicación y las plataformas tecnológicas no son únicamente canales de transmisión. También intervienen en qué conocimiento se produce, se amplifica y adquiere legitimidad.

La tecnología puede reforzar estas inercias o contribuir a transformarlas, dependiendo de cómo se diseñe y gobierne.

El riesgo de confiar demasiado en las recomendaciones automatizadas

En los Servicios Sociales, las consecuencias de una recomendación sesgada pueden ser especialmente graves. Si una herramienta proporciona una interpretación incorrecta de una situación familiar o personal, puede afectar a las decisiones que toman las profesionales.

Por eso es necesario ser especialmente rigurosos con la introducción de sistemas de inteligencia artificial en procesos que afectan a derechos.

Algunas aplicaciones están prohibidas y otras requieren evaluaciones previas de impacto, supervisión y garantías adicionales.

Se destacó también el riesgo de sobreconfianza en la máquina: la tendencia a atribuir una autoridad excesiva a sus recomendaciones, incluso cuando no comprendemos cómo se han generado.

En Barcelona se ha creado un **consejo asesor de inteligencia artificial** al que se trasladan propuestas de proyectos después de realizar los correspondientes estudios de impacto.

Este trabajo se conecta con otras iniciativas vinculadas al humanismo tecnológico, las buenas prácticas y la construcción de acuerdos de ciudad.

Educación, alfabetización mediática y pensamiento crítico

La educación apareció como otro de los grandes ámbitos de transformación.

Se mencionaron los debates relacionados con los resultados del informe PISA y la necesidad de abordar de manera más profunda la incorporación de las tecnologías digitales a los procesos educativos.

No se trata de responsabilizar a las pantallas de todos los problemas. **La relación con las tecnologías ha cambiado, pero algunas preocupaciones no son completamente nuevas. Antes también podíamos pasar cuatro horas al día viendo la televisión. Lo que antes estaba en el salón pasó a nuestras habitaciones y ahora lo llevamos permanentemente en el bolsillo.**

La cuestión es cómo introducir una utilización crítica de estas herramientas.

El alumnado accede a sistemas completos, con sus posibilidades y sus sesgos. No necesariamente existen mecanismos que permitan filtrar o identificar esas limitaciones.

Por eso resulta importante reforzar la alfabetización mediática, la capacidad de interpretación, el pensamiento crítico y la comprensión de los procesos mediante los que se construyen las respuestas.

Retos compartidos por los diferentes ámbitos

En la conversación se identificaron varios desafíos.

1. En primer lugar, la necesidad de avanzar hacia una **regulación más completa y coordinada**. Las tecnologías evolucionan rápidamente y los procedimientos legislativos tienen dificultades para seguir ese ritmo. Se señalaron lagunas y problemas de articulación entre diferentes normas, especialmente ante la expansión de la inteligencia artificial generativa.
2. En segundo lugar, se planteó que las organizaciones no deberían incorporar estas herramientas únicamente **desde una lógica industrial de optimización de recursos**.
 - a. En el periodismo y las industrias creativas existen finalidades que no pueden reducirse a producir más contenidos con menos personas. Los posibles ahorros deberían orientarse también a **mejorar la calidad del trabajo y de los servicios**.
 - b. En educación, esto exige repensar el lugar que ocupan los resultados frente a los **procesos de aprendizaje**. La inteligencia artificial puede proporcionar respuestas de manera inmediata, pero necesitamos seguir desarrollando la capacidad de razonar, interpretar y construir conocimiento.

- c. En los Servicios Sociales, el principal reto es mantener la **capacidad de supervisión y decisión profesional ante una evolución tecnológica muy rápida.**

La pregunta compartida no es únicamente qué puede hacer la inteligencia artificial, sino qué queremos que haga, para qué y bajo qué condiciones.

4. Tendencias tecnológicas, evaluación de sesgos y desarrollo de modelos propios

Mesa con Olatz Pérez (HiTZ), Guillermo Villate (TECNALIA), Harritxu Gete (Vicomtech) y Xabier Saralegi (Orai NLP)

Sabemos que hay sesgos, pero todavía nos faltan herramientas para comprenderlos

Una de las primeras reflexiones de esta mesa fue que la inteligencia artificial ha evolucionado con tanta rapidez que todavía no comprendemos suficientemente qué tenemos entre manos.

Sabemos que los modelos incorporan sesgos, pero necesitamos diagnósticos mucho más precisos sobre cómo funcionan y cuáles son sus efectos.

Cuando preguntamos directamente a un modelo por determinados estereotipos, puede ofrecer respuestas aparentemente correctas y evitar expresiones discriminatorias. Sin embargo, cuando esos mismos estereotipos aparecen de manera indirecta dentro de una tarea, las respuestas pueden reproducirlos.

Esto hace que la evaluación sea especialmente compleja.

No basta con preguntar a los sistemas si están sesgados. Necesitamos diseñar pruebas que permitan observar cómo actúan en situaciones reales, en diferentes contextos y ante distintas instrucciones.

Los sesgos políticos y culturales son especialmente relevantes. También lo son las interacciones entre categorías como el género, la edad o el origen.

Actualmente nos faltan tanto conocimiento como herramientas para elaborar diagnósticos suficientemente completos.

¿Qué significa eliminar los sesgos?

La conversación abordó una cuestión de fondo: **qué entendemos por un modelo sin sesgos. No todas las personas compartimos los mismos valores ni tenemos la misma interpretación de lo que constituye una respuesta adecuada.**

Se utilizó como ejemplo la distancia entre distintos modelos de sociedad y el hecho de que algunas grandes empresas tecnológicas estén controladas por personas con posiciones políticas muy concretas.

