



Implicaciones Éticas, Jurídicas, Económicas y Sociales de la IA

DANIEL LÓPEZ CASTRO (INSTITUTO DE FILOSOFÍA, CSIC)

Dos sesiones

TEÓRICA

PRÁCTICA

¿Por qué **filosofía**?

- ❑ Amor por las preguntas, curiosidad (Emilio Lledó)
- ❑ Puente entre las ciencias y las humanidades
 - ❑ Relación bidireccional – Es posible...
 - ❑ ¿Hacer ciencia sin preguntarse las implicaciones sociales de los avances tecnocientíficos?
 - ❑ ¿Hacer filosofía sin conocer los avances en neurociencia, psicología o inteligencia artificial?
- ❑ Incertidumbre y falta de seguridad
 - ❑ Riesgos: consecuencias imprevisibles
 - ❑ Oportunidades: descubrimiento y creación
- ❑ No se pueden plantear como cuestiones estrictamente científicas (lo que sabemos o nos falta por saber), sino que también son cuestiones de preferencia, cultura y valores.

¿Por qué **ética aplicada**?

- Minimización del daño (seguridad y protección)
- Beneficencia o “maximización del beneficio”
- Respeto y autonomía
- Inclusión y justicia?
- Medio ambiente y generaciones futuras
 - Futuro del trabajo
 - Cambio climático y colapso energético

¿Por qué **ética** de la tecnología?

- ❑ Importancia de la tecnología
 - ❑ ¿Qué es?
 - ❑ ¿Un instrumento?
 - ❑ Condición humana → Condición tecnológica

- ❑ No neutralidad
 - ❑ Moralidad y política

¿Por qué **ética** de la información?

Cómo las TICs **afectan** a:

- nuestra concepción del yo (quiénes somos),
- nuestras relaciones (cómo socializamos),
- nuestra concepción de la realidad (nuestra metafísica) y
- nuestras interacciones con la realidad (nuestra agencia) (Floridi 2015).

¿Por qué **ética de la información**?

Esto tiene lugar a través de cuatro **transformaciones**:

- ❑- la difuminación de la frontera entre realidad y virtualidad;
- ❑- la difuminación de la frontera entre humano, máquina y naturaleza;
- ❑- el paso de la escasez de información a la sobreabundancia de la misma;
- ❑- y el paso de la primacía de objetos independientes, propiedades y relaciones binarias, a la primacía de interacciones, procesos y redes (Floridi 2015)

❑ Somos **INFORGS**, embebidos en la **INFOESFERA**

¿Por qué **ética** de la IA?

- Narrativas políticas
 - EEUU → Eficiencia empresarial
 - China → Control
 - Europa → Ética
- Privacidad
- Sesgos
- Responsabilidad
- Explicabilidad

Privacidad

- Problema desde los inicios de internet
- Propiedad → ¿bienes comunes?
- Caso Schrems (I, II, ¿y III?)
- Asimetrías de poder (transparencia vs vigilancia)
- Vulnerabilidad

Privacidad (mente extendida y neuroderechos)



- Condiciones: criterio de constancia, disponibilidad, endoso + criterio de individualización y afianzamiento (añadido por Sterelny, 2010)
- Delegación de procesos cognitivos: memoria, orientación geográfica
- Delegación de decisiones: Restaurantes, amistades, ¿pareja?
- ¿Qué delegar y qué no? ¿A quién delegar? ¿Con qué garantías?

- Neuroderechos

Sesgos

- Sesgo de automatización
- Recomendación de decisiones injustas o poco equitativas.
- Una Fuente de sesgos: los datos.
- También los sesgos vienen de nosotros: si hay sesgo dentro, habrá sesgo fuera.
- ¿Perpetuación de sesgos y denegación de futuro?
- Correlaciones espúreas
 - Confusión de correlación con causalidad

Responsabilidad

- Responsabilidad de actuar → responsabilidad moral
- IA – Capacidad de agencia, pero no responsabilidad
- Problema de muchas manos y muchas cosas
- Dificultad de señalar responsable cuando algo sale mal
- Problema del control
- Problema del conocimiento

Explicabilidad

- Deep Learning / Redes neuronales
- Problema de la caja Negra
- Relación con la responsabilidad
- Falta de transparencia → menos confianza y posibilidad de aceptación y apropiación social.
- Implica:
 - Derecho a una explicación comprensible

Ejemplos prácticos

- Civio – Bosco – Bono Social Eléctico
- COMPAS – ProPublica
- Donación de órganos

- Muchos ejemplos en *Armas de Destrucción Matemática* (Cathy O’Neil, 2016)

Hacia una ética (también) positiva



iGracias!

DANIEL.LOPEZ@CSIC.ES