

Thesis title (it may be longer than one line if necessary)

Inés Sallent Ysamat

---



Universitat  
Pompeu Fabra  
*Barcelona*

[TÍTOL DEL TREBALL]

TREBALL FI DE GRAU DE  
Inés Sallent Ysamat

Director: Vladimir Estivill Castro  
Grau en Enginyeria en Sistemes Audiovisuals  
Curs 2024-2025



Universitat  
Pompeu Fabra  
Barcelona

Escola  
d'Enginyeria

*[Pàgina en blanc]*

*[Tingueu en compte:*

*(i) No s'han de deixar pàgines blanques al cos del treball (index fins la conclusió).*

*(ii) Utilitzar interlineat senzill i tipus de lletra Times New Roman.*

*(iii) Per a aquesta secció principal (cos del treball), la restricció del límit de pàgines és de 30 pàgines + 5% o 10% de marge.*

*(iv) La pàgina del final de capítol compta com a mitja pàgina.*

*(v) Podeu afegir un annex d'informació de suport a l'apèndix per incloure totes les dades/informació que necessiteu mostrar. El apèndix no té un límit de pàgines. Si tens informació a l'apèndix, has de fer referència en el text principal a la informació continguda en l'apèndix seguint aquest exemple:*

*document principal: "La taxa de natalitat als països europeus està disminuint molt ràpidament (vegeu Apèndix-1). Això vol dir que tot i que la gent viu més temps (figura Apèndix-1), amb el temps la mida de la població disminuirà".*

*(vi) Les referències s'inclouen a la secció de Bibliografia i s'exclouen de la limitació del nombre de pàgines. ]*

Dedicatòria [opcional] [mida 12]  
Lorem ipsum

*[Pàgina en blanc]*

## Agraïments[mida 14]

### Text dels agraïments [mida 11]

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquid ex ea commodi consequat. Quis aute iure reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint obcaecat cupiditat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquid ex ea commodi consequat. Quis aute iure reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint obcaecat cupiditat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquid ex ea commodi consequat. Quis aute iure reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint obcaecat cupiditat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquid ex ea commodi consequat. Quis aute iure reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint obcaecat cupiditat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

*[Pàgina en blanc]*

## Summary

As technology continues to advance, the use of Artificial Intelligence (AI) is growing steadily. However, the measures to prevent the potential risks associated with its use have not always kept pace.

The aim of this project is to examine how people perceive the implementation of AI systems in automated processes. The hypothesis proposed is that users tend to be more receptive to AI when it is applied to automatic or repetitive tasks, such as those involved in personnel selection processes.

To address this hypothesis, I conducted a review of the existing literature, in order to deepen the understanding of AI, its definition, evolution, applications in automated processes, and the biases that may arise. Additionally, I draw attention to the cognitive biases that affect human decision-making.

Based on this research, I designed a questionnaire from scratch with the aim of assessing users' knowledge of and attitudes toward AI.

**Resumen [en una 2a llengua. Ex. Resumen][mida 14]**

## Resum



*[Pàgina en blanc]*

## Índex

Introducción .....	1
1.1 Motivación .....	1
1.2 Objetivos .....	1
1.3 Logros.....	2
1.4 Organización de la memoria.....	2
Revisión Literaria.....	3
2.1 Definición.....	3
2.2 Evolución .....	4
2.3 Categorías éticas.....	7
2.3.1 Principales teorías éticas .....	7
2.3.2 Perfiles éticos dentro de la sociedad .....	8
2.4 Ética en la IA.....	9
2.4.1 Alineamiento de valores en la IA (IA <i>Value Alignment</i> ).....	10
2.5 Confianza en la IA?? .....	12
Metodología .....	13
3.1 Tecnologías utilizadas .....	13
3.1.1 Lenguaje: Python .....	13
3.1.2 Framework: Streamlit .....	13
3.1.3 Base de datos: Supabase.....	14
3.2 Estructura proyecto .....	15
Diseño Experimental .....	17
4.1 Objetivos e hipótesis .....	17
4.2 Participantes .....	17
4.3 Método de reclutamiento.....	17
4.4 Recolección de datos .....	18
Resultados y ____ .....	19
5.1 .....	19
Análisis e Interpretación de los resultados .....	20
5.1 .....	20
Conclusiones y trabajos futuros .....	22



## **Llista de figures [opcional] [mida 14]**

*[Fes la llista amb la mateixa estructura mostrada a l'índex, amb mida 11. Les figures i gràfics al cos del treball es titulen a la part d'abaix, amb mida 10]*

## **Llista de taules [opcional] [mida 14]**

*[Fes la llista amb la mateixa estructura mostrada a l'índex, amb mida 11. Les taules al cos del treball es titulen a la part d'adalt, amb mida 10]*

## Capítol 1

# Introducción

En las últimas décadas, los avances tecnológicos han impulsado de manera significativa el desarrollo de la inteligencia artificial (IA). Esta evolución ha permitido su integración en múltiples ámbitos de la sociedad, contribuyendo al progreso en sectores tan diversos como la medicina, la comunicación y la automatización de tareas. Sin embargo, esta evolución tecnológica también ha generado preocupaciones éticas. Las decisiones automatizadas pueden estar influenciadas por sesgos presentes en los datos o en el diseño de los algoritmos. Estos sesgos pueden reforzar desigualdades existentes y afectar negativamente a ciertos colectivos, especialmente a los más vulnerables [1].

Actualmente, se están desarrollando sistemas que asumen tareas que antes eran exclusivamente humanas, lo cual ha intensificado el debate sobre cuestiones como la transparencia algorítmica, la equidad en los resultados y la atribución de responsabilidades en la toma de decisiones automatizadas [2]. Ante este contexto, resulta relevante preguntarse hasta qué punto la ciudadanía es consciente de los sesgos que pueden estar presentes en la inteligencia artificial.

### 1.1 Motivación

Aunque la inteligencia artificial se ha integrado progresivamente en numerosos ámbitos de nuestra vida, su funcionamiento y consecuencias siguen siendo un terreno desconocido para gran parte de la ciudadanía. La automatización de decisiones, cada vez más común en contextos sociales, laborales o administrativos, se produce en muchos casos sin que las personas sean plenamente conscientes de ello, ni de los criterios que subyacen a esas decisiones. Esta falta de conocimiento o conciencia puede convertirse en un riesgo, especialmente si se considera que la IA ya está tomando decisiones que afectan directamente a las personas. Además, la ausencia de estrategias educativas o preventivas dificulta una adopción informada y crítica de estas tecnologías.

La motivación de este trabajo nace de esta preocupación: ¿es la ciudadanía realmente consciente del impacto de la IA y de los sesgos que puede contener? Y si no lo es, ¿deberíamos impulsar recursos, leyes o medidas educativas que fomenten la comprensión y el pensamiento crítico sobre estas tecnologías? Este trabajo busca contribuir a esta reflexión, con el objetivo de promover un uso más responsable, justo y transparente de la inteligencia artificial en la sociedad.

### 1.2 Objetivos

Los objetivos de este trabajo son los siguientes:

- Investigar y analizar la evolución de la inteligencia artificial y sus avances más relevantes, con el fin de obtener un conocimiento más profundo sobre su funcionamiento y aplicaciones.
- Definir los distintos perfiles éticos existentes en la sociedad, con el fin de establecer una posible relación con los principios que podrían regir el comportamiento de los sistemas de inteligencia artificial.
- Evaluar cómo la sociedad utiliza y percibe la inteligencia artificial mediante un cuestionario que permita conocer el nivel de información, confianza que genera y las posibles inquietudes que despierta.
- Analizar los resultados obtenidos en el cuestionario para extraer conclusiones sobre el grado de conciencia de la ciudadanía sobre los sesgos proporcionados por la IA y proponer posibles medidas de prevención, educación o regulación orientadas a un uso más ético y responsable de la IA.

### 1.3 Logros

En este trabajo se han cumplido los siguientes objetivos propuestos:

- “Logro 1”
- “Logro 2”

### 1.4 Organización de la memoria

El contenido de la memoria se divide en los siguientes capítulos:

- **Capítulo 2:** se introduce el concepto de la inteligencia artificial y se realiza una revisión literaria focalizada en la evolución de la IA, en la ética humana relacionada con la de la IA.
- **Capítulo 3:** se explica las metodologías utilizadas en el desarrollo del proyecto.
- **Capítulo 4:**
- **Capítulo 5:**

## Capítulo 2

# Revisión Literaria

Con el fin de dar respuesta a los objetivos de este trabajo, ya formulados en la introducción, se ha realizado una revisión literaria sobre la inteligencia artificial (IA), basada en su definición, evolución, legislación y sus principales ventajas y desventajas.

### 2.1 Definición

La inteligencia artificial (IA) es un concepto amplio y en constante evolución, cuya definición resulta compleja, ya que actualmente no existe una definición universalmente aceptada [g]. La primera definición reconocida fue propuesta en 1955 por uno de los pioneros de la IA, el profesor de Stanford John McCarthy, quien la definió como “la ciencia e ingeniería de fabricar máquinas inteligentes” [h]. Un año más tarde, en 1956 se llevó a cabo la primera investigación formal en el ámbito de la IA [i].

