

**Universitat Oberta de Catalunya**

**Jornada de estudio: El futuro marco regulatorio de la Inteligencia Artificial en la Unión Europea**

**Barcelona, 18 de noviembre de 2022**

**EVALUACIÓN DE IMPACTO EN DERECHOS HUMANOS DE LOS SISTEMAS DE INTELIGENCIA ARTIFICIAL**

Alessandro Mantelero

Universidad Politécnica de Turín | Cátedra Jean Monnet de Derecho y Sociedades Digitales del Mediterráneo



© Mantelero 2022

Project: 101047818 - DIGIMED - ERASMUS-JMO-2021-CHAIR

< | >

Desde la protección de datos como 'the law of everything' al AI Act



## The Risk Society and AI

- ✓ El papel sustitutivo que durante muchos años ha jugado la protección de datos
- ✓ La protección de datos como derecho habilitante pero... no es 'the law of everything'
  - Variedad de impactos de la IA en la sociedad y los individuos
  - Poca experiencia en la evaluación de los derechos humanos en las DPIAs
  - Casos de impacto sobre los derechos humanos que no tienen que ver con el tratamiento de datos personales (p.ej., IA para la gestión de infraestructuras críticas)
- ✓ Dimensión global
  - Escenario tecnológico (grandes empresas, diferentes niveles de desarrollo, etc.)
  - Escenario geopolítico (manipulación de la opinión pública, ciberguerra, etc.)
  - Necesidad de evitar que se creen normas de facto
- ✓ De los remedios ex post (casos individuales) a la regulación preventiva general (riesgo)



## Un enfoque basado en el riesgo: un difícil equilibrio

- ✓ Política industrial y regulación de la innovación tecnológica (por ejemplo, las primeras leyes de protección de datos)
- ✓ Modelo basado en el riesgo y variedad de enfoques del riesgo
  - Riesgo industrial (producto y producción): seguridad, cumplimiento y estándares
  - Riesgos que afectan a las personas y a la sociedad
    - Evaluación de riesgos y beneficios
    - Evaluación centrada en los derechos humanos
- ✓ Al Act: una respuesta "proporcionada" al riesgo (aplicaciones de alto riesgo y aceptabilidad social)
- ✓ Niveles de riesgo y cuestiones de clasificación (p. ej., las amplias categorías del anexo III)



## Evaluación y diseño

- ✓ Necesidad de desarrollar modelos de evaluación de riesgos
- ✓ Modelos centrados en la medición del riesgo (DPIA, HRESIA, evaluación de los riesgos de ciberseguridad) frente a modelos basados en recomendaciones (por ejemplo, la tradicional HRIA)
- ✓ Papel clave de la participación ciudadana y necesidad de metodologías adecuadas
- ✓ Enfoque holístico y diseño de una IA verdaderamente centrada en el ser humano
- ✓ Papel limitado de los estándares y de los organismos de normalización en el ámbito de los derechos humanos (p. ej., calidad de los datos, determinados procedimientos de detección de sesgos)



< II >

**Evaluar el impacto sobre los derechos humanos: una propuesta concreta**



## HRIA en la IA

- No es la tradicional HRIA, centrada en la concienciación/el apoyo a la toma de decisiones
  - Enfoque en el producto/servicio de IA (poca relevancia de la dimensión local/territorial)
  - Enfoque previo y due diligence
  
- Regulación de la IA: énfasis en umbrales de riesgo cuantificables
  - Aceptabilidad de los productos/servicios de IA
  - Transparencia en la determinación de los casos de alto riesgo
  - Obligaciones legales y responsabilidad (ya no es una herramienta voluntaria)
  - Comparación entre las posibles opciones de diseño

## Un modelo de evaluación del riesgo

- Identificación de riesgos (HRs), probabilidad (L) y gravedad (S)
  - Probabilidad
    - Probabilidad de consecuencias desfavorables
    - Exposición
  - Gravedad
    - Gravedad del perjuicio
    - Esfuerzo para superar el perjuicio causado y revertir los efectos desfavorables



- Una escala de cuatro niveles (bajo, medio, alto, muy alto) para evitar cualquier riesgo de posicionamiento medio
- Uso de escalas cardinales (combinaciones y no multiplicaciones)
- Evaluación basada en expertos



Tab. 2 Probability

	Probability	
Low	The risk of prejudice is improbable or highly improbable	1
Medium	The risk may occur	2
High	There is a high probability that the risk occurs	3
Very high	The risk is highly likely to occur	4

Tab. 3 Exposure

	Exposure	
Low	Few or very few of the identified population of rights-holders are potentially affected	1
Medium	Some of the identified population are potentially affected	2
High	The majority of the identified population is potentially affected	3
Very high	Almost the entire identified population is potentially affected	4

Tab. 4 Likelihood table(L)

		Probability			
		1	2	3	4
Exposure	1	1	2	3	4
	2	2	3	5	9
	3	3	5	9	12
	4	4	7	12	15

Tab. 4 Gravity of the prejudice

	Gravity of the prejudice	
Low	Affected individuals and groups may encounter only minor prejudices in the exercise of their rights and freedoms.	1
Medium	Affected individuals and groups may encounter significant prejudices.	2
High	Affected individuals and groups may encounter serious prejudices.	3
Very high	Affected individuals and groups may encounter serious or even irreversible prejudices.	4

Tab. 5 Effort to overcome the prejudice and to reverse adverse effects.