Si quienes diseñan y gobiernan los modelos tienen determinadas concepciones del mundo, estas pueden influir sobre los sistemas que producen.

Por tanto, eliminar sesgos no es únicamente una operación técnica. ¿Quién decide qué es un sesgo y qué valores deben orientar las respuestas de una inteligencia artificial?

Desde un punto de vista tecnológico, podemos intervenir en los datos y ajustar el comportamiento de los modelos. Sin embargo, el problema es también social y político. Si no existe un consenso completo sobre determinados derechos, valores o interpretaciones de la realidad, tampoco podemos esperar que los modelos reflejen una neutralidad universal.

Se habló incluso de un **choque entre utopías: diferentes concepciones sobre la sociedad que queremos construir.**

La ausencia de representación de determinados colectivos también constituye una forma de sesgo. Esto complica todavía más la aspiración de construir sistemas completamente objetivos.

Sesgos que aparecen en las traducciones

Otra línea de trabajo se centra en observar cómo las **herramientas de traducción automática pueden reproducir o introducir sesgos.**

En Vicomtech están investigando especialmente los sesgos de género durante los procesos de traducción.

Se mencionó un ejemplo: cuando hablamos de pilotos de avión, los modelos pueden tender a utilizar formas masculinas porque los datos de entrenamiento contienen más referencias a hombres desempeñando esa profesión. En algunos experimentos, las referencias a la política o al ejercicio del poder seguían apareciendo asociadas mayoritariamente a hombres.

Así, los estereotipos presentes en internet vuelven a aparecer en las traducciones y contribuyen a mantener las mismas asociaciones.

Los sesgos pueden resultar menos visibles en lenguas que no marcan de la misma manera determinadas categorías gramaticales, como ocurre con el género en

euskera. Sin embargo, que un sesgo no aparezca directamente en la estructura de una lengua no significa que haya desaparecido de nuestras interpretaciones o de los modelos.

También se observó que **proporcionar más contexto puede mejorar determinadas traducciones, aunque no siempre elimina los patrones de desigualdad.**

Auditorías y nuevas formas de evaluación

Las personas participantes coincidieron en que necesitamos **avanzar en la evaluación y auditoría de los modelos.**

En TECNALIA están trabajando en estas cuestiones, conscientes de que todavía no se realizan suficientes análisis sistemáticos.

Una de las posibilidades planteadas fue que los sistemas no ofrecieran necesariamente una única respuesta, sino distintas alternativas o interpretaciones.

Esto permitiría que las personas profesionales pudieran contrastarlas y tomar decisiones desde el conocimiento del contexto.

La falta de contexto aparece como una de las principales limitaciones de muchas aplicaciones actuales. También se están desarrollando nuevas soluciones relacionadas con los chatbots, las salvaguardas y los mecanismos de control.

Sin embargo, los modelos están adquiriendo cada vez más autonomía, lo que obliga a revisar continuamente las formas de supervisarlos.

¿Es posible construir una inteligencia artificial sin sesgos?

La posibilidad de desarrollar un modelo completamente libre de sesgos fue cuestionada durante la conversación. Se planteó como una aspiración difícil de alcanzar, del mismo modo que resulta imposible pensar en una sociedad humana completamente libre de prejuicios, interpretaciones y marcos culturales. La objetividad absoluta no existe.

Aunque técnicamente podamos modificar los datos y orientar los comportamientos de los modelos, siempre aparecerá la pregunta sobre quién define los criterios de evaluación.

Además, las grandes corporaciones que desarrollan estas herramientas pueden tener intereses en conservar determinados sesgos. No todas las discriminaciones se mantienen por desconocimiento. En ocasiones conocemos sus efectos y, aun así, las reproducimos porque existen intereses económicos o políticos.

La tecnología permite intervenir sobre los modelos, pero no resuelve por sí sola los conflictos de valores que existen en la sociedad.

Modelos abiertos, regulación y autonomía tecnológica

La conversación abordó también la tensión entre el desarrollo tecnológico en Europa y el avance de Estados Unidos y China.

Existe incertidumbre sobre cuál debe ser el papel de los modelos libres y de código abierto. Desde la investigación, disponer de modelos abiertos resulta especialmente importante. Permite estudiar su funcionamiento, modificarlos, evaluar sus limitaciones y desarrollar nuevas soluciones. Al mismo tiempo, es necesario establecer límites sobre determinados usos.

Se defendió la importancia de una regulación europea, aunque también se señalaron las dificultades de regular una tecnología que evoluciona a mayor velocidad que los procedimientos normativos.

Parte del marco jurídico se ha construido a partir de categorías y tecnologías anteriores a la aparición de los grandes modelos generativos.

La cuestión es encontrar un equilibrio: necesitamos reglas y garantías, pero no podemos intentar regular cada detalle de una realidad que cambia constantemente.

El euskera y la necesidad de desarrollar tecnología propia

La situación del euskera ocupó una parte importante de la conversación. **Se valoró positivamente que las tecnologías más avanzadas puedan interactuar y traducir cada vez mejor en euskera.**

Hace apenas unos años, muchas de estas posibilidades parecían difíciles de imaginar.

Sin embargo, se expresó una preocupación compartida: **no podemos confiar el futuro digital del euskera únicamente a las grandes corporaciones tecnológicas.** Estas empresas desarrollan sus modelos en función de intereses económicos y estrategias comerciales que pueden cambiar. Por ejemplo, una actualización que mejora el comportamiento general de un modelo puede empeorar sus resultados en euskera.

También se han observado diferencias importantes cuando se utilizan modelos para evaluar respuestas o actuar como jueces en determinados procesos. En algunos experimentos, los resultados obtenidos en euskera son peores.

Por eso resulta necesario seguir creando y cuidando los datos lingüísticos, desarrollar productos propios e investigar sobre las necesidades concretas de la comunidad euskaldun.

