

Desafíos éticos de la IA



Del 26 al 28 de noviembre, tuvo lugar en la Universidad de Deusto el **Congreso Ahotsak. El papel de las ciencias sociales y humanas ante los retos del presente**, organizado por la Facultad de Ciencias Sociales y Humanas. Como señala la Decana, Ane Ferran, “un congreso donde queremos reivindicar nuestras voces: voces que explican la realidad, que contribuyen colaborativamente a la transformación y la justicia social” (ver el programa [aquí](#)).

Voy a recoger aquí algunas de las ideas que salieron en la mesa: **“Resituando (y gobernando) el entusiasmo: desafíos éticos de la irrupción de la IA”**. Tuvo como ponentes a: [María López](#) — Dra. en Derechos Humanos —, [Borja Sanz](#) — Dr. en Sistemas de Información —, [Ruth Carbajo](#) — Dra. en Ingeniería, Energía y Tecnologías de la Información — y [Laura Marrón](#) — Directora de Basque Artificial Intelligence Center ([BAIC](#)) —. Y como moderador, [Peru Sasía](#) — Dr. en Química Macromolecular —.

El punto de partida fue que si hiciéramos un análisis ético profundo no asumiríamos que la IA es inevitable y, mucho menos, ingobernable. Antes de hablar de los efectos conviene plantearse la pertinencia del uso y desarrollo de la IA. En primer lugar se presentaron algunos de los **principales desafíos éticos** de la irrupción de la IA:

- Debemos reflexionar no sólo sobre lo que podemos hacer, sino sobre **lo que debemos hacer**. La implantación de las innovaciones no siempre asume los retos de la cohesión social, la democratización, etc.
- Actualmente vivimos en un **tecnofeudalismo** — véase [Robledo \(2024\)](#) —, un capitalismo gobernado por un pequeño grupo de hombres — no hay mujeres CEO de las grandes empresas tecnológicas — que controla tanto el dinero como los datos, lo que provoca una profunda desigualdad social que amenaza la democracia. Además, se está dando una pérdida de relevancia de los marcos normativos. Los nuevos señores feudales necesitan que los marcos normativos les opriman lo menos posible.
- ¿Cómo se pueden diseñar algoritmos justos si no hay fórmulas a aplicar? Los tecnólogos se encuentran con sistemas complejos, muy complejos, de difícil

explicabilidad — no se sabe cómo funcionan y por qué lo hacen así—. Un caso sugerente es [Anthropic](#), empresa dedicada a la seguridad y la investigación en inteligencia artificial, que pretende crear sistemas de IA fiables, interpretables y controlables y que publican los fallos que encuentran en sus sistemas.

Vistos los desafíos, el siguiente paso fue plantear **los tratamientos, las medidas**, a aplicar:

- **Auditar los algoritmos** puede ayudar a identificar problemas, aunque esta medida sea a posteriori y su alcance no sea general. Es muy importante la pregunta por la trazabilidad del dato. La base de conocimiento de la IA es todo internet. [Esta intervención me recordó la charla que dio Gema Galdón hace dos años en Deusto Forum, [ver entrada](#)].
- **Poner a las personas en el centro**. Corremos el peligro de acabar pareciéndonos a las máquinas, en lugar de al revés.
- **Mapear** cómo y para qué se usa la IA y así valorar el nivel de riesgo. Y a partir de ahí hacer **planes de capacitación**, entendiendo que no sólo se trata de saber usarla, sino de conocer qué datos uso, de quién son, y preguntarse si se pueden compartir.
- **A las universidades nos corresponden algunas tareas fundamentales** como: desarrollar un posicionamiento crítico respecto de la IA, encontrar un idioma común desde la interdisciplinariedad, una reflexión profunda sobre la innovación y cómo esta debe revertir a la sociedad (democratización del sistema de innovación).
- No se puede dar un paso atrás en el ámbito normativo. **Hay que asegurar un marco normativo fuerte que proteja los derechos de todas las personas.**
- Hay que **generar consensos y transversalizar la perspectiva ética**: qué entendemos por sostenibilidad, por explicabilidad, incluir el concepto de rendición de cuentas.

Para terminar la mesa cada participante eligió una palabra o concepto. Suscribo todos ellos: potencial de la colaboración, perplejidad, responsabilidad compartida y militancia. No se puede ser un ciudadano, una ciudadana, responsable sin preguntarse y posicionarse ante los retos éticos de la IA. Y como dice el proverbio africano: "**Si quieres ir rápido, ve solo; si quieres llegar lejos, ve acompañado**".

Referencias

Anthropic <https://www.anthropic.com/company>

Basque Artificial Intelligence Center <https://baic.eus/>

Borja Sanz <https://www.linkedin.com/in/bosanz/>

Congreso Ahotsak <https://live.eventtia.com/es/ahotsak>

Echaniz, Arantza (2023, 27 noviembre). Ni inteligencia, ni artificial.

<https://echanizbarrondo.blogspot.com/2023/11/ni-inteligencia-ni-artificial.html>

Laura Marrón <https://www.linkedin.com/in/laura-marron/>

María López

<https://www.linkedin.com/in/mar%C3%ADa-l%C3%B3pez-belloso-13078b140/>

Peru Sasia <https://www.linkedin.com/in/pedro-m-sasia-096496216/>

Robledo, Ricardo (2024, 1 agosto). ¿Del capitalismo al tecnofeudalismo? Disensiones con Varoufakis.

<https://conversacionsobrehistoria.info/2024/08/01/del-capitalismo-al-tecnofeudalismo-disensiones-con-varoufakis/>

Ruth Carbajo <https://www.linkedin.com/in/ruthcarbajo/>

Etiquetas: ética, IA,