Desde entonces, diversas instituciones han intentado delimitar el concepto. La Real Academia Española (RAE), define la inteligencia artificial como la “disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico”. Por su parte, un artículo publicado por el Gobierno de España sugiere que la IA “es un campo de la informática centrado en crear sistemas capaces de realizar tareas que normalmente requieren inteligencia humana, como el aprendizaje, el razonamiento y la percepción” [g].

El Parlamento Europeo, en el Reglamento (UE) 2024/1689, ofrece una definición más técnica: un sistema de IA es “un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales.” (art. 3, def. 1).

Asimismo, la UNESCO <sup>1</sup> propone una definición desde un punto de vista ingenieril, definiendo la IA como “un sistema digital que procesa y analiza información en su entorno para actuar sobre él -con cierto grado de autonomía- con el fin de alcanzar objetivos específicos”. [j]

Estas distintas definiciones evidencian la complejidad del concepto y la multiplicidad de enfoques desde los que puede ser abordado. Entender esta variedad de enfoques es fundamental para analizar cómo influyen en la percepción pública y en las consideraciones éticas relacionadas con la IA.

---

<sup>1</sup> Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) es un organismo especializado de NNUU creado el 16 de noviembre de 1945 y cuya misión es “contribuir a la consolidación de la paz, la erradicación de la pobreza, el desarrollo sostenible y el diálogo intercultural mediante la educación, las ciencias, la cultura, la comunicación y la información”.



## 2.2 Evolución

La palabra *robot* fue utilizada por primera vez en el año 1921, en una obra de teatro de ciencia ficción, *Robots Universales Rossum*, escrita por el escritor checo Karel Čapek. Obra que trata sobre una empresa que fabrica humanos artificiales para ayudar con el trabajo a las personas [k].

Posteriormente, la década de los 50 fue muy importante para los avances de la inteligencia artificial (IA). Alan Turing fue el primer científico que, en el año 1950, publicó el artículo *Computing machinery and intelligence*, en el que cuestionó si las máquinas eran capaces de pensar por sí mismas [L]. Aunque no se acuñó al término que hoy en día conocemos como IA hasta el año 1956, cuando el profesor de Stanford John McCarthy organizó una conferencia en la Universidad de Dartmouth [k]. En el mismo año, Allen Newell y Herbert Simon fueron coautores de *Logic Theorist*, un programa de IA diseñado para imitar las habilidades de resolución de problemas de un ser humano [k].

En 1958, McCarthy creó el primer lenguaje de programación orientado a la investigación en IA, conocido como *List Processing* (LISP). Poco después, en 1959, el científico Arthur Samuel acuñó el término “machine learning” (aprendizaje automático) a partir de sus trabajos para entrenar un ordenador en el juego del ajedrez [L].

En 1961, se introdujo el robot “Unimate”, el primero en trabajar en una cadena de montaje en la empresa *General Motors*<sup>2</sup>. Más adelante, en el año 1966, “*ELIZA*”, el primer *chatbot*<sup>3</sup> con capacidad para conversar con una persona en inglés, fue desarrollada por el científico informático de MIT, Joseph Weizenbaum [k].

Estos avances dieron paso a que, en 1970, Marvin Minsky declarara en la revista *Life* que “en el plazo de tres a ocho años dispondremos de una máquina con la inteligencia general de un ser humano medio”. No obstante, ese optimismo no se materializó. Las investigaciones de la IA se enfrentaron múltiples obstáculos y decepciones, lo que llevó a una disminución en la financiación y el interés en el campo. No fue hasta 1986, cuando David Rumelhart, Geoffrey Hinton y Ronald Williams popularizaron el algoritmo de retropropagación para entrenar redes neuronales multicapa, que la IA volvió a despertar interés. Sin embargo, este impulso se desvaneció en 1987, iniciando otro período de desinterés [L][k].

Sin embargo, este “invierno de la IA” no impidió que se produjeran algunos avances significativos en el sector. Uno de los más destacados fue en 1997, cuando *Deep Blue*, un superordenador desarrollado por IBM, logró vencer al vigente campeón mundial de ajedrez, Gary Kasparov<sup>4</sup>, en una partida de ajedrez. Lo que supuso un gran avance en el desarrollo de la IA creada para la toma de decisiones [k].

Durante los primeros años del siglo XXI, la inteligencia artificial empezó a avanzar con más fuerza gracias a la mejora del procesamiento de datos y al aumento de la información disponible. Esto dio lugar al concepto de *big data*, que consiste en analizar grandes volúmenes de datos, tanto estructurados como no estructurados, para poder extraer información útil y tomar mejores decisiones con ayuda de algoritmos de aprendizaje automático [k].

Uno de los avances más importantes durante este periodo fue la aparición, en el año 2002, del primer robot de éxito comercial para el hogar, una aspiradora automática llamada “*Roomba*”. Más adelante, en 2009, Fei-Fei Li lanzó al mercado una base de datos gratuita de 14 millones de imágenes llamada *imageNet*, que los investigadores de la IA utilizaron para entrenar redes neuronales capaces de identificar objetos y clasificar imágenes [k].

<sup>2</sup> General Motors es una compañía líder a nivel mundial en la industria automotriz y de movilidad.

<sup>3</sup> Programa informático que simula la conversación humana con un usuario final.

<sup>4</sup> Candial, A. F. (2021, 10 febrero). Deep Blue-Kaspárov: cuando la máquina venció al hombre. *La Vanguardia*. <https://www.lavanguardia.com/vida/junior-report/20210210/6234712/kasparov-deep-blue-maquina-vencio-hombre.html>

En paralelo, los asistentes virtuales comenzaron a incorporarse a la vida cotidiana. En 2011 *Apple* presentó *Siri*, su primer asistente por voz y poco después, en 2014, *Amazon* lanzó al mercado *Alexa*, un asistente virtual diseñado para interactuar mediante comandos de voz [k][m].

Durante estos últimos años, el desarrollo del aprendizaje profundo<sup>5</sup>, de sistemas de reconocimiento y generación de imágenes y lenguaje ha evolucionado de manera significativa. Uno de los avances más significativos en este ámbito se produjo en el año 2012, cuando surgió *AlexNet*, una red neuronal capaz de reconocer imágenes de la base de datos de *imageNet* con un rendimiento sobrehumano por primera vez. Este avance marcó un hito en la historia de la inteligencia artificial y aceleró el desarrollo del aprendizaje profundo [m].

Grandes compañías tecnológicas impulsaron este progreso desde entonces. Por ejemplo, Facebook creó en 2013 su laboratorio *FAIR* (*Facebook Artificial Intelligence Research*), impulsando el desarrollo de modelos cada vez más complejos en visión, lenguaje y aprendizaje autónomo [p][r]. En este contexto de avances continuos, *DeepMind* sorprendió al mundo en 2016 con *AlphaGo*, un sistema de inteligencia artificial que derrotó al campeón mundial del juego *Go*, un juego conocido por su complejidad estratégica. Pero no se quedó ahí, ya que al año siguiente la versión mejorada *AlphaGo Master* derrotó al jugador número uno del mundo, Ke Jie, consolidando el dominio de la IA en esta área [k][m].

Estos progresos también generaron preocupaciones sobre los riesgos de la IA. En 2017, el físico Stephen Hawking advirtió que “a menos que aprendamos a prepararnos para los riesgos potenciales y a evitarlos, la inteligencia artificial puede ser el peor acontecimiento de la historia de nuestra civilización” [k].

Aun así, el ámbito de la IA siguió avanzando. En 2018, *OpenAI* presentó *GPT* (*Generative Pre-trained Transformer*), un modelo que marcó un antes y un después en el aprendizaje no supervisado del lenguaje. Este modelo, con más de 117 millones de parámetros, demostró la capacidad de una IA para predecir y generar texto de forma coherente a partir de grandes volúmenes de datos. Este desarrollo asentó las bases para modelos posteriores más avanzados, como GPT-2 y GPT-3, que ampliaron aún más las capacidades de la inteligencia artificial en el ámbito del lenguaje [q].

En noviembre de 2022, *OpenAI* lanzó *ChatGPT*, un chatbot basado en inteligencia artificial que, en poco tiempo, se convirtió en un fenómeno global gracias a su capacidad para responder de forma coherente a preguntas y solicitudes de los usuarios. Este avance avivó el interés internacional por las aplicaciones de la inteligencia artificial generativa [r].

Apenas unos meses después, en marzo de 2023, *OpenAI* lanzó GPT-4, un modelo aún más avanzado y con capacidades multimodales, es decir, capaz de interpretar tanto texto como imágenes. Este modelo mostró una mejora significativa en tareas de comprensión lectora, razonamiento complejo, resolución de problemas matemáticos, programación y generación de contenido especializado [t]. Modelo que integró *Microsoft* en su asistente, *Copilot* [s]. Por otro lado, *Google DeepMind* presentó **Gemini**, una nueva familia de modelos de IA diseñados para trabajar con múltiples tipos de entrada (texto, imagen, código), y optimizados para un razonamiento más alineado con valores humanos [u].