Se compartieron avances vinculados a Latxa y a su evolución hacia nuevas capacidades y modalidades.

La existencia de modelos propios permite evaluar mejor su adecuación a la lengua y la cultura, sin depender completamente de decisiones tomadas en Silicon Valley.

También se destacó que las organizaciones de investigación y desarrollo tecnológico del territorio están más cerca de las necesidades de las personas y las empresas que las grandes plataformas internacionales.

En este sentido, el conocimiento local constituye una ventaja.

La investigación y el apoyo institucional

El desarrollo de modelos adaptados al euskera requiere **recursos, investigación y continuidad.**

No basta con crear una herramienta y dar por completado el trabajo. **Es necesario seguir alimentando los modelos con buenos datos, evaluar su funcionamiento y adaptarlos a nuevas necesidades.**

Cada vez aparecen más demandas de las empresas relacionadas con el uso de inteligencia artificial en euskera.

En centros como TECNALIA se están explorando alternativas para avanzar hacia modelos más propios y soluciones adaptadas a necesidades concretas.

Se destacó la importancia del apoyo institucional para sostener estos procesos de investigación y desarrollo.

La autonomía tecnológica no significa necesariamente construirlo todo desde cero, sino desarrollar la capacidad para participar en las decisiones, adaptar las herramientas y evitar dependencias absolutas de proveedores externos.

Notas del Seminario Internacional «Tecnología, poder y ciudadanía global: ¿Qué futuro nos depara la Inteligencia Artificial?» 20 de octubre de 2026 · Encuentro impulsado por Alboan y la Universidad de Deusto

Estas notas recogen algunas de las ideas, reflexiones y preguntas surgidas en las ponencias y conversaciones del seminario. Se organizan por bloques temáticos, manteniendo el carácter de apuntes personales, las ideas principales y los interrogantes que aparecieron durante el encuentro.

1. Inteligencia artificial, poder y regulación: un sistema jurídico que parece haber quedado obsoleto

Conversación con Paloma Llana y Carlos Bajo Erro

El sistema jurídico como ficción compartida

Una de las primeras ideas tiene que ver con entender el sistema jurídico como una ficción compartida, podríamos decir que intersubjetiva. Las fronteras, el dinero o los propios sistemas legales son construcciones colectivas que hemos desarrollado para organizar nuestras relaciones, establecer acuerdos y regular posibles conflictos.

Una frontera puede coincidir con un accidente geográfico, pero el concepto de frontera es una construcción abstracta. Su existencia y legitimidad dependen de que compartamos una determinada interpretación de la realidad.

El problema es que nuestro sistema jurídico y los mecanismos de regulación que hemos construido parten de una distribución del poder en la que la soberanía recae fundamentalmente en los Estados. Las leyes funcionan porque hemos delegado en ellos la capacidad de ejercer la fuerza para garantizar su cumplimiento.

¿Qué ocurre cuando las empresas tecnológicas operan en una realidad que desborda las fronteras territoriales y los mecanismos tradicionales de regulación?

Las grandes empresas tecnológicas actúan a escala global y los modelos de lenguaje tienen obligaciones todavía difíciles de concretar y hacer cumplir. Si las empresas no están suficientemente vinculadas a un territorio, ¿quién garantiza el cumplimiento de las leyes?, ¿quién puede sancionarlas?, ¿cómo se hacen efectivos los límites?

Desde esta perspectiva, el sistema de regulación parece haber nacido parcialmente obsoleto. Se centra en determinadas cuestiones, como la gestión de los datos, pero no alcanza a comprender ni regular toda la complejidad de lo que está ocurriendo.

La externalización de la cognición humana

La inteligencia artificial generativa está abriendo un proceso de externalización de una parte de la cognición humana. Estamos delegando en las máquinas actividades que hasta ahora asociábamos a nuestras capacidades de análisis, interpretación, elaboración y decisión. También es cierto que parte de esta delegación ya había comenzado con anterioridad a la llegada de la IA generativa llevando a otros países parte del trabajo de explotación de datos, por ejemplo.

En este contexto, resulta especialmente difícil establecer mecanismos de protección suficientemente sólidos. Los posibles efectos y derivas de estas tecnologías son difíciles de anticipar y controlar.

Aparecía una imagen especialmente sugerente: estamos incorporando modelos de lenguaje en nuestras organizaciones como si fueran personas trabajadoras descontroladas, a las que permitimos intervenir en nuestros procesos sin comprender del todo cómo funcionan ni qué consecuencias pueden tener sus actuaciones.

La tensión entre regulación y competitividad

La Unión Europea está intentando establecer límites a través de la regulación, mientras Estados Unidos apuesta por un escenario mucho más desregulado. Esta diferencia genera una tensión importante: las normas que pretenden proteger a la ciudadanía pueden convertirse también en una dificultad competitiva para las empresas europeas.

Además, las grandes tecnológicas operan con niveles de complejidad extraordinarios. Esa complejidad no es únicamente una característica de sus productos, sino que también puede convertirse en una herramienta para dificultar su supervisión y regulación.

Cuando los mecanismos de control intentan alcanzar una determinada realidad tecnológica, las grandes empresas ya están trabajando en una nueva escala de complejidad.

¿Dónde podemos poner límites cuando los instrumentos tradicionales de regulación parecen insuficientes?

Una de las posibilidades planteadas es intervenir a través de la fiscalidad. También utilizar las propias herramientas tecnológicas y su complejidad como mecanismos de supervisión y contrapoder. Se mencionaron estrategias de poder blando (soft power) y formas de resistencia más cercanas a la guerrilla.

