
ÉTICA, LEGISLACIÓN Y PROFESIÓN

“‘APRENDER’” A HACER TU TRABAJO DE UNA FOMRA ÉTICA Y LEGAL

ALEJANDRO BARRACHINA ARGUDO

GRADO EN INGENIERÍA INFORMÁTICA
FACULTAD DE INFORMÁTICA
UNIVERSIDAD COMPLUTENSE DE MADRID

Índice general

1. Introducción a la Ética y la Legislación	5
1.1. Introducción a la Ética	5
1.2. Introducción a la Legislación	5
1.3. Cuestiones éticas relacionadas con el diseño de aplicaciones informáticas	6
1.3.1. Manipulación online	6
1.3.2. El salto a la IA (Inteligencia Artificial)	7
1.3.3. Guía ética para una IA confiable (UE (Unión Europea))	7
Glosario	9
Índice de figuras	10
Índice de cuadros	11

1 | Introducción a la Ética y la Legislación

1.1. Introducción a la Ética

La ética es una parte de la filosofía, la cual implica cuestionarse sobre problemas relativos al bien, al deber, a la virtud o al vicio. Los objetivos de la ética son:

1. Llegar a alguna conclusión acerca de lo que es correcto o no
2. Que la conclusión pueda ser defendida con argumentos.
3. Justificar las normas que regulan el comportamiento en diferentes ámbitos de la vida

La reflexión es la consideración minuciosa de un asunto, esto implica dedicarle tiempo y no quedarse con la primera impresión. Para llevar a cabo esta consideración minuciosa debemos:

- Informarnos: contrastar fuentes y poder citarlas.
- Pros y contrastar.
- Partes afectadas.
- Circunstancias.
- Posibles escenarios futuros.
- Contrastar: comentar, dialogar, escuchar otros puntos de vista diferentes al propio.
- Discernir, analizar

Es importante hacerse muchas preguntas para poder reconocer y tener en cuenta a todas las partes implicadas, dirigir esfuerzos a la hora de buscar información, ampliar nuestra visión del tema y compartir o hacer esas preguntas a otras personas para ampliar nuestra perspectiva. Los elementos claves para una reflexión ética son:

1. Identificar los valores del conflicto.
2. Empatía: personas/colectivos afectados: ponerse en el lugar de los otros, percibir a los demás como seres sintientes.
3. Analizar las circunstancias.
4. Pros, contras y su peso.
5. Alcance de las consecuencias en el espacio y el tiempo.
6. Posibles escenarios futuros partiendo del presente.

Tras una reflexión ética debemos llegar a una conclusión clara, directamente relacionada con los argumentos y ejemplos reales utilizados. Esta conclusión debe ser coherente con los argumentos y ejemplos usados. Si esta conclusión está matizada demostrará un correcto proceso de reflexión.

1.2. Introducción a la Legislación

El derecho es la técnica de dar a cada persona lo que le corresponde. También es un conjunto de normas organizadas, escritas o no, cuyo cumplimiento es obligatorio y pueden ser impuestas de forma coactiva, que sirven para asegurar la pacífica convivencia.

El Código Civil en su art. 1.1 señala como fuentes del Derecho:

1. La Ley: norma vigente.
2. La costumbre: conductas repetidas desde tiempos inmemoriales. Solo rige en defecto de ley aplicable y siempre que no contradiga la moral o al orden público y tiene que ser probada.

3. Los PGD (Principios Generales del Derecho): principios fundamentales. Se aplican en defecto de la Ley o costumbre. Además, su contenido está siempre presente en el ordenamiento jurídico.

Estas fuentes de derecho tienen una jerarquía, lo que hace que una disposición no carezca de valor si contradice a una de rango superior. Ley > costumbre > PGD. La legislación refleja una serie de valores (Derechos, protección) y una serie de conflictos (delitos, sanción, penas).

1.3. Cuestiones éticas relacionadas con el diseño de aplicaciones informáticas

Tanto el diseño como el software son políticos, ya que podemos considerar que “Los algoritmos son opiniones incrustadas en código”.

El diseño de un software no es neutro, ya que se tienen que considerar distintos puntos al hacer el programa:

- Qué función o funciones realiza.
- Qué restricciones ponemos al usuario y cuales no.
- Qué datos vamos a recolectar y/o analizar.
- Qué podemos deducir sobre los usuarios.
- Qué mecanismos se van a utilizar para crear o no adicción entre los usuarios.

Un ejemplo de software político puede ser BOSCO, una aplicación financiada por el Gobierno Español utilizada por las compañías eléctricas para decidir quién tiene derecho a un descuento en la factura de la luz.

Otro ejemplo sería el sistema VeriPol utilizado por la Policía Nacional para detectar denuncias falsas.