	Effort	
Low	Suffered prejudice can be overcome without any problem (e.g. time spent amending information, annoyances, irritations, etc.)	1
Medium	Suffered prejudice can be overcome despite a few difficulties (e.g. extra costs, fear, lack of understanding, stress, minor physical ailments, etc.).	2
High	Suffered prejudice can be overcome albeit with serious difficulties (e.g. economic loss, property damage, worsening of health, etc.).	3
Very high	Suffered prejudice may not be overcome (e.g. long-term psychological or physical ailments, death, etc.).	4

Tab. 6 Severity table (S)

		Gravity			
		1	2	3	4
Effort	1	1	2	4	6
	2	2	3	5	8
	3	3	5	8	10
	4	5	8	10	12

Tab. 7. Table of envisaged risks

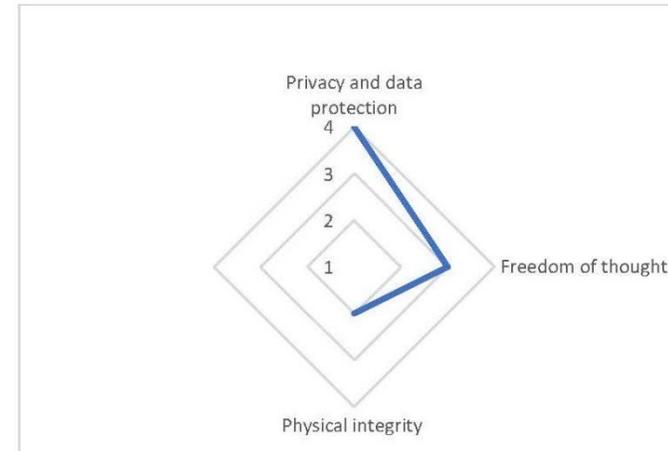
	L	S	overall impact
R1			
R2			
....			
Rn			

Tab. 8. Overall risk impact table

		Severity [impacted right/freedom]			
		Low	Medium	High	Very high
Likelihood	Low				
	Medium				
	High				
	Very high				

- Planificación y alcance basados en preguntas
- Enfoque participativo
- Gráfico radial
- Proceso circular

Radial graph (impact)



- 1 Low impact
- 2 Medium impact
- 3 High impact
- 4 Very high impact

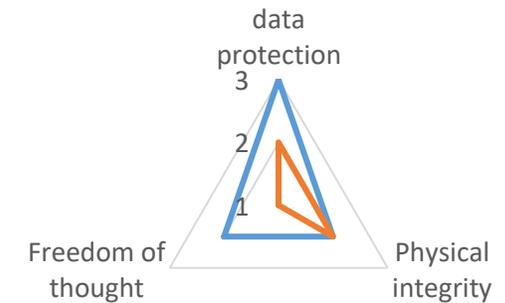
Tab. 9. Comparative risk impact analysis table (before/after mitigation measures and excluding factors)

	L	S	Overall impact	EFs	MMs	rL	rS	Final Impact
R1								
R2								
....								
Rn								



- Evaluación (efectos de las medidas adoptadas)

Riesgo	L	S	Overall impact	MMs	rL	rS	Final impact
Impacto en la privacidad y la protección de datos	VH	M	H	Yes	M	M	M
Impacto en la libertad de pensamiento	M	L	M	Yes	L	L	L
Impacto en el derecho a la seguridad psicológica y física	L	M	M	Yes	L	M	M



1 Impacto bajo  
 2 Impacto medio  
 3 Impacto alto





**Alessandro Mantelero**  
**alessandro.mantelero@polito.it**  
**@mantelero**



**Politecnico  
di Torino**

**Jean Monnet Chair in Mediterranean  
Digital Societies and Law**



**Co-funded by  
the European Union**

**Project: 101047818 - DIGIMED - ERASMUS-JMO-2021-CHAIR**



**Open Access**  
[https://link.springer.com/book/10.1007/  
978-94-6265-531-7](https://link.springer.com/book/10.1007/978-94-6265-531-7)