En el fondo, la pregunta va más allá de cómo adaptar las normas existentes. Quizá el sistema, tal y como lo conocíamos, ha dejado de funcionar para regular esta nueva realidad y necesitamos construir otro.

De lo contrario, corremos el riesgo de quedarnos en narrativas y respuestas emocionales que no consiguen frenar lo que se describió como una subversión democrática impulsada desde Estados Unidos.

Europa, Estados Unidos y China: distintos escenarios de poder

En este proceso existe también una disputa geopolítica entre Europa, Estados Unidos y China.

California ocupa un lugar especialmente significativo. Es uno de los grandes centros de concentración de poder tecnológico y económico del mundo. Se mencionó que, si fuera un Estado independiente, ocuparía una de las primeras posiciones entre las grandes economías mundiales.

Precisamente allí están apareciendo movimientos sociales que cuestionan la expansión de los centros de datos, su consumo de agua y sus consecuencias sobre los territorios.

La batalla por la regulación también se está librando en las calles. Estos movimientos de resistencia pueden ofrecer aprendizajes interesantes sobre cómo construir capacidad de agencia frente a las grandes tecnológicas.

En China, más allá de su configuración política autoritaria y del modelo particular de capitalismo que está desarrollando, también aparecen algunas prácticas que merece la pena observar. **Una de las cuestiones mencionadas fue la existencia de límites a la antropomorfización de las inteligencias artificiales generativas. Llamar a una herramienta «chatín» o relacionarnos con ella como si fuera una persona no es una cuestión inocente. Esa humanización puede reforzar los mecanismos de vinculación emocional y dependencia que generan estas tecnologías.**

Aquí aparece una cuestión especialmente relevante: **¿hasta qué punto las estrategias de diseño de estas herramientas buscan producir adicción como parte de su modelo de negocio?**

Datos de entrenamiento, agentes y transformación del trabajo

Otro de los ámbitos de discusión tiene que ver con la **legitimidad de los datos utilizados para entrenar los modelos y con las diferencias entre la inteligencia artificial generativa abierta y los agentes especializados.**

Solemos insistir en que la IA se equivoca, inventa respuestas y puede alejarse de la realidad. Sin embargo, los agentes que trabajan dentro de entornos más acotados, con conocimientos propios y mecanismos de supervisión, pueden alcanzar niveles de fiabilidad mucho mayores en determinadas tareas.

Esto cambia la conversación sobre el impacto laboral.

Una parte de los trabajos que antes deslocalizábamos hacia otros países puede comenzar a realizarse mediante agentes de inteligencia artificial. Aparecen nuevas posibilidades de automatización y nuevas formas de reorganizar los procesos productivos.

En ese contexto, necesitamos preguntarnos cómo redefinir el trabajo, qué funciones queremos delegar, cómo cambia el valor de las capacidades profesionales y qué relaciones de poder están apareciendo en el ámbito laboral.

Fiscalidad, cotizaciones y sostenibilidad del sistema de bienestar

Una de las cuestiones más relevantes es cómo adaptar nuestros mecanismos de protección social a un escenario en el que una parte importante del trabajo humano pudiera ser sustituida por sistemas automatizados.

Se recuperó una propuesta que ya había aparecido en los debates sobre robótica: establecer cotizaciones vinculadas a los robots o, en este caso, a los agentes de inteligencia artificial que sustituyen trabajo humano.

Si una empresa deja de contratar a una persona porque puede realizar sus funciones mediante un agente, ¿debería existir algún mecanismo de cotización que contribuya a sostener el sistema de protección social?

Durante la conversación se mencionaron previsiones de pérdida o transformación de hasta un 40 % del trabajo. Son estimaciones inciertas, en una realidad que todavía no se ha estabilizado, pero permiten comprender la magnitud del problema.

Si una parte significativa del empleo desaparece, también pueden reducirse los ingresos fiscales y las cotizaciones que sostienen las pensiones y otros servicios públicos. Se planteó la posibilidad de establecer tasas sobre determinados servicios automatizados y destinar esos recursos a fondos públicos.

En definitiva, parece necesario comenzar a hablar de estas cuestiones antes de que los efectos sobre la financiación del bienestar sean difíciles de abordar.

El riesgo de perder capacidades profesionales

Apareció también una reflexión sobre los límites de la automatización y los posibles errores que cometemos cuando delegamos determinadas funciones en las máquinas.

La inteligencia artificial no es tan infalible como a menudo se presenta. **Una delegación poco cuidadosa puede deteriorar los procesos, producir peores datos y terminar generando productos o servicios de menor calidad.**

Pero quizá uno de los riesgos más importantes esté relacionado con los procesos de aprendizaje y desarrollo profesional.

Las empresas pueden dejar de contratar a personas jóvenes para realizar tareas que ahora resuelven mediante inteligencia artificial. Sin embargo, esas tareas también eran una puerta de entrada al mundo laboral y una oportunidad para desarrollar experiencia.

Si dejamos de contratar profesionales junior, ¿cómo podrán convertirse en profesionales senior? ¿Cómo se construirá la experiencia? ¿Dónde se desarrollarán los conocimientos prácticos? ¿Cómo seguiremos cuidando el propósito profesional de las nuevas generaciones?

La inteligencia artificial puede reforzar la precarización laboral, especialmente entre las personas jóvenes, y modificar las trayectorias de aprendizaje que hasta ahora sostenían muchas profesiones.