1.3.1. Manipulación online

Algunos de los casos más conocidos son: Facebook con su publicidad dirigida a adolescentes estresados y deprimidos, Uber con estrategias para conseguir que sus conductores trabajen más horas o en determinadas zonas y Cambridge Analytica y sus mensajes políticos personalizados basados en perfiles de Facebook.

Dentro de la manipulación online podemos ver varios tipos:

- **Persuasión (influencia):** cualquier forma de influencia e influencia a través de la discusión racional.
- **Persuasión en sentido explícito:** es visible, consciente y resistible (hay alternativas).
- **Manipulación:** es oculta, inconsciente y explota vulnerabilidades cognitivas emocionales y estructurales o individuales. La víctima no es consciente de que ha sido dirigida a una decisión determinada. Hay un beneficio para la parte manipuladora.
- **Manipulación + TIC (Tecnologías de la información y la comunicación):** hay grandes cantidades de información personal y algoritmos que analizan estos datos que son invisibles y omnipresentes. Generan campañas individualizadas con anuncios y mensajes a medida.

Estas acciones pueden producir daños materiales como gastar dinero en cosas que no necesitamos o gastar más de lo que queríamos o daños en cuanto a violación de la autonomía, con impacto individual y/o social y político.

Al estudiar un programa tenemos que hacernos las siguientes preguntas:

- Quién decide qué y cómo se diseñan.
- Si tiene mayor o menor potencial manipulador.
- Si está diseñada para ser adictiva.
- Si está diseñada para potenciar el comportamiento impulsivo y disminuir la capacidad analítica.
- Si se da al usuario una falsa sensación de autonomía.

1.3.2. El salto a la IA

La evolución de la IA nos ha traído sistemas muy complejos con aprendizajes profundos (encuentran patrones en un conjunto de datos) que sacan conclusiones estadísticas y son fáciles de engañar.

Al ser entrenadas mediante conjuntos de datos, ¿Pueden sus diseñadores predecir su comportamiento? Algunas arquitecturas de IA funcionan como una caja negra en la que no sabemos el porqué de las decisiones de la máquina, aunque no todas las arquitecturas son tan opacas.

Hay que tener en cuenta el hecho de que el sesgo cognitivo y los prejuicios de los desarrolladores de una IA pueden influir de manera negativa en su toma de decisiones. Estos sesgos pueden ser incoscientes o heredados (por el diseño o por el conjunto de datos que se usa para el entrenamiento) o prejuicios deliberados que nazcan de una determinada política o idea.

Para evitar IAs con este tipo de problema tenemos que buscar siempre que el algoritmo sea explicable.

Cómo asegurar que el algoritmo es justo, cómo asegurar que el algoritmo es interpretable y explicable: todo eso está todavía bastante lejos.

Nihar Shah

Es de interés para las personas que desarrollan el algoritmo, para abogados y jueces y para toda la ciudadanía entender la toma de decisiones de un algoritmo.

Parte de evitar problemas de bias viene en forma de auditorías, para las cuales hay que tener acceso a:

- Diseño: explicabilidad, trazabilidad y reproducibilidad.
- Datos de entrenamiento o conjunto de datos para probar el sistema.
- Código fuente

1.3.3. Guía ética para una IA confiable (UE)

To everyone who shapes technology today
We live in a world where technology is consuming society, ethics, and our core existence.
It is time to take responsibility for the world we are creating. Time to put humans before business. Time to replace the empty rhetoric of building a better world with a commitment to real action. It is time to organize, and to hold each other accountable.

The Copenhagen Letter, 2017

La UE define una IA fiable como una que es lícita, ética y robusta. Una IA fiable debe respetar cuatro principios éticos (a parte de respetar los derechos fundamentales):

- Respeto de la autonomía humana.
- Prevención del daño.
- Equidad.
- Explicabilidad.

Una IA fiable debe cumplir también siete requisitos clave:

- Acción y supervisión humanas.
- Solidez técnica y seguridad.

- Transparencia.
- Diversidad, no discriminación y equidad.
- Bienestar social y ambiental.
- Rendición de cuentas.

Para aplicar estos principios hay que hacerlo desde el propio diseño del algoritmo: poniendo el respeto al ser humano por encima de cualquier otro tipo de interés como centro del diseño, respetar la autonomía individual y colectiva, no discriminar, revisar los sesgos, hacer un algoritmo explicable y transparente, respetar la privacidad y los DDAA (derechos de autor) (licencia y opciones de protección de los DDAA) y prever y tratar de evitar posibles usos fraudulentos y/o delictivos del algoritmo.

Glosario

DDAA derechos de autor

IA Inteligencia Artificial

PGD Principios Generales del Derecho

TIC Tecnologías de la información y la comunicación

UE Unión Europea

Índice de figuras

Índice de cuadros