



TEKgunea release minor 3.7

AA Administrazio publikoan. IA en la Administración Pública.

Retos éticos en el uso de la IA avanzada en la Administración Pública

¿QUÉ SIGNIFICA
PENSAR
ÉTICAMENTE?

Bien

Principios

Fines

Derechos,
normas...

Virtudes,
valores

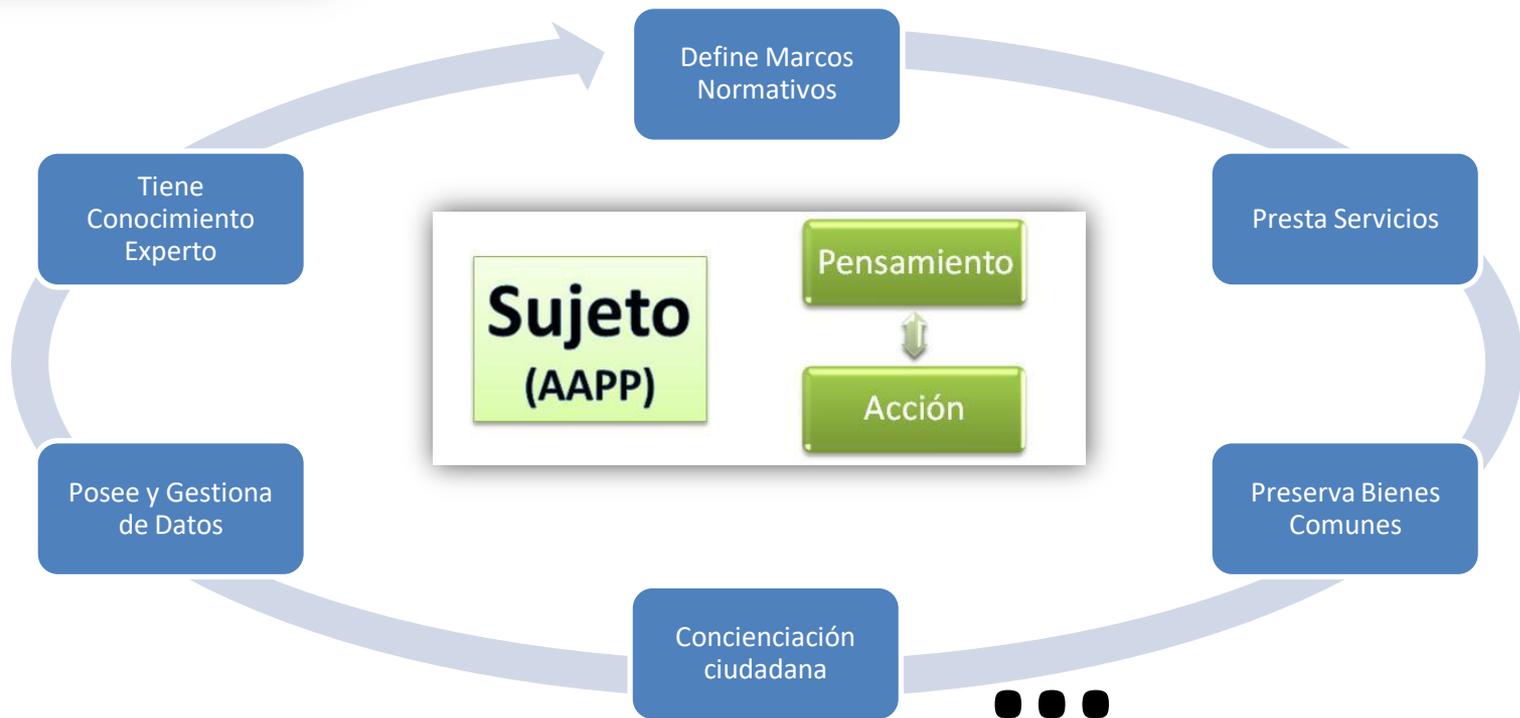
Procesos,
participación

Utilidades,
bien común

**Sujeto
(AAPP)**

Pensamiento

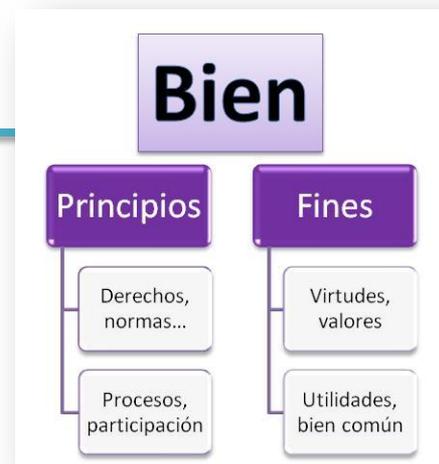
Acción



ÉTICA e IA: ¿Qué preguntas nos hacemos?

Una primera aproximación ética:
¿Qué es una **BUENA** decisión?

- ***RESPETA DERECHOS***
- ***LO HEMOS ACORDADO***
- ***CONTRIBUYE AL BIEN COMÚN***
- ***RESPONDE A UNOS VALORES***



ÉTICA e IA: ¿Qué preguntas nos hacemos?