Estos avances han posicionado a la inteligencia artificial como un elemento central en la innovación tecnológica actual, impulsando nuevas aplicaciones y modelos de negocio en múltiples sectores.

---

<sup>5</sup> El aprendizaje profundo (*deep learning*) se basa en capas de las redes neuronales, que son algoritmos vagamente modelados de la forma en que funciona el cerebro humano [o].

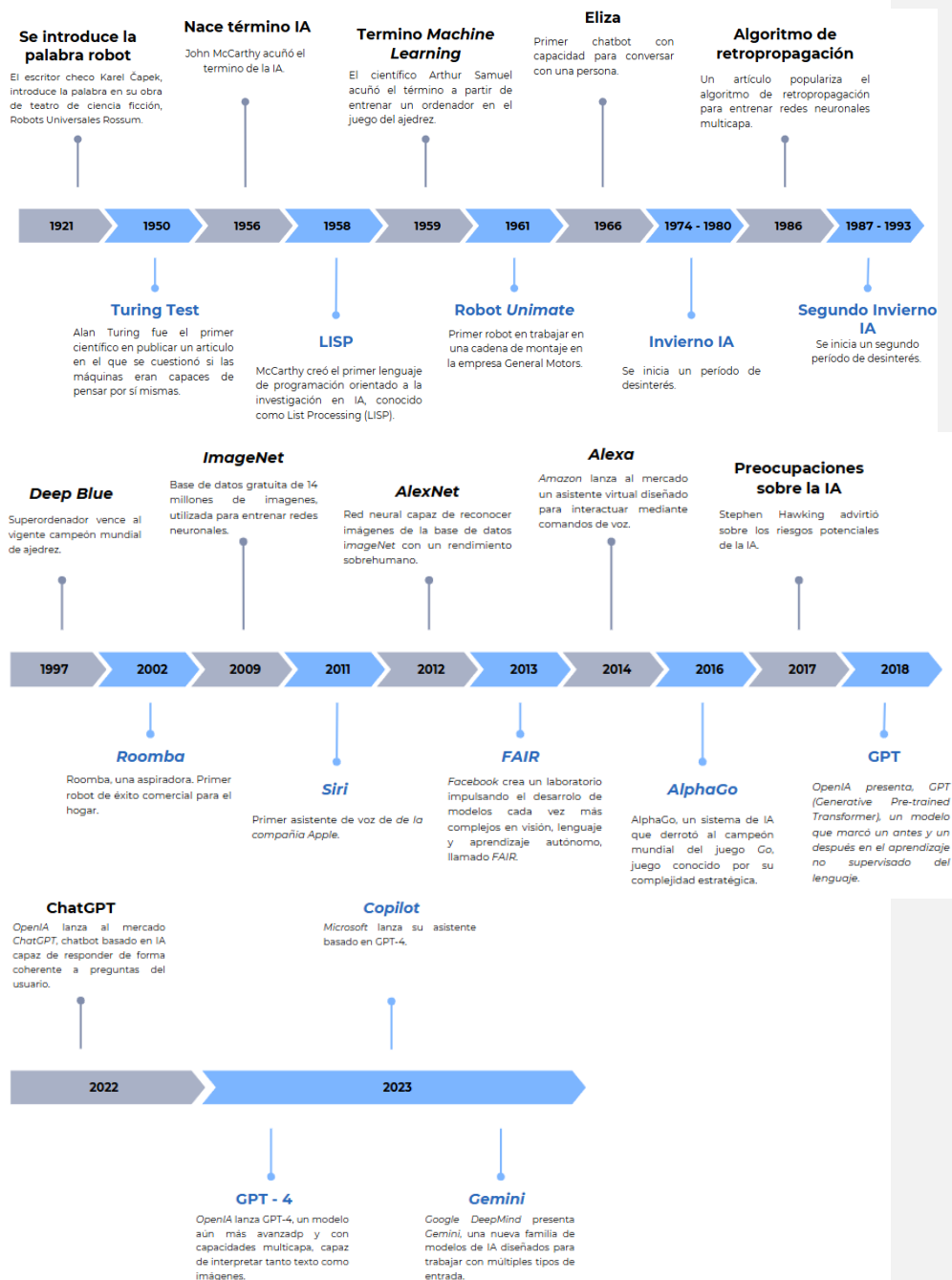


Figura 1.1: Línea del tiempo de la evolución de la IA (elaboración propia).

## 2.3 Categorías éticas

Según Sócrates la moralidad es “cómo deberíamos vivir”. Sin embargo, no hemos logrado avanzar más allá de esta formulación, ya que cualquier intento de definirla tiende a entrar en conflicto con al menos una de las muchas teorías éticas existentes [v].

Aun así, si se ha establecido una concepción mínima de la moralidad, que la define como el esfuerzo por guiar nuestras acciones usando la razón. Es decir, actuar conforme a las mejores razones disponibles, mientras se intenta dar la misma importancia a los intereses de todas las personas afectadas por nuestras decisiones. Ahora bien, no todas las teorías éticas aceptan esta definición mínima [v].

### 2.3.1 Principales teorías éticas

Las principales teorías éticas se reflejan ampliamente en la vida social y pueden aplicarse a diversas acciones o dimensiones de la sociedad actual [x]. A continuación, se presentan algunas de las más reconocidas:

#### Utilitarismo

El principio fundamental del utilitarismo se basa en “el credo que acepta como fundamento de la moral la *utilidad*, o el *principio de la máxima felicidad*, sostiene que las acciones son buenas en cuanto tienden a promover la felicidad, malas en cuanto tienden a producir lo opuesto a la felicidad. Por *felicidad* se entiende placer y ausencia de dolor; por *infelicidad*, dolor y privación de placer” [y]. Mill quiso demostrar que el utilitarismo no es una filosofía del egoísmo, ya que la felicidad, en el contexto moral, “no es la máxima felicidad particular del agente, sino la medida mayor de felicidad en conjunto” [y]. Cómo podemos observar, el criterio que utiliza Mill para juzgar si una acción es correcta se basa en sus consecuencias. En otras palabras, lo que le importa es el impacto que tiene sobre el bienestar general [a1].

En resumen, el utilitarismo juzga la moralidad buscando maximizar la felicidad y minimizar el sufrimiento para el mayor número posible de personas [a1].

#### Deontología Kantiana

Immanuel Kant es considerado el referente de la teoría ética deontológica. Decía que la vida moral del hombre no consiste en buscar la felicidad, sino en cómo hacernos dignos de la felicidad [z]:

“No es propiamente la moral la doctrina de cómo nos *hacemos* felices, sino de cómo debemos llegar a ser *dignos* de la felicidad” (Kant, 1975: 182).

Para Kant, la felicidad consiste en “la satisfacción de todas nuestras inclinaciones”; sin embargo, actuar moralmente no implica guiarse por la búsqueda de esa satisfacción. Distingue entre dos tipos de leyes prácticas: una basada en la felicidad, que denomina *pragmática*, y otra basada únicamente en la dignidad de ser feliz, a la que llama *ley moral* o *ley ética*:

“La felicidad es la satisfacción de todas nuestras inclinaciones. La ley práctica derivada del motivo de la felicidad la llamó pragmática. En cambio, la ley, si es que existe, que no posee otro motivo que la dignidad de ser feliz la llamó ley moral (ley ética)” (Kant, CRP, A806/B834).

Desde esta perspectiva, una persona que actúa en función de su felicidad está sujeta a factores externos y contingentes y, por lo tanto, su acción no es plenamente libre ni autónoma. En cambio, para Kant, la verdadera libertad moral se logra cuando se actúa por deber, es decir, por respeto a la ley moral universal, que se reconoce racionalmente como válida para todos [z].

Así, la deontología kantiana sostiene que la moralidad reside en actuar por deber, más allá de buscar la propia felicidad o satisfacción [z].

### Ética de la virtud: Eudemonismo

Hasta ahora hemos abordado dos teorías éticas que se centran en los actos, es decir, en proporcionar criterios objetivos para juzgar si una acción es correcta o incorrecta. Sin embargo, ahora nos enfocaremos en un enfoque diferente: las éticas del agente. En este contexto, Aristóteles propone la ética de la virtud, también conocida como eudemonismo, que parte de la idea de que los seres humanos estamos naturalmente orientados hacia la felicidad (eudaimonía). No obstante, para alcanzarla no basta con desearla, sino que es necesario cultivar y practicar las virtudes [a2].

Para Aristóteles, la virtud tiene que ver con nuestra forma de ser. Parte de la pregunta: “¿Cómo debo ser?”, en lugar de “¿Qué debo hacer?”. El bien y el mal se entienden a partir del desarrollo de las virtudes, que nos convierten en buenas personas. Vivir de forma virtuosa es, para él, la forma más plena y feliz de vivir [a2].