También se mencionaron experiencias de empresas que han consumido rápidamente sus presupuestos de tokens y han comenzado a preocuparse por limitar el gasto. **Los costes de las licencias pueden aumentar, aunque todavía puedan seguir siendo competitivos frente al coste de contratar personas.**

Una sociedad preparada para regalar sus datos

La expansión de la inteligencia artificial generativa no se produce en el vacío. **Llega a una sociedad que durante años ha entregado sus datos de manera casi inconsciente, que ha desarrollado fuertes dependencias del teléfono móvil y de las redes sociales, y que vive procesos crecientes de precarización e individualización.**

En ese contexto, la adopción de ChatGPT y otras herramientas generativas ha sido extraordinariamente rápida. **En muy poco tiempo hemos incorporado estas**

tecnologías a nuestra vida cotidiana, probablemente sin comprender todavía todas sus consecuencias.

Una parte importante del negocio de las grandes tecnológicas está vinculada a los datos. Como personas consumidoras aceptamos dispositivos y aplicaciones más baratos a cambio de ceder información sobre nuestras vidas.

No parece que valoremos suficientemente la privacidad o el uso ético de los datos.

Aparecía un ejemplo: ¿estaríamos dispuestos a pagar 300 euros más por una televisión que respetara nuestra privacidad y no utilizara nuestros datos para alimentar mecanismos de dependencia y consumo?

Una frase de la ponencia resume especialmente bien esta preocupación:

Los datos que regalamos desde 2008 nos están quitando el trabajo hoy. Los que regalamos ahora pueden quitarles el trabajo a nuestros hijos e hijas.

No se trata únicamente de proteger nuestra información personal, sino de comprender cómo el conocimiento que entregamos puede convertirse en una herramienta para automatizar nuestro propio trabajo.

2. Inteligencia artificial y sostenibilidad ambiental: hacia algoritmos verdes

Ponencia de Verónica Bolón. Universidade da Coruña

La segunda ponencia se centró en el impacto medioambiental de la inteligencia artificial y en la necesidad de incorporar la sostenibilidad al diseño y utilización de estas tecnologías.

La tecnología nunca es neutra

La inteligencia artificial no es una tecnología nueva. Sus orígenes se sitúan en 1956 y las primeras redes neuronales comenzaron a desarrollarse en la segunda mitad de los años cincuenta.

Lo que ha cambiado en los últimos años es la combinación de tres factores: la disponibilidad de cantidades enormes de datos, una capacidad de cómputo mucho mayor y un volumen extraordinario de financiación.

Esta combinación está haciendo posible la expansión acelerada que estamos viviendo.

Tradicionalmente, cuando hablamos de una inteligencia artificial de calidad, solemos pedirle tres cosas:

- **Robustez:** que funcione de manera fiable, cometa pocos errores y sea resistente a ataques o manipulaciones.
- **Transparencia:** que podamos comprender, en la medida de lo posible, cómo toma sus decisiones y cómo se relaciona el algoritmo con las personas.
- **Ética:** que no contribuya a generar discriminaciones, que respete la privacidad y que tenga en cuenta sus consecuencias sociales.

A estas tres dimensiones necesitamos incorporar una cuarta: **la sostenibilidad ambiental.**

No es suficiente con desarrollar algoritmos fiables, transparentes y éticos. También necesitamos preguntarnos cuánta energía consumen, qué recursos materiales requieren y qué impacto generan sobre el planeta.

La huella ambiental de la inteligencia artificial

La inteligencia artificial contamina. Y lo hace tanto durante el entrenamiento de los modelos como durante su utilización cotidiana.

Durante la ponencia se compartieron diferentes comparaciones que ayudan a visualizar esta realidad.

- Se mencionó que entrenar un único modelo de inteligencia artificial puede generar emisiones de CO₂ equivalentes a las de cinco coches durante toda su vida útil.
- También se compartió una estimación según la cual el entrenamiento de GPT-3 habría requerido una cantidad de energía comparable al consumo anual de 126 hogares daneses.

Son referencias que permiten hacernos una idea de la escala del problema, aunque conviene revisar las fuentes originales y las condiciones de cada estimación antes de utilizarlas como datos generales.

Una de las principales líneas de investigación consiste precisamente en reducir el volumen de información y las operaciones necesarias para que los modelos puedan funcionar, disminuyendo así su consumo energético.

La nube es una infraestructura física

Cuando hablamos de servicios en la nube, podemos tener la impresión de que estamos utilizando una tecnología inmaterial. Sin embargo, detrás de esa nube hay grandes centros de datos que ocupan enormes extensiones de terreno, consumen electricidad y necesitan agua para refrigerar sus equipos.

Se mencionó el caso de Aragón, donde están proliferando este tipo de instalaciones. El impacto de los centros de datos no afecta únicamente al consumo energético global, sino también a los territorios y comunidades donde se instalan.

Durante la ponencia se citaron estimaciones sobre incrementos muy significativos en las **facturas de electricidad de comunidades cercanas a centros de datos en Estados Unidos, llegando a mencionarse un 267 %.**

También se habló del enorme consumo de agua de estas infraestructuras y de las consecuencias sobre los recursos hídricos disponibles.

Otra cuestión especialmente interesante fue el efecto sobre la temperatura del entorno. **Los centros de datos actúan como islas de calor y se mencionaron estimaciones de incrementos cercanos a dos grados en sus proximidades.**

Todo esto refuerza la necesidad de observar la inteligencia artificial no solo como una herramienta digital, sino como una infraestructura material con consecuencias territoriales, energéticas y ambientales.

No todas las interacciones tienen el mismo impacto

Otra idea importante es que el consumo de recursos varía mucho según la tarea que realizamos. Una conversación de texto con un chatbot requiere, por lo general, menos recursos que la generación de imágenes.

Se utilizó una comparación especialmente gráfica: generar una imagen mediante determinados modelos de inteligencia artificial puede representar un consumo energético comparable al necesario para cargar completamente un teléfono móvil.

También se recordó la campaña en la que millones de personas comenzaron a transformar sus fotografías para obtener imágenes con una estética similar a la de los estudios Ghibli. Se mencionó que alrededor de un millón de personas utilizaron la herramienta en apenas una hora para este tipo de transformaciones.

Es un ejemplo de cómo una práctica aparentemente pequeña puede adquirir una dimensión ambiental extraordinaria cuando se multiplica por millones de personas.

Incluso nuestras interacciones más cotidianas, como escribir «por favor» o «gracias» a las máquinas, generan operaciones y un consumo de energía asociado.

La cuestión no es culpabilizar cada interacción, sino hacer visible que existe un coste material que normalmente permanece oculto.

Entrenamiento e inferencia: dos momentos diferentes

Una de las distinciones más interesantes de la ponencia fue la diferencia entre entrenar un modelo y utilizarlo.

- Cuando entrenamos una inteligencia artificial para distinguir fotografías de perros y gatos, necesitamos proporcionarle una cantidad muy importante de imágenes para que pueda aprender a reconocer patrones. Este proceso de entrenamiento puede requerir una gran cantidad de recursos computacionales y energéticos.
- Sin embargo, una vez que el modelo está entrenado, mostrarle una fotografía y preguntarle si aparece un perro o un gato puede representar un consumo mucho menor. El problema es que esta segunda operación, conocida como inferencia, puede realizarse millones de veces al día.

Por tanto, aunque cada interacción sea relativamente barata en términos energéticos, el consumo acumulado puede ser enorme.

Aprovechar el calor y repensar las infraestructuras

Entre las alternativas planteadas aparecieron algunas relacionadas con la propia infraestructura tecnológica.

Los centros de datos generan grandes cantidades de calor residual. Una posibilidad es aprovechar ese calor para otros usos, como la climatización de viviendas o el funcionamiento de invernaderos.

También se mencionó la posibilidad de desarrollar pequeños centros de datos distribuidos, incluso integrados en edificios o viviendas, que pudieran contribuir a sistemas de calefacción.

Estas propuestas permiten pensar en modelos de aprovechamiento energético diferentes, aunque no eliminan por sí mismas el consumo de recursos.

Otra línea de actuación tiene que ver con la **defensa de los territorios y la capacidad de las comunidades para participar en las decisiones sobre dónde y cómo se instalan estas infraestructuras.**

Diseñar algoritmos que contaminen menos

Más allá de los centros de datos, una de las principales propuestas de Verónica Bolón es trabajar en el propio diseño de los algoritmos.

No todos los modelos necesitan tener un tamaño enorme ni utilizar cantidades masivas de información para resolver las tareas que les pedimos.

Entre las estrategias mencionadas aparecen la reducción del tamaño de las redes neuronales, la destilación de modelos grandes para obtener otros más pequeños, la reutilización de modelos existentes y el uso de energías renovables.

Se trata de avanzar hacia una inteligencia artificial que alcance resultados adecuados con el menor consumo posible de recursos.

Aquí aparece una cuestión importante: no siempre necesitamos la inteligencia artificial más potente; necesitamos la herramienta más adecuada para cada tarea.

La inteligencia artificial también puede ayudar al planeta

Existen muchas aplicaciones que pueden contribuir a mejorar la sostenibilidad ambiental: la eficiencia energética, la previsión meteorológica, la planificación de rutas, la modelización de la demanda, la anticipación de inundaciones o el desarrollo de políticas ambientales.

Sin embargo, se insistió en la diferencia entre inteligencia artificial tradicional e inteligencia artificial generativa.

Para muchas de estas aplicaciones no necesitamos modelos generativos de gran tamaño. Los sistemas tradicionales pueden resolverlas de forma mucho más eficiente.

Se mencionaron diferencias importantes de consumo, incluso de varios órdenes de magnitud en determinadas comparaciones.

Desde esta perspectiva, el despliegue indiscriminado de modelos generativos esta dificultando más que ayudando a la lucha contra la crisis climática.

El reto consiste en utilizar la inteligencia artificial allí donde genera un beneficio que justifique su coste ambiental.

¿Qué podemos hacer como personas usuarias?

Además de los cambios tecnológicos, institucionales y regulatorios, aparecieron algunas propuestas relacionadas con nuestros hábitos cotidianos.

- La primera tiene que ver con **revisar y reutilizar**. Antes de generar algo nuevo mediante inteligencia artificial, podemos preguntarnos si ya existe un recurso que nos sirva. Por ejemplo, no necesitamos generar imágenes nuevas para todas nuestras presentaciones. En muchas ocasiones podemos utilizar fotografías o recursos gráficos existentes.
- Tampoco necesitamos acudir a la inteligencia artificial generativa para cualquier búsqueda de información. **Una búsqueda directa en la web puede ser suficiente y más eficiente.**
- **Lo mismo ocurre con los correos electrónicos y otras tareas sencillas. ¿Realmente necesitamos activar un modelo de lenguaje para redactar cada mensaje? La idea es evitar matar moscas a cañonazos.**

Otra propuesta es **mejorar la información disponible sobre los consumos energéticos de estas herramientas. Del mismo modo que conocemos el consumo de otros dispositivos y servicios, deberíamos poder comprender mejor el coste de las operaciones que realizamos con inteligencia artificial.**

Finalmente, apareció una llamada a generar conciencia, compartir esta información y contribuir a que el impacto ambiental de las tecnologías digitales forme parte de las conversaciones públicas.

3. Minerales críticos, extractivismo y resistencias desde los territorios

Janvier Zihahirwa de República democrática del Congo y Martha Lucía Márquez de Colombia

El encuentro incorporó también una mirada desde Colombia y el Congo, con intervenciones que abordaron los minerales críticos, el extractivismo y las resistencias de las comunidades afectadas.

Estas aportaciones permitieron ampliar la reflexión sobre el impacto de la tecnología. **La inteligencia artificial y el conjunto de infraestructuras digitales necesitan materiales, minerales, energía y procesos industriales que tienen consecuencias concretas sobre determinados territorios.**

Se habló de los conflictos vinculados a la minería, de las guerras y de la destrucción de los territorios, así como de las diferentes formas de activismo y resistencia que están apareciendo frente a estos procesos extractivos.

Esta perspectiva introduce una dimensión importante: el desarrollo tecnológico que se produce y consume principalmente en determinados lugares del mundo depende también de recursos y procesos que generan costes sociales y ambientales en otros territorios.

Los apuntes recogidos sobre estas intervenciones son más breves, pero permiten situar una cuestión de fondo: no podemos comprender la inteligencia artificial sin observar también las cadenas materiales, económicas y políticas que hacen posible su existencia.

Resumen post

1. La inteligencia artificial no es neutra: aprende nuestros valores y reproduce nuestros sesgos

Los modelos de inteligencia artificial aprenden de la información que hemos producido como sociedad. Por tanto, **no solo incorporan nuestros conocimientos, sino también nuestros prejuicios, estereotipos y maneras de interpretar la realidad.**

Las investigaciones presentadas mostraron cómo la apariencia física, el género, la edad o el origen racial percibido pueden condicionar las respuestas de los modelos. El denominado **lookismo algorítmico** es un ejemplo: las personas consideradas más atractivas (en los experimentos usaban “filtros de belleza” disponibles en redes sociales) tienden a ser asociadas con cualidades positivas, como la inteligencia, la competencia o la confianza..

Además, los sesgos no son características fijas de un modelo. Cambian según la tarea, el idioma, el contexto y las instrucciones que utilizamos.

2. Estamos creando un bucle que puede reforzar las desigualdades existentes

La inteligencia artificial no se limita a reproducir la realidad. También contribuye a construirla.

Los modelos aprenden de contenidos digitales que ya incorporan desigualdades y estereotipos. **Posteriormente generan nuevos textos, imágenes y representaciones que pueden reforzar esos mismos patrones.** Estos contenidos vuelven a circular y condicionan nuestras percepciones y comportamientos.

Se produce así una dinámica de retroalimentación: la sociedad alimenta los modelos y los modelos influyen sobre la sociedad.

La pregunta no es únicamente cómo eliminar los sesgos de una herramienta, sino cómo evitar que las tecnologías consoliden las desigualdades que pretendemos superar.

3. Las grandes tecnológicas apuesta a complejidad y desbordan nuestra capacidad de regulación

Los mecanismos jurídicos tradicionales se construyeron sobre una distribución del poder basada principalmente en los Estados y sus territorios. Sin embargo, **las grandes empresas tecnológicas operan a escala global, con una capacidad económica y tecnológica que dificulta su supervisión.**

La Unión Europea intenta establecer garantías mediante la regulación, mientras otros entornos, especialmente Estados Unidos, siguen estrategias de desregulación extrema. **Aparece una tensión entre proteger derechos, mantener la competitividad y desarrollar autonomía tecnológica.**

Pero el problema va más allá de actualizar las leyes. Necesitamos revisar quién tiene capacidad real para decidir sobre el desarrollo tecnológico, establecer límites y garantizar su cumplimiento.

En este contexto, las iniciativas ciudadanas, las resistencias territoriales y la fiscalización aparecen como formas de contrapoder que merecen atención.

4. Estamos externalizando capacidades cognitivas y transformando el sentido del trabajo

La inteligencia artificial generativa permite delegar tareas de análisis, interpretación, elaboración de contenidos y toma de decisiones que hasta ahora realizábamos las personas.

Esto abre oportunidades para mejorar procesos, pero también plantea riesgos de dependencia y pérdida de capacidades profesionales.

Una preocupación especialmente relevante es la desaparición de tareas que tradicionalmente permitían aprender un oficio. Si las organizaciones dejan de contratar personas jóvenes porque determinadas funciones pueden automatizarse, ¿cómo se desarrollará la experiencia necesaria para convertirse en profesionales con mayor conocimiento y criterio?

Además, la automatización plantea un reto para los sistemas de bienestar. **Si disminuye el trabajo humano remunerado, también pueden reducirse las**

cotizaciones y los ingresos fiscales. Se recuperó la posibilidad de establecer mecanismos de contribución vinculados a los agentes automatizados.

5. No deberíamos incorporar inteligencia artificial sin preguntarnos antes para qué

La experiencia compartida desde los Servicios Sociales de Barcelona resultó especialmente ilustrativa. Frente a la presión por incorporar inteligencia artificial, optaron por detenerse y analizar dónde podía aportar valor y bajo qué condiciones.

Se identificaron aplicaciones para facilitar trámites, mejorar la accesibilidad, organizar información o reducir tareas burocráticas. Pero también riesgos importantes cuando los sistemas intervienen en decisiones que afectan a derechos y situaciones de vulnerabilidad.

La incorporación de estas herramientas requiere protocolos, evaluación de impacto, auditorías, formación y supervisión profesional.

El reto no es solamente aprender a utilizar inteligencia artificial, sino desarrollar criterios compartidos para decidir cuándo utilizarla, cuándo no y qué responsabilidades deben permanecer en las personas.

6. La inteligencia artificial puede concentrar el conocimiento y debilitar los relatos locales

Los modelos se entrenan con enormes cantidades de información, pero esta no representa de manera equilibrada a todas las comunidades, culturas, territorios e idiomas.

Los lugares que producen menos contenido digital corren el riesgo de quedar invisibilizados o de ser explicados desde perspectivas externas. En el periodismo, esta tendencia puede reforzar la centralización de los relatos y generar nuevos desiertos informativos.

Esta preocupación conecta con el desarrollo tecnológico en euskera. **Aunque los grandes modelos están mejorando sus capacidades lingüísticas, no podemos confiar exclusivamente en las decisiones comerciales de unas pocas corporaciones.**

Necesitamos seguir desarrollando datos, investigación, herramientas y modelos propios, como Latxa y Kimu, que respondan a las necesidades culturales y lingüísticas de nuestra comunidad.

La autonomía tecnológica no significa necesariamente construirlo todo desde cero, sino conservar capacidad para comprender, adaptar y decidir sobre las tecnologías que utilizamos.

7. La inteligencia artificial tiene una huella ambiental que permanece demasiado invisible

La llamada nube depende de infraestructuras físicas: centros de datos, redes, dispositivos, minerales, electricidad y agua.

El entrenamiento de los modelos requiere importantes recursos, pero también su utilización cotidiana genera consumos que, multiplicados por millones de interacciones, adquieren una dimensión considerable.

Los centros de datos afectan a los territorios donde se instalan, especialmente por su demanda de energía, agua y suelo, así como por el calor que generan.

La sostenibilidad debería incorporarse como una dimensión fundamental de la calidad tecnológica, junto con la robustez, la transparencia y la ética.

Esto implica desarrollar algoritmos más eficientes, reutilizar modelos, reducir su tamaño y elegir tecnologías proporcionadas a las necesidades reales.

8. No siempre necesitamos la inteligencia artificial más potente

Una de las propuestas más prácticas de las jornadas fue cuestionar la utilización indiscriminada de modelos generativos.

Muchas tareas pueden resolverse mediante tecnologías más sencillas, sistemas tradicionales de inteligencia artificial o recursos que ya existen.

La generación de imágenes, por ejemplo, puede requerir considerablemente más energía que determinadas interacciones de texto. Y no siempre necesitamos generar contenidos nuevos cuando podemos reutilizar materiales existentes.

La cuestión no consiste en renunciar a estas tecnologías, sino en valorar su utilidad en relación con sus costes económicos, sociales y ambientales.

No necesitamos utilizar inteligencia artificial para todo. Necesitamos utilizar la herramienta adecuada para cada tarea.

9. Los beneficios y los costes del desarrollo tecnológico están distribuidos de manera desigual

Las aportaciones desde Colombia y el Congo permitieron conectar el desarrollo de la inteligencia artificial con las cadenas materiales que lo sostienen.

Los dispositivos y las infraestructuras digitales necesitan minerales críticos, energía y procesos industriales que generan conflictos sociales y ambientales en determinados territorios.

Mientras los beneficios económicos y tecnológicos se concentran en algunas regiones y corporaciones, una parte importante de los costes se desplaza hacia otras comunidades.

Esta perspectiva amplía la conversación sobre la sostenibilidad. No basta con reducir el consumo energético de los algoritmos. **También necesitamos preguntarnos por las condiciones de extracción, las relaciones económicas y la capacidad de decisión de las comunidades afectadas.**

10. La gobernanza de la inteligencia artificial es una conversación sobre la sociedad que queremos construir

Una de las preguntas más relevantes de ambas jornadas fue quién decide qué valores deben incorporar los modelos de inteligencia artificial.

Eliminar sesgos no es una operación exclusivamente técnica. Lo que consideramos justo, adecuado o deseable depende también de nuestras interpretaciones culturales, posiciones políticas y acuerdos sociales.

Las empresas que desarrollan estas tecnologías toman decisiones que influyen en los contenidos que recibimos, las representaciones que generamos y las posibilidades de actuación que tenemos.

Por eso necesitamos ampliar los espacios de deliberación, investigación, regulación y participación ciudadana.

La cuestión de fondo no es únicamente qué puede hacer la inteligencia artificial, sino qué queremos que haga, al servicio de quién y con qué mecanismos de responsabilidad y control democrático.

Una reflexión final: recuperar capacidad colectiva de decisión

Las dos jornadas muestran que resulta insuficiente analizar la inteligencia artificial como una herramienta aislada. Necesitamos comprenderla dentro de un sistema de

relaciones que conecta a las empresas tecnológicas, las instituciones públicas, las organizaciones, las personas trabajadoras, las comunidades y los territorios.

Los sesgos se relacionan con nuestras estructuras sociales. La automatización del trabajo afecta a los procesos de aprendizaje y a la sostenibilidad del bienestar. La concentración de información modifica nuestra capacidad para interpretar la realidad. El desarrollo tecnológico depende de infraestructuras materiales cuyos impactos no se distribuyen de manera equitativa.

Desde esta perspectiva, uno de los mayores riesgos es perder progresivamente nuestra capacidad de comprender, cuestionar y decidir sobre tecnologías que están adquiriendo una presencia creciente en nuestras vidas.

Al mismo tiempo, las jornadas permitieron conocer investigaciones, experiencias institucionales, desarrollos tecnológicos propios y movimientos ciudadanos que están intentando intervenir sobre estas dinámicas.

Quizá el principal aprendizaje sea que no se trata solamente de adaptarnos a la inteligencia artificial, sino de **construir capacidades individuales, organizativas y colectivas para orientar su desarrollo.**

Y quedan algunas preguntas para seguir conversando: **¿qué capacidades queremos preservar y desarrollar como personas y organizaciones? ¿Qué dependencias estamos dispuestas a asumir? ¿Qué decisiones necesitamos mantener bajo control humano? ¿Dónde están apareciendo alternativas y formas de colaboración que podemos reforzar?**

En definitiva, hablar de inteligencia artificial es también hablar de poder, derechos, conocimiento, sostenibilidad y democracia. Y, sobre todo, de nuestra capacidad para participar en la construcción del futuro que queremos.