• RESPECTA DERECHOS

- ✓ Salud
- ✓ Vida digna
- ✓ Información
- ✓ Privacidad
- ✓ Propiedad
- ✓ ...

Bien

Principios

Derechos,
normas...

Procesos,
participación

Fines

Virtudes,
valores

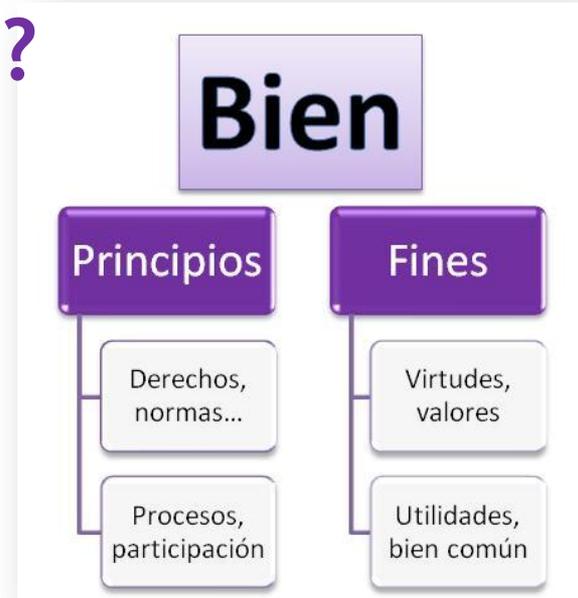
Utilidades,
bien común



ÉTICA e IA: ¿Qué preguntas nos hacemos?

- ***LO HEMOS ACORDADO***

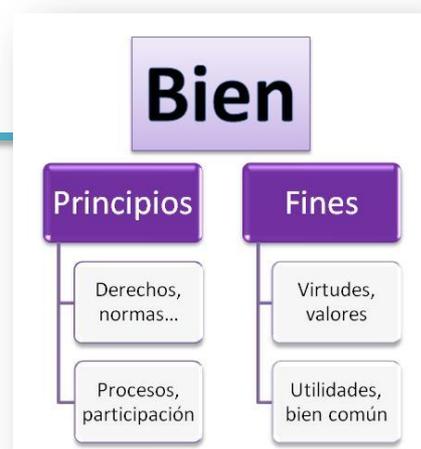
- ✓ ¿Todos los afectados/as tienen voz?
- ✓ ¿Se verifica el cumplimiento?
- ✓ ¿Cómo se decide?
- ✓ ¿Quién decide?
- ✓ ...



ÉTICA e IA: ¿Qué preguntas nos hacemos?

• **CONTRIBUYE AL BIEN COMÚN**

- ✓ ¿Conozco las consecuencias?
- ✓ ¿A quién beneficia?
- ✓ ¿A quién perjudica?
- ✓ ¿Gestiono los riesgos?
- ✓ ¿Hay víctimas?
- ✓ ¿Es sostenible?...



OBJETIVOS DE DESARROLLO SOSTENIBLE



ÉTICA e IA: ¿Qué preguntas nos hacemos?

- **RESPONDE A UNOS VALORES**

- ✓ Solidaridad
- ✓ Cooperación
- ✓ Transparencia
- ✓ Honestidad
- ✓ Austeridad
- ✓ Sobriedad
- ✓ ...



... centrándonos en las AAPP, es interesante observar cómo, considerando que la sensibilidad ética constituye una competencia naturalmente insertada en su cultura, las primeras reacciones han sido, en muchos casos, de perplejidad, ante un fenómeno que, como decimos, ha irrumpido de forma vertiginosa...

... no es, por lo tanto, extraño, que las respuestas iniciales, a la espera de marcos jurídicos que delineen un perímetro normativo adecuado para poder incorporar la IA en su quehacer diario con las suficientes garantías éticas, hayan sido de una naturaleza que podemos considerar esencialmente defensiva, asumiendo estrategias de mitigación de posibles daños o riesgos éticos específicos que su uso pueda generar.

Estrategias de mitigación

- Políticas
- Mecanismos de estandarización (*)
- Mecanismos de Certificación (*)

() riesgo de desplazamiento*

***Guías y Códigos
(documentos
de pautas
éticas)***

***Medidas específicas de
Gobernanza***

Una
“*cartografía ética*” ...

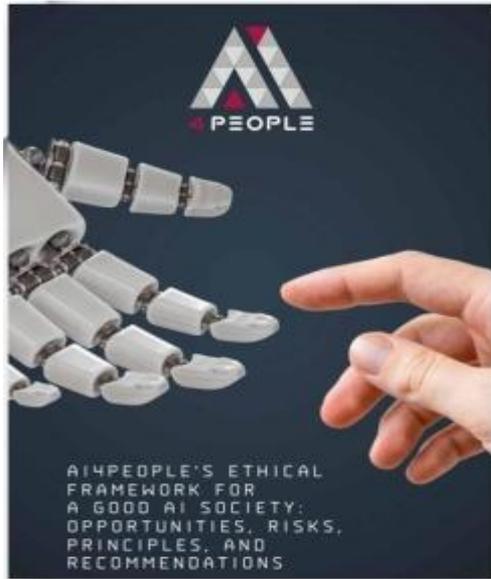
*Guías y Códigos
(documentos
de pautas
éticas)*

... con muy pocas “*señalizaciones*”

Uno de los principales desafíos éticos de la inteligencia artificial (IA) **es la forma en que se utiliza y se implementa**. Estos son algunos de los principales retos:

- 1. Sesgo de datos:** Los sistemas de IA pueden ser entrenados con conjuntos de datos que contienen prejuicios y discriminación, lo que puede generar decisiones erróneas y sesgadas.
- 2. Privacidad y seguridad:** La IA recopila grandes cantidades de datos personales, lo que puede poner en riesgo la privacidad y seguridad de las personas.
- 3. Responsabilidad y accountability:** En situaciones en las que la IA toma decisiones, no siempre es claro quién es el responsable y quién debe asumir la responsabilidad.
- 4. Transparencia:** Es importante que los sistemas de IA sean transparentes y que las decisiones tomadas por la IA sean explicadas de manera comprensible para las personas.
- 5. Impacto en el empleo:** A medida que los sistemas de IA y automatización se vuelven más eficientes, podrían reemplazar a ciertos trabajos y provocar desempleo.
- 6. Responsabilidad social:** En general, la IA debe ser utilizada para mejorar la vida humana y no para dañarla, lo que significa que se deben considerar los posibles efectos secundarios en la sociedad.





Buen uso de la IA

Ayuda a la autorrealización
del ser humano

Refuerza la agencia
humana

Aumenta las capacidades
de la sociedad

Aumenta la cohesión social

Mal uso de la IA

Debilita las capacidades
humanas

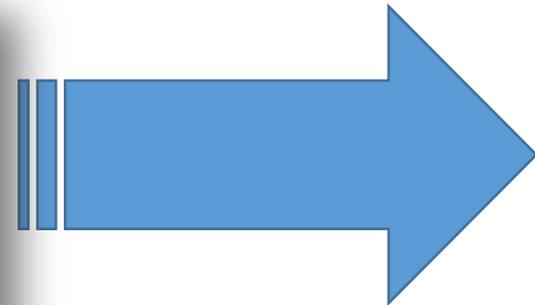
Elimina la responsabilidad
humana

Reduce el control humano

Erosiona la auto
determinación

Estrategias de mitigación: Guías y Códigos

Del análisis de 84 documentos de Guías y Códigos de ética en la IA aparecen **once principios y valores éticos** dominantes a tener en cuenta



Estrategias de mitigación: Guías y Códigos → Principios y Valores

- ✓ **Transparencia**
- ✓ **Justicia, imparcialidad, equidad**
- ✓ **No maleficencia**
- ✓ **Responsabilidad y Rendición de cuentas**
- ✓ **Privacidad**
- ✓ **Beneficencia**
- ✓ **Libertad y autonomía**
- ✓ **Confianza**
- ✓ **Sostenibilidad**
- ✓ **Dignidad**
- ✓ **Solidaridad**

Estrategias de mitigación: Guías y Códigos → Principios y Valores

Comentarios generales

- De los once grupos de principios éticos, los referentes a **transparencia**, **justicia**, **no-maleficencia**, **responsabilidad** y **privacidad** aparecen en más de la mitad de los documentos.
 - La **transparencia**, como “posibilitadora o perjudicadora de otras practicas éticas o principios” explica su prevalencia.
 - La **justicia**, **no-maleficencia**, **responsabilidad** y **privacidad** puede entenderse como una llamada de cautela a la comunidad global contra el potencial riesgo del al IA.
- Las referencias a la **no-maleficencia** son mayores que las de **beneficencia** demuestra una preocupación con la obligación moral de prevenir el daño

“Principio de precaución”

Estrategias de mitigación: Guías y Códigos → Principios y Valores

Comentarios generales

- La sostenibilidad, dignidad y solidaridad están significativamente poco representadas comparadas con las otras dimensiones éticas.
- **Todavía hay mucha incertidumbre sobre que principios éticos deben priorizarse y como se resuelve el conflicto entre principios.**



Beneficencia: promover al bienestar, preservar la dignidad, sostener el planeta

No-maleficencia: privacidad, seguridad y “principio de precaución”

Autonomía: el poder de decidir, o si queremos decidir

Justicia: promover la prosperidad y preservar la solidaridad

Explicabilidad: hace posibles los otros principios a través de la inteligibilidad y la rendición de cuentas

**Estrategias de mitigación para un uso responsable de
la IA en las Administraciones Públicas:**

Prioridades

Estrategias de mitigación para un uso responsable de
la IA en las Administraciones Públicas:

Prioridades

Generar confianza en la sociedad sobre el desarrollo o uso de la IA:

- realización de auditorías externas independientes
- asunción de códigos o certificaciones
- rendición de cuentas clara y transparente
- establecimiento de mecanismos de protección de aquellos profesionales que identifiquen malos comportamientos (*whistleblowing*)
- compromiso con que los instrumentos utilizados sean comprendidos y explicables.

Estrategias de mitigación para un uso responsable de
la IA en las Administraciones Públicas:

Prioridades

Gestión de los impactos que el uso de la IA puede generar:

no solo debemos preguntarnos por las oportunidades de la IA. Es también importante preguntarse a quién afectará y de qué manera. El uso de la IA en las AAPP puede generar impactos muy distintos en muchas personas y colectivos.

Por lo tanto, es importante resaltar el uso de lo que podemos considerar **antenas éticas de vigilancia**, que permitan a la organización adelantarse a posibles impactos negativos y gestionarlos convenientemente, con especial atención a las brechas y a la vulnerabilidad.

Estrategias de mitigación para un uso responsable de
la IA en las Administraciones Públicas:

Prioridades

**Exigencia de Responsabilidad de los retos que enfrenta el conjunto de la
sociedad respecto a la IA:**

- Colaboración en el establecimiento de marcos normativos adecuados que protejan el bien común
- Relación de cooperación con los diferentes *stakeholders*
- Orientación de sus recursos y conocimiento experto al desarrollo de una IA segura

Una gobernanza ética de la IA en las Administraciones Públicas:

Algunas alertas

Una gobernanza ética de la IA
en las Administraciones Públicas:

Algunas alertas

Procesos de toma de decisión:

la existencia de mecanismos basados en algoritmos excesivamente automatizados (y, lo que es más importante, empoderados) puede conducir a una cesión (intencionada o no) de funciones indelegables por parte de las AAPP, renunciando (conscientemente o no) a sus responsabilidades.

Una gobernanza ética de la IA
en las Administraciones Públicas:

Algunas alertas

Mecanismos de análisis de la realidad:

basados en dinámicas automatizadas de monitorización, se corre el riesgo de producir alertas sesgadas o inadecuadamente interpretadas. Esto puede generar dificultades para un adecuado sistema de vigilancia ética que, como indicábamos anteriormente, constituye un elemento imprescindible entre las estrategias de gestión de la IA, que basa su fiabilidad en una honesta y afinada lectura e interpretación de la realidad.

Una gobernanza ética de la IA
en las Administraciones Públicas:

Algunas alertas

Procesos de participación:

que involucren especialmente a los *stakeholders* más afectados por el uso de la IA. Se enfrenta a los conocidos problemas que genera la complejidad de interacciones humano-máquina, debilitando varias de las condiciones que exige una justa participación de estos grupos de interés en los procesos de toma de decisión. En este aspecto, es necesario recordar que dichas interacciones son normalmente generadas (al menos de momento) bajo las reglas que impone la propia naturaleza de la máquina, como, por ejemplo, la inmediatez y el intercambio basado en respuestas sencillas, habitualmente sobre la base de opciones predeterminadas.

Algunas cautelas éticas respecto a la IA

Cuatro pre-concepciones sobre los avances tecnológicos

- **Son benignos:** tendemos a pensar que más/nueva tecnología es siempre **bueno**.
- **Son espectaculares:** son algo ajeno a nuestra intención y/o decisión. Una especie de nueva **magia**.
- **Son inevitables:** el avance no puede pararse, cambiarse o redirigirse. Las nuevas tecnologías acabarán **emergiendo**.
- **Son moralmente neutrales:** la tecnología en sí **no es ni mala ni buena** (si algo malo pasa, es debido a un uso incorrecto o inapropiado).



La tentación de la inocencia tecnológica

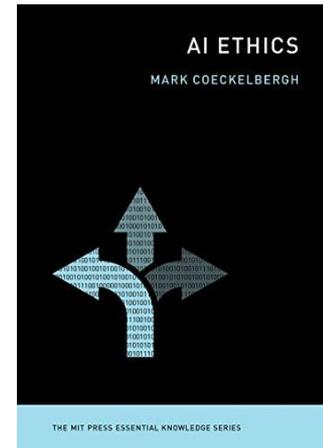
 Peru Sasia | 9 abril, 2018

ÉTICA EMPRESARIAL

ÉTICA Y TECNOLOGÍA



“En el trasfondo de la discusión sobre la IA hay grandes desacuerdos sobre la naturaleza del ser humano, la inteligencia humana, la mente, el entendimiento, la consciencia, la creatividad, el significado, el conocimiento humano, la ciencia y así sucesivamente. Si de verdad se trata de una batalla esta concierne tanto a lo humano como a la IA”



Marc Coeckelbergh