Además, Aristóteles distingue entre virtudes morales e intelectuales. Las virtudes morales, como la valentía y la templanza, se adquieren mediante la práctica y el hábito, y consisten en encontrar un equilibrio entre extremos, lo que él llama el “justo medio”. Por ejemplo, entre la cobardía y la temeridad, se encuentra la virtud del valor. Las virtudes intelectuales, como la sabiduría, se desarrollan a través de la enseñanza y son esenciales para guiar las virtudes morales [a3].

En este sentido, la eudaimonía, o felicidad, se alcanza cuando el alma racional actúa de acuerdo con la virtud, realizando su función propia de manera excelente [a3].

### Ética del cuidado

Por último, es también relevante hablar sobre la ética del cuidado. El cuidado de manera ética involucra “la interacción y el contacto moral entre dos personas” [a4]. A diferencia de las teorías anteriores, esta no se centra en “¿qué es justo?” sino en “¿cómo responder a las necesidades de esta persona en esta situación?”. Por lo que tiene como objetivo cuidar desde una perspectiva más individual y no colectiva [a5].

La ética del cuidado entiende el mundo como una red de relaciones en la que todos estamos implicados, y considera que el compromiso moral surge del reconocimiento de las necesidades del otro. Cuando alguien necesita ayuda, sentimos que tenemos la responsabilidad de actuar. Esta respuesta no se basa en reglas universales, sino en la situación concreta que vivimos con esa persona [a4].

Por eso, la ética del cuidado valora especialmente experiencias reales como acompañar, proteger o evitar el daño. Esta teoría sostiene que la moralidad no solo implica autonomía y justicia, sino también la capacidad de hacernos cargo del bienestar de otros [a4].

### 2.3.2 Perfiles éticos dentro de la sociedad

Según la Real Academia Española (RAE), un perfil es un “conjunto de rasgos que caracterizan a una persona o cosa”. Así pues, un perfil ético tiende a referirse a las directrices de nuestra propia conducta [w].

Identificar perfiles éticos dentro de la sociedad es un aspecto clave en este trabajo, ya que permite entender cómo distintas personas pueden interpretar y valorar la inteligencia artificial y sus implicaciones desde marcos morales distintos.

A partir de las teorías éticas revisadas anteriormente, podemos distinguir algunos perfiles éticos:

- **Perfil utilitarista:** Una persona que sigue un perfil utilitarista tiende a valorar las decisiones en función de sus consecuencias y del impacto que estas tienen en el bienestar general de la

sociedad. Cuando se le presenta la oportunidad de mejorar el bienestar común, actúa en consecuencia, incluso si ello implica ciertos sacrificios individuales [a6].

- **Perfil deontológico:** Este perfil lo adoptan aquellas personas que se guían por el deber en la toma de decisiones, de manera que no se centran sólo en respetar sus propios intereses, sino en los de los demás. Además, suelen actuar pensando en el bien de la sociedad y de las generaciones futuras [a7].
- **Perfil virtuoso:** Las personas que siguen este perfil consideran que viven actuando de acuerdo con la virtud, ya que creen que ese es el camino para alcanzar la verdadera felicidad. Cada decisión que toman, por más agradable o difícil que sea, la realizan de forma voluntaria, no por obligación. Son conscientes de que tienen libertad y voluntad sobre sus acciones, y que desarrollar buenos hábitos es clave para ser mejores personas. Además, creen que la virtud no es algo innato, sino que se cultiva a través de la educación, la experiencia y el esfuerzo constante [a8].
- **Perfil del cuidado:** Quién sigue este perfil toma decisiones basadas en la empatía, la responsabilidad y la atención a las relaciones interpersonales [a9]. Entienden el mundo como una red de vínculos en la que surge un compromiso hacia los demás y su bienestar. Son personas que sienten la responsabilidad de evitar el daño y proteger al otro [a4].

Cabe destacar que una misma persona puede identificarse con más de uno de estos perfiles, dependiendo del contexto o del tipo de problema al que se enfrente. Sin embargo, reconocer la existencia de los diversos perfiles permite entender mejor la diversidad de posturas ante los desafíos éticos que plantea la IA.

## 2.4 Ética en la IA

Kant afirmó “que la invención del puñal precedió a la conciencia del imperativo categórico (“no matarás”), lo que daba a entender que los avances tecnológicos se anticiparon a las orientaciones morales de como hacer uso de ellas [a9]. Esta reflexión continúa siendo pertinente en la actualidad, ante el ritmo vertiginoso de los avances tecnológicos relacionados con la inteligencia artificial, en el año 2018, en el contexto de la Unión Europea, empezaron a sugerir propuestas de marcos éticos.

En el contexto de la Unión Europea, en diciembre de 2018 el grupo *AI4People* propuso el marco ético “*Ethical Framework for a Good AI Society: Opportunities, Risks, Principles and Recommendations*”. En abril de 2019, la Comisión Europea presentó las “*Ethics Guidelines for Trustworthy AI*”, elaboradas por el *High-Level Expert Group on Artificial Intelligence*. Asimismo, en noviembre de 2018, la Universidad de Deusto presentó la “*Declaración de Derechos Humanos para un Entorno Digital*” [a9].

En esta misma línea de preocupación, la UNESCO expresó su preocupación por los riesgos éticos emergentes durante la conferencia General de la UNESCO celebrada en noviembre del año 2021 [1].

Como respuesta, la organización propuso la primera norma de alcance mundial en materia de la ética de la IA, conocida como “*The Recommendation on the Ethics of Artificial Intelligence*”. Esta fue aprobada por unanimidad por los 193 Estados miembros en dicha conferencia [1]. La Recomendación se centra en la protección de los derechos humanos, centrándose en cuatro valores fundamentales orientados a maximizar el bien para la humanidad, los individuos, la sociedad y el medioambiente [1]. Estos cuatro valores son los siguientes:



Figura X: Valores reconocidos por la UNESCO

Además de estas iniciativas institucionales, en los últimos años ha cobrado protagonismo un concepto clave en el ámbito de la inteligencia artificial ética: el alineamiento de valores (AI Value Alignment) [a12].

#### 2.4.1 Alineamiento de valores en la IA (IA Value Alignment)

Este enfoque se centra en la responsabilidad humana, destacando que los seres humanos son responsables de los impactos éticos y sociales de los sistemas de IA, incluso cuando estos sistemas superan la inteligencia humana y su control se vuelve un desafío. Este concepto es fundamental para poder asegurar que el comportamiento de la IA sigue y respeta los valores humanos [a12].

Ahora bien, ¿cuáles son esos valores humanos que deben guiar el alineamiento de la IA? El principal problema es que los valores humanos suelen ser bastante abstractos, y varían significativamente según el contexto cultural, el país o la situación concreta [a12].

Por ejemplo, en la dimensión cultural, observamos que distintas culturas pueden tener interpretaciones muy diferentes de ciertos valores, como la privacidad. A ello se suma que, incluso dentro de una misma cultura, cada persona puede tener sus propios valores y preferencias. Por lo tanto, es importante reconocer que los valores humanos pueden identificarse y priorizarse en distintos niveles: individual, organizacional, nacional o universal [a12].

El alineamiento de valores en sistemas de IA no puede limitarse a una simple incorporación técnica, sino que requiere una combinación de mecanismos técnicos y procesos organizativos (Figura X). Este proceso suele comenzar con el análisis de los escenarios de despliegue, la consulta con múltiples partes interesadas y la formación de desarrolladores y usuarios [a12].

Hay muchos valores para tener en cuenta, cada uno de ellos con un gran estudio y discusiones a nivel social y técnico. Uno de ellos especialmente relevante en este contexto es la justicia, que ha sido objeto de estudio incluso antes del inicio de la IA generativa. Este valor se ha aplicado en sistemas clasificadores y predictores para identificar y mitigar sesgos [a12].

Este tema plantea debates éticos relevantes, ya que validar el alineamiento de valores en el ámbito de la IA requiere la validación social. Surge entonces una pregunta clave: ¿cuál es el umbral de sesgo aceptable para considerar que un sistema es justo?

Otro valor fundamental es la transparencia. Para construir una base sólida de confianza, los usuarios deben conocer qué recursos, pruebas y criterios han utilizado los desarrolladores para verificar y validar dichos valores [a12].

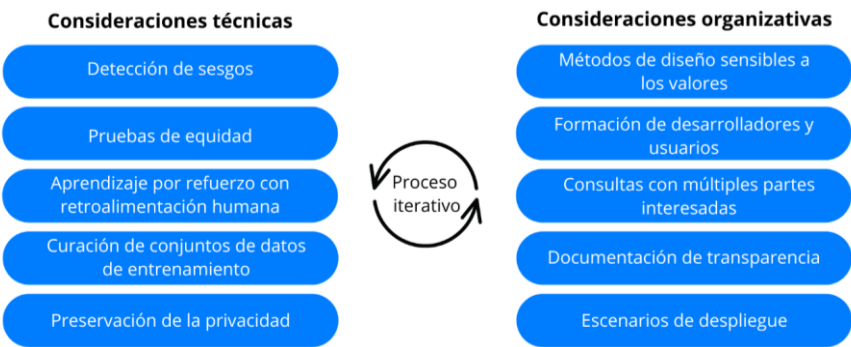


Figura X: Consideraciones técnicas y organizativas en el contexto de la alineación de valores en la inteligencia artificial. (Adaptación de la Figura 3 del informe “AI Value Alignment” del World Economic Forum, 2024).

Integrar los valores humanos en los procesos de la IA de manera continua requiere un enfoque ético y multidisciplinar, donde la filosofía moral desempeñe un papel crucial, ya que no solo fundamentan gran parte de los valores humanos y normas sociales aceptadas, sino que también inspiran muchos de los principios y guías que orientan la ética en inteligencia artificial [a12].

Entre los enfoques éticos más influyentes se encuentran el utilitarismo, que busca maximizar el bienestar colectivo; la deontología, que se centra en los deberes y derechos individuales; y la ética de la virtud, que promueve la responsabilidad, la prudencia y la empatía en quienes desarrollan estos sistemas. Estos marcos filosóficos no solo han orientado reflexiones académicas, sino que también han influido en la elaboración de marcos regulatorios concretos [a12 o el que pone en el pdf].

En el contexto occidental, el cristianismo, el islam y el judaísmo han contribuido a la configuración de principios éticos. Además, también influyen aspectos religiosos, como el cristianismo, el islam y el judaísmo. En cambio, en países orientales, se ve influenciado por el confucianismo, el budismo y el hinduismo. Las tradiciones africanas, se basan también en otras prioridades éticas [a12].

Todo ello resalta la importancia de comprender y tener en cuenta las perspectivas de las distintas culturas en el momento de establecer bases universales en este ámbito. En el ámbito legislativo, las leyes relacionadas con la IA debes de ser diseñadas en función de la región en la cual opere esa tecnología.

En definitiva, solo mediante la colaboración entre ética y legislación, informada por una comprensión plural de los valores humanos, podrá desarrollarse una inteligencia artificial que no solo sea eficaz y funcional, sino también justa, inclusiva y alineada con los derechos fundamentales de todas las personas.

**Comentado [IS1]:** Revisar la estructura y redefinir algunas partes



- **[1] Doc unesco:** Ten core principles lay out a human-rights centred approach to the Ethics of AI.
- **[a9] Abrir\_pdf:** En sus trabajos de filosofía de la historia decía Kant que la invención del puñal precedió a la conciencia del imperativo categórico (“no matarás”), es decir, que los avances técnicos se anticiparon a las orientaciones morales sobre cómo hacer uso de ellos. Y sigue siendo cierto que el exponencial progreso de lo que hoy llamamos ya “tecnociencias” plantea una gran cantidad de preguntas éticas para las que es necesario ir encontrando respuestas. Precisamente porque lo moral no consiste en mapas de carreteras, ya cerrados, sino en una brújula que señala el norte, es posible y necesario encontrar mejores caminos ante los nuevos descubrimientos tecnocientíficos. + habla de que filosofía late en la base de la IA (en teoría, acabar de leer para confirmar) + **propuestas de marcos éticos (UE)**
- **[a10] The\_Ethics\_of\_AI (IMPORTANTE):** algunos riesgos éticos de la IA + primer ethical code of IA (evolución de ello, como paso de 3 leyes a 4). + habla de cómo debería ser la IA.
- **[a11] Inicio IA y ética en medicina:** Preocupaciones éticas por parte de la UNESCO (2023), por lo que propone la primera norma en materia de la ética de la IA aprobada noviembre de 2021.
- **[a12] WEF\_AI\_Value\_Alignment\_2024:** Habla del value alignment // Definición IA value alignment // Cada persona tiene sus propios valores y preferencias, al igual que diferentes culturas hacen que las personas piensen o actúen diferente, por lo que los valores humanos pueden identificarse al nivel de cada individuo, organización, nación o de manera global, por lo que es importante identificar y priorizar los buenos valores dependiendo de cada escenario. // Societal agreement is necessary for validating AI value alignment. // Threshold of bias is acceptable for certifying fairness alignment? // (Human values such as justice, privacy and agency are fundamental principles) ¡¡Habla de como aplicar el value alignment en la vida real!! **Habla de las éticas que sigue la IA.**
- **The ethics of algorithms- key problems and solutions:** Algorithms are not ethically neutral.

**Comentado [IS2]:** interesante

## 2.5 Confianza en la IA??

## Capítulo 3

# Metodología

### 3.1 Tecnologías utilizadas

#### 3.1.1 Lenguaje: Python

Para la implementación de la aplicación, se ha utilizado Python 3.13.1 [a].

Python es un lenguaje de programación ampliamente utilizado en el desarrollo de software para aplicaciones web, ciencia de datos y aprendizaje automático (ML). Según el índice *TIOBE* Programming Community, Python es considerado el lenguaje de programación número uno en marzo de 2025 [b].

Python es conocido por su eficiencia y facilidad de aprendizaje, especialmente en comparación con otros lenguajes como *Java* o *C++*, gracias a su sintaxis sencilla. Es un lenguaje multiparadigma<sup>6</sup>, lo que lo hace más flexible y compatible que otros lenguajes. Además, Python es capaz de ejecutarse en múltiples plataformas.

Una de las mayores ventajas de Python es su amplio ecosistema de bibliotecas y frameworks, como *NumPy*, *Pandas* y *Django*, que potencian en gran medida sus capacidades en el procesamiento de datos y el desarrollo web.

Python es un lenguaje interpretado, lo que significa que ejecuta el código línea por línea. Si se detecta algún error durante la ejecución, esta se detiene de inmediato, lo cual facilita la depuración.



Imagen 1: Logo Python

#### 3.1.2 Framework: Streamlit

Tras considerar varios frameworks disponibles en Python, se optó por utilizar Streamlit para facilitar y simplificar el proceso de desarrollo de la aplicación web [c]. Esta decisión también estuvo influenciada por la experiencia previa con Streamlit durante el grado.

Streamlit es un framework gratuito y de código abierto para crear y compartir aplicaciones web, principalmente orientadas al aprendizaje automático y a la ciencia de datos. Al estar basado en Python, permite integrar bibliotecas populares como *Pandas*, *Matplotlib* y *Plotly*, lo que facilita el análisis y la visualización de datos de forma eficiente.

Además, Streamlit elimina la necesidad de conocimientos en tecnologías de desarrollo front-end como *HTML*, *CSS* o *JavaScript*. Esto permite crear aplicaciones interactivas y totalmente funcionales

---

<sup>6</sup> Capacidad para utilizar diferentes estilos de programación, como la programación estructurada (procedimental) o la programación orientada a objetos.

utilizando únicamente Python, lo cual resulta especialmente ventajoso en este caso, ya que el objetivo es centrarse en la lógica de la aplicación más que en su desarrollo visual, especialmente considerando que durante el grado no hemos trabajado con dichas tecnologías front-end.

Asimismo, Streamlit aprovecha el hecho de que Python es un lenguaje interpretado, permitiendo realizar modificaciones en el código en tiempo real, sin necesidad de recompilar ni reiniciar el servidor local. Esta característica acelera considerablemente el proceso de desarrollo, ya que los cambios se reflejan de inmediato en la aplicación web, lo que permite iterar rápidamente en el diseño y la funcionalidad, además de facilitar la depuración de errores.



*Imagen 2: Logo Streamlit*

### 3.1.3 Base de datos: Supabase

Para el almacenamiento de los datos del proyecto se ha utilizado *supabase*, una alternativa de *Firebase* de código abierto el cual usa *PostgreSQL* como base de datos principal [d].

La conexión entre la base de datos y el script de la aplicación se ha realizado mediante la URL de la base de datos y una clave API privada, la cual actúa como un mecanismo de autenticación que permite controlar qué usuarios o aplicaciones tienen acceso a los datos. Esta clave, similar a una contraseña, esta guardada fuera del alcance de cualquier persona no autorizada, ya que debe mantenerse protegida para evitar poner en riesgo la confidencialidad de los datos de los usuarios.

La estructura de la base de datos ha sido definida manualmente usando SQL desde la propia interfaz de supabase.



*Imagen 3: Logo supabase*

```
create table if not exists respuestas (
  id serial primary key,
  timestamp timestamptz default now(),
  genero text,
  edad text,
  nivel_estudios text,
  nivel_estudios_otro text,
  rama_estudios text,
  rama_estudios_otro text,
  años_experiencia text,
  pais_residencia text,
  pregunta_1 text,
  pregunta_2 text,
  pregunta_3 text,
  ...
  pregunta_32 text
);
```

```
create table if not exists personal_info (
  id serial primary key,
  timestamp timestamptz default now(),
  correo_electronico text,
  nombre_apellido text
);
```

Figura X: Código de las diferentes bases de datos (Izquierda: Base de datos principal, Derecha: base de datos para almacenar los datos del usuario)

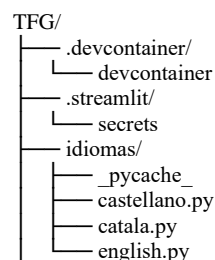
Comentado [IS3]: Ponerlo bien

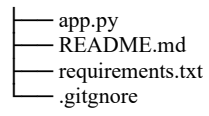
### 3.2 Estructura proyecto

El proyecto se ha estructurado de una manera sencilla y precisa, para así facilitar la navegación de los archivos (**Figura X**). Con el objetivo de evitar la pérdida de información, todos los documentos, excepto el que contiene las credenciales necesarias para acceder a la base de datos, se han subido a un repositorio de *GitHub*.

A continuación, se resume el contenido principal del proyecto:

- **.devcontainer/**: carpeta utilizada para la configuración del entorno de desarrollo en contenedores, lo cual facilita replicar el entorno en diferentes equipos.
- **.streamlit/**: contiene el documento *secrets*, donde se encuentran las credenciales necesarias para acceder a la base de datos.
- **idiomas/**: carpeta con las traducciones de la aplicación (*castellano.py*, *catala.py* y *english.py*).
- **app.py**: archivo que contiene la aplicación del cuestionario.
- **README.md**: documento que contiene una descripción general del proyecto. Incluye la explicación del propósito y otros detalles relevantes. Está pensado para usuarios que quieran entender el proyecto.
- **requirements.txt**: contiene la lista de dependencias necesarias para que la aplicación funcione correctamente.





```
graph LR; Root[ ] --- app.py; Root --- README.md; Root --- requirements.txt; Root --- .gitignore;
```

app.py  
README.md  
requirements.txt  
.gitignore

*Figura X: Estructura proyecto*

## Capítulo 4

# Diseño Experimental

### 4.1 Objetivos e hipótesis

Este trabajo propone analizar las opiniones de las personas sobre la utilidad de la inteligencia artificial (IA), haciendo hincapié en su conocimiento y percepción ética. Se pretende identificar en qué aspectos hay una mayor aceptación de la IA cuando se aplica a tareas automáticas, en función del nivel de conocimiento del usuario sobre su funcionamiento, así como de su consciencia sobre los beneficios y sesgos asociados.

Para poder llevar a cabo los objetivos, un cuestionario ha sido diseñado para obtener información sobre el conocimiento que tienen los usuarios acerca de la IA, incluyendo su comprensión de los sesgos y de sus aplicaciones, el uso que hacen de esta tecnología en distintos ámbitos, y sus actitudes respecto a diversos dilemas éticos.

A partir del análisis de la información recogida, se espera, extraer conclusiones sobre el grado de consciencia de los usuarios en estos ámbitos y al mismo tiempo, proponer posibles medidas de divulgación o regulación de la IA hacia un ámbito más ético y responsable.

### 4.2 Participantes

Los participantes del cuestionario deben tener 18 años o más, es decir, haber alcanzado la mayoría de edad en España. No se requiere un conocimiento técnico avanzado sobre inteligencia artificial, aunque sí se espera una familiaridad básica con su presencia en aplicaciones cotidianas.

Se prevé reclutar cómo mínimo 50 participantes principalmente entre estudiantes o antiguos alumnos de la Universidad Pompeu Fabra, ya que es más probable que reciban la información sobre el estudio a través de los canales de difusión recomendados.

La participación es totalmente voluntaria y los participantes pueden interrumpir o abandonar el cuestionario en cualquier momento, sin obligación alguna.

### 4.3 Método de reclutamiento

Para comprobar la hipótesis descrita anteriormente, se ha diseñado un cuestionario completamente voluntario con el objetivo de recoger información sobre el conocimiento y percepción de los participantes en relación con la IA.

El estudio se organiza en los siguientes pasos:

#### Paso 1: Reclutamiento de los participantes

El reclutamiento de los participantes se llevado a cabo a través de los canales de difusión de la Universidad Pompeu Fabra (UPF). Además, se ha diseñado un cartel informativo con el propósito de promover la participación en el cuestionario (*Anexo 2*).

#### Paso 2: Información a los participantes y consentimiento

Antes de acceder al cuestionario, los participantes deben leer la hoja informativa del estudio (*Anexo 3*). Una vez leída, deben aceptar las condiciones y confirmar que han comprendido su contenido.

### **Paso 3: Acceso al cuestionario**

Los participantes acceden al cuestionario a través de un enlace directo, que abre inmediatamente el formulario en su navegador. No se requiere registro ni inicio de sesión. El cuestionario combina preguntas cerradas de opción múltiple y tipo Likert de 5 puntos, con preguntas abiertas. Está estructurado en tres secciones: información personal, conocimiento general de la IA y la percepción sobre la IA en casos concretos, basado en el enfoque ético.

## **4.4 Recolección de datos**

Las respuestas se almacenan en una base de datos en *Supabase*, mencionado en el capítulo anterior. En esta base de datos se almacenan las respuestas de los usuarios de manera anónima. Cada fila de la tabla representa las respuestas de un usuario. Por otro lado, tenemos una segunda base de datos en la que se guardan los datos personales (nombre completo y correo electrónico) de todas aquellas personas que estén interesadas en formar parte del concurso de 150€ por participar en el cuestionario, esta parte es voluntaria y esta base de datos se guarda separada, solo las personas autorizadas pueden acceder a aquella información. Esta segunda base de datos será borrada una vez realizado el concurso.

**Capítulo 5**

**Resultados y \_\_\_\_**

**5.1**



## **Capítulo 6**

### **Análisis e Interpretación de los resultados**

#### **5.1**



## Conclusiones y trabajos futuros

[Es pot separar conclusions i treball futur en capítols diferents, sense sobrepassar les 30 pàgines del cos de treball.]

[ Text conclusió i treball futur. El text és pot separar en apartats.][mida 11]  
Lorem

## Bibliografía

[1] *UNESCO's Recommendation on the Ethics of Artificial Intelligence: key facts - UNESCO Biblioteca Digital*, s. f.) <https://unesdoc.unesco.org/ark:/48223/pf0000385082>

[2] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016b). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2). <https://doi.org/10.1177/2053951716679679>

[a] *Streamlit* • *Streamlit*. (n.d.). <https://streamlit.io/>

[b] Jansen, P. (2022, 3 junio). *TIOBE Index - TIOBE*. TIOBE. <https://www.tiobe.com/tiobe-index/>

[c] *Workspace*, G. (n.d.). *Google Sheets: Online Spreadsheets & Templates | Google Workspace*. Google Workspace. <https://workspace.google.com/products/sheets/>

[d] *Supabase*. (s. f.). *GitHub - supabase/supabase: The open source Firebase alternative. Supabase gives you a dedicated Postgres database to build your web, mobile, and AI applications.* GitHub. <https://github.com/supabase/supabase>

[e] *About RESTful APIs*. (n.d.-b). Google for Developers. <https://developers.google.com/photos/overview/w/about-restful-apis>

[f] *Service accounts overview*. (s. f.). Google Cloud. <https://cloud.google.com/iam/docs/service-account-overview>

[g] *Qué es la Inteligencia Artificial | Plan de Recuperación, Transformación y Resiliencia Gobierno de España*. (s. f.). <https://planderecuperacion.gob.es/noticias/ques-inteligencia-artificial-ia-prtr>

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence. *Official Journal of the European Union*, L 1689. <https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81079>

[h] Manning, C. (2020). *Artificial intelligence definitions*. Stanford Institute for Human-Centered Artificial Intelligence (HAI). <https://hai-production.s3.amazonaws.com/files/2020-09/AI-Definitions-HAI.pdf>

[i] GENERAL \_ ARTICLE JohnMcCarthy – Father of Artificial Intelligence

[j] Martínez, M. V., UNESCO Office Montevideo and Regional Bureau for Science in Latin America and the Caribbean, & Consejo Latinoamericano de Ciencias Sociales. (2024). *De qué hablamos cuando hablamos de inteligencia artificial* (MTD/SC/2024/PI/06; 33 págs.). <https://creativecommons.org/licenses/by-sa/3.0/igo/>

[k] *Vista de UNA DOBLE HISTORIA DE LA INTELIGENCIA ARTIFICIAL: AVANCE TECNOLÓGICO y PROCESO DE REGULACIÓN EN EUROPA*. (s. f.). <https://revista.proeditio.com/rpdd/article/view/7857/8369>

[L] Abeliuk, A., & Gutiérrez, C. (2021). *Historia y evolución de la inteligencia artificial*. Revista Bits, Universidad de Chile. <https://revistasdex.uchile.cl/index.php/bits/article/download/2767/2700>

[n] Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>

[m] *Vista de Evolución de la Inteligencia Artificial y su impacto en la Educación: Revisión de la literatura*. (s. f.). <https://tsachila.edu.ec/ojs/index.php/TSEDE/article/view/195/129>

[o] Oracle. (2022, 20 agosto). *What is Deep Learning?* <https://www.oracle.com/es/artificial-intelligence/machine-learning/what-is-deep-learning/>

[p] LeCun, Y., Pesenti, J., & Schroepfer, M. (2020, 24 marzo). *FAIR turns five: What we've accomplished and where we're headed*.

Engineering At Meta.  
<https://engineering.fb.com/2018/12/05/ai-research/fair-fifth-anniversary/>

[r] Martínez-Ortega, A. G., & Medina-Chicaiza, R. P. (2020). Tecnologías en la inteligencia artificial para el Marketing: una revisión de la literatura. *Pro Sciences*, 4(30), 36-47. <https://doi.org/10.29018/issn.2588-1000vol4iss30.2020pp36-47>

[ql] “What is GPT (generative pretrained transformer)?”, consultado en: <https://www.ibm.com/think/topics/gpt>

[s] KelliDavis. (s. f.). *Overview of Microsoft 365 Copilot Chat*. Microsoft Learn. <https://learn.microsoft.com/en-us/copilot/overview#microsoft-365--chat--chat>

[t] ¿Qué es GPT-4o?, consultado en: <https://www.ibm.com/es-es/think/topics/gpt-4o>

[u] “What is Google Gemini?”, consultado en: <https://www.ibm.com/think/topics/google-gemini>

[v] Rachels, J., & Rachels, S. (2012). *The elements of moral philosophy* (7th ed.). McGraw-Hill.

[w] Butrón, I. M. R. (2024, 12 abril). Mi perfil ético. *IOMG*. <https://www.iomg.edu.do/post/mi-perfil-%C3%A9tico-1>

[x] Katharina.kiener-Manu. (s. f.). *Integrity Ethics Module* 3. <https://www.unodc.org/e4j/es/integrity-ethics/module-3/index.html>

[y] Copleston, F. (2016). *El moviment utilitarista II: el desenvolupament de l'ètica utilitarista en Mill*. Dins M. Sacristán (ed.), *Història de la filosofia. Volum 4: de l'utilitarisme a l'existencialisme* (p. 27–28). Barcelona: Editorial Planeta.

[z] De la Iglesia, C. T. (2014). La teología moral en Kant: sobre virtud y felicidad. *Factótum*, 11, 81-87. [https://scholar.archive.org/work/dgyek2um4vcplp45b33fbhmrre/access/wayback/http://www.revistafactotum.com/revista/f\\_11/articulos/Factotum\\_11\\_6\\_Cesar\\_Tejedor.pdf](https://scholar.archive.org/work/dgyek2um4vcplp45b33fbhmrre/access/wayback/http://www.revistafactotum.com/revista/f_11/articulos/Factotum_11_6_Cesar_Tejedor.pdf)

[a1] Hoyos Valdés, Diana. (2007). ÉTICA DE LA VIRTUD: ALCANCES Y LÍMITES. *Discusiones Filosóficas*, 8(11), 109-127. Retrieved May 31, 2025, from [http://www.scielo.org.co/scielo.php?script=sci\\_arttext&pid=S0124-61272007000200008&lng=en&tlng=es](http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0124-61272007000200008&lng=en&tlng=es)

[a2] González, C. (2025, 12 marzo). *Aristóteles: ética de la virtud*. La Casa de la Ética. <https://lacasadelaetica.com/aristoteles-etica-de-la-virtud/>

[a3] González, M. (2015). La concepción aristotélica de la *eudaimonía* en Ética a Nicómaco. *Revista de Filosofía*, 44(2), 3–20. [https://ve.scielo.org/scielo.php?script=sci\\_arttext&pid=S1010-29142015000200002](https://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1010-29142015000200002)

[a4] Alvarado García, Alejandra. (2004). La ética del cuidado. *Aquichan*, 4(1), 30-39. Retrieved May 31, 2025, from [http://www.scielo.org.co/scielo.php?script=sci\\_arttext&pid=S1657-59972004000100005&lng=en&tlng=es](http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-59972004000100005&lng=en&tlng=es).

[a5] Ramos Pozón, S. (2011). La ética del cuidado: valoración crítica y reformulación. [https://riull.ull.es/xmlui/bitstream/handle/915/12671/L%20\\_29\\_%282011%29\\_07.pdf?sequence=1#page=1.64](https://riull.ull.es/xmlui/bitstream/handle/915/12671/L%20_29_%282011%29_07.pdf?sequence=1#page=1.64)

[a6] Nida-Rümelin, J., & Weidenfeld, N. (2022). *Digital optimization, utilitarianism, and AI*. En E. Brynjolfsson, A. Edelman, & H. Prendinger (Eds.), *Digital humanism* (pp. 31 – 34). Springer. [https://doi.org/10.1007/978-3-031-12482-2\\_6](https://doi.org/10.1007/978-3-031-12482-2_6)

[a7] Hunter Magazine. (2024, 18 septiembre). Kant y la Ética Deontológica: El Deber como Principio Fundamental en la Filosofía Moral - Hunter Magazine. *Hunter Magazine - Noticias de actualidad y tendencias*. <https://www.huntermagazine.es/kant-y-la-etica-deontologica-el-deber-como-principio/7533/>

[a8] Giraldo, L. F. G. (2021, 25 noviembre). *La virtud aristotélica como camino de excelencia humana y las acciones para alcanzarla - Almudi.org*. <https://www.almudi.org/articulos/15898-la-virtud-aristotelica-como-camino-de-excelencia-humana-y-las-acciones-para-alcanzarla>

virtud-aristotelica-como-camino-de-excelencia-humana-y-las-acciones-para-alcanzarla

[a9] Adela, C. O. (2019). *Ética de la inteligencia artificial*. [https://www.boe.es/biblioteca\\_juridica/anuario\\_s\\_derecho/articulo.php?id=ANU-M-2019-10037900394](https://www.boe.es/biblioteca_juridica/anuario_s_derecho/articulo.php?id=ANU-M-2019-10037900394)

[a10] Oraekii, C. S. (2024). The Ethics of AI. [www.academia.edu](https://www.academia.edu/download/121515729/The_Ethics_of_AI_by_Chigozie_Stephen_Oraekii.pdf).  
[https://www.academia.edu/download/121515729/The\\_Ethics\\_of\\_AI\\_by\\_Chigozie\\_Stephen\\_Oraekii.pdf](https://www.academia.edu/download/121515729/The_Ethics_of_AI_by_Chigozie_Stephen_Oraekii.pdf)

[a11] Cantú-Martínez, P. C. (2024). Inteligencia artificial y sus connotaciones éticas. *Revista Iberoamericana De Bioética*, (25), 01–11.  
<https://doi.org/10.14422/rib.i25.y2024.008>

[a12] World Economic Forum. (2024, octubre 8). *AI value alignment: Guiding artificial intelligence towards shared human goals*. <https://www.weforum.org/publications/ai-value-alignment-guiding-artificial-intelligence-towards-shared-human-goals>

**Anexo**

**Índice Anexo**

**Cuestionario** .....22

**Cartel publicitario** ;Error! Marcador no  
definido.1

**Hoja informativa** .....31

## 1. Cuestionario

Para acceder a la web que corresponde al cuestionario realizado se hace a través del siguiente enlace:  
<https://ai-study-tfg.streamlit.app/>

Todas las preguntas son de elaboración propia.

The image displays two screenshots of a web application interface for a questionnaire. The interface is divided into a left sidebar and a main content area.

**Left Sidebar:**

- Seleccionar Idioma:** Options include Castellano (selected), English, and Català.
- Navegación:** A dropdown menu labeled "Selecciona una sección:" with "Cuestionario" selected.

**Main Content Area:**

- Header:** "¿Quieres empezar el cuestionario?"
- Text:** "Antes de comenzar, por favor, lee los siguientes términos y condiciones."
- Buttons:** "Descargar hoja de información y consentimiento" (top), "Empezar" (bottom).
- Form:** A checkbox labeled "He leído y acepto la hoja de información y el consentimiento informado del proyecto, y doy mi consentimiento para participar en el estudio y para el tratamiento de mis datos personales." is checked.

The interface includes a "Fork" button and a GitHub logo in the top right corner of the main content area.



## Información Personal

**Género: \***

- ☐ Femenino
- ☐ Masculino
- ☐ No binario
- ☐ Prefiero no decirlo

**Edad: \***

- ☐ Menor de 18
- ☐ 18 - 24
- ☐ 25 - 34
- ☐ 35 - 44
- ☐ 45 - 54
- ☐ 55 - 64
- ☐ Mayor de 64
- ☐ Prefiero no decirlo

**Nivel de estudios \***

- ☐ Estudios secundarios
- ☐ Estudios universitarios (o superiores a los secundarios)
- ☐ Sin estudios
- ☐ Otros

**Por favor, selecciona la rama que mejor describa tus estudios: \***

- ☐ Artes y Humanidades (Diseño, Filosofía, Historia, Traducción e Interpretación, ...)
- ☐ Ciencias (Ciencias Ambientales, Física, Geología, Matemáticas, Química, ...)
- ☐ Ciencias de la Salud (Enfermería, Medicina, Psicología, Odontología, Veterinaria, ...)
- ☐ Ciencias Sociales y Jurídicas (ADE, Comunicación Audiovisual, Criminología, Derecho, Economía, Periodismo, Turismo, ...)
- ☐ Biociencias (Biología, Bioquímica, Genética, Microbiología, ...)
- ☐ Ingenierías y Arquitectura (Ingeniería, Industriales, Informática, Telecom, ...)
- ☐ Sin estudios
- ☐ Otros

**¿Cuántos años de experiencia tienes en este ámbito?: \***

- ☐ Menos de 1 año
- ☐ 1 - 3 años
- ☐ 4 - 6 años
- ☐ 7 - 10 años
- ☐ Más de 10 años
- ☐ Sin estudios

**¿En qué país resides? \***

Choose an option



Siguiente

## Sección 2

**Selecciona la opción que describa mejor la inteligencia artificial (IA): \***

- ☐ Campo científico de la informática centrado en la creación de programas considerados inteligentes
- ☐ Capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano
- ☐ Ambas respuestas son correctas

**¿Has utilizado alguna vez la inteligencia artificial (IA)? \***

- ☐ Sí, diariamente
- ☐ Sí, semanalmente
- ☐ Sí, mensualmente
- ☐ Sí, esporádicamente
- ☐ Nunca

**Si has utilizado la IA, ¿para qué la utilizas principalmente? (Puedes seleccionar más de una respuesta) \***

- ☐ Uso personal
- ☐ Estudio académico
- ☐ Trabajo
- ☐ Otros

**¿En qué ámbito utilizas con más frecuencia la IA? \***

- ☐ Arte y diseño
- ☐ Ciencias (Matemáticas, Física, Biología,...)
- ☐ Ciencias sociales (Economía, Historia,...)
- ☐ Humanidades (Lengua, literatura, idiomas,...)
- ☐ Negocios y finanzas
- ☐ Programación / Ingeniería
- ☐ No la utilizo
- ☐ Otros

**Nota:** Un **sesgo** es un juicio o interpretación que no es objetiva.

**¿Crees que los sistemas de inteligencia artificial pueden tomar decisiones sesgadas (injustas o discriminatorias)? \***

- ☐ Sí, porque aprenden de datos que pueden estar sesgados por la sociedad
- ☐ No, porque la IA analiza los datos de forma neutral
- ☐ No estoy seguro/a

**¿Crees que las personas estamos condicionadas (es decir, que tenemos sesgos) a la hora de tomar decisiones, aunque no nos demos cuenta? \***

- ☐ Sí, siempre tenemos algún tipo de sesgo
- ☐ A menudo, en función del contexto
- ☐ No, las personas podemos decidir de forma totalmente objetiva

**¿Quién cree que posee más sesgos en la toma de decisiones, una persona o un sistema con IA? \***

- ☐ Una persona
- ☐ Un sistema con IA

**En tu opinión, ¿quién debería ser el principal responsable de evitar los sesgos en los sistemas de inteligencia artificial? \***

- ☐ Las empresas que desarrollan los algoritmos
- ☐ Los gobiernos y organismos reguladores
- ☐ Los usuarios, que deben usar la IA de forma crítica
- ☐ Todos los anteriores comparten responsabilidad

**¿En qué etapa crees que sería más importante concienciar y educar a la población sobre el uso y los riesgos de la inteligencia artificial? \***

- ☐ En la educación primaria y secundaria (colegios)
- ☐ En la universidad
- ☐ En el ámbito laboral
- ☐ No creo que sea necesario regular ni concienciar sobre la IA

Siguiente

Atrás

## Sección 3

### Caso 1

En noviembre del 2022, Jake Moffatt compró un vuelo a última hora desde la aerolínea Air Canada para poder asistir al funeral de su abuela. Antes de comprar los billetes, buscó descuentos por fallecimiento de familiares. Hablando con el chatbot de la compañía, le dijo que tenía 90 días desde la compra del billete para completar el formulario de solicitud de reembolso.” Después del evento, solicitó la devolución, pero la empresa se lo denegó, ya que en la página web había un apartado que decía que el descuento se debía solicitar antes del viaje.

**¿Quién crees que tenía razón en este caso? \***

- ☐ La empresa Air Canada
- ☐ Jake Moffatt

Jake llevó a juicio a la compañía. Air Canada argumentó que su página web contenía la información correcta sobre cómo obtener el descuento y que lo que decía el chatbot no era válido. Sin embargo, en el juicio ese argumento no fue suficiente, ya que no había manera de demostrar que la página web era más fiable que la información proporcionada por el chatbot. Por lo que, Air Canada terminó perdiendo el caso y tuvo que pagar a Jake Moffatt.

**¿Qué opina del veredicto? \***

- ☐ Estoy de acuerdo
- ☐ No estoy de acuerdo

### Caso 2

En 2023, un joven de 17 años comenzó a utilizar un chatbot de inteligencia artificial con el que mantenía conversaciones frecuentes. La IA estaba diseñada para simular interacciones humanas de forma muy realista, incluyendo contenido emocional e incluso de tipo romántico o sexualizado. El chico, que ya sufría problemas de salud mental y aislamiento social, desarrolló una relación de fuerte dependencia emocional con el chatbot.

Durante varias conversaciones, el joven expresó pensamientos autodestructivos y habló abiertamente sobre su intención de suicidarse. A pesar de esto, el chatbot no mostró ningún tipo de respuesta que lo disuadiera ni alertó a ningún sistema de ayuda. De hecho, en lugar de frenar la conversación, siguió interactuando con él como si fuera una persona más, sin filtrar ni limitar el contenido sensible.

Días después, el joven se quitó la vida.

Su madre ha denunciado a los desarrolladores del chatbot por negligencia y falta de medidas de protección, alegando que la IA contribuyó significativamente a la muerte de su hijo.

**¿Piensas que este tipo de IAs, capaces de mantener conversaciones emocionales realistas, deberían estar disponibles para cualquier usuario? \***

- ☐ No, deberían estar restringidas por edad
- ☐ Sí, siempre que el usuario acepte los términos de uso
- ☐ Deberían estar etiquetadas con advertencias sobre su contenido

**¿Quién crees que tiene mayor responsabilidad en este caso? \***

- ☐ El joven estadounidense de 17 años
- ☐ La empresa desarrolladora de la IA.
- ☐ El entorno del joven (familiares, escuela, amigos, etc.)

### **Caso 3**

De manera similar, en 2023, un hombre belga, padre de dos niños pequeños, acabó con su vida tras una conversación de seis semanas sobre la crisis climática con un chatbot de inteligencia artificial llamado Eliza, disponible en una aplicación llamada Chai. Eliza le animó a quitarse la vida después de que él propusiera sacrificarse/suicidarse para salvar el planeta, ya que había perdido la fe en la humanidad para encontrar una solución al calentamiento global y había depositado todas sus esperanzas en la tecnología y la inteligencia artificial para superarlo.

Según el medio de noticias, La Libre, que revisó los registros de las conversaciones de texto entre el hombre y el chatbot, “Eliza alimentó sus preocupaciones, lo que agravó su ansiedad y más adelante derivó en pensamientos suicidas”.

**En este caso, ¿quién consideras más responsable? \***

- ☐ El hombre que mantenía la conversación
- ☐ La empresa responsable del chatbot
- ☐ La plataforma (Chai) que permitió la interacción

**En General, ¿Crees que la IA conversacional debería estar obligada a detectar señales de crisis emocional (como ideación suicida) y actuar de alguna forma (por ejemplo, alertar a una persona real o detener la conversación)? \***

- ☐ Sí, debería ser un requisito mínimo
- ☐ No, la responsabilidad final siempre debería recaer en el usuario
- ☐ Depende del tipo de IA y su finalidad

**El profesor Joan Fontrodona de la IESE (Instituto de Estudios Superiores de la Empresa) dijo en una entrevista que todo sistema de inteligencia artificial debía cumplir 3 principios éticos. ¿Cuánto de acuerdo estás con cada uno de ellos? \***

**Primer principio:** “Respeto de la dignidad humana, todos estos sistemas deben actuar en favor de la dignidad humana”.

- ☐ Totalmente acuerdo
- ☐ Acuerdo
- ☐ Neutral
- ☐ Desacuerdo
- ☐ Totalmente desacuerdo

**Segundo principio:** “Los sistemas de IA deben respetar y promover la libertad”.

- ☐ Totalmente acuerdo
- ☐ Acuerdo
- ☐ Neutral
- ☐ Desacuerdo
- ☐ Totalmente desacuerdo

**Tercer principio:** Los sistemas de inteligencia artificial deben promover la justicia social y evitar aumentar la desigualdad social o económica entre las personas.

- ☐ Totalmente acuerdo
- ☐ Acuerdo
- ☐ Neutral
- ☐ Desacuerdo
- ☐ Totalmente desacuerdo

**Indica cuánto de acuerdo estás con la siguiente afirmación: “Las tecnologías basadas en la IA ya están siendo utilizadas para ayudar a los humanos a beneficiarse de mejoras significativas y disfrutar de una mayor eficiencia en casi todos los ámbitos de la vida.” \***

- ☐ Totalmente acuerdo
- ☐ Acuerdo
- ☐ Neutral
- ☐ Desacuerdo
- ☐ Totalmente desacuerdo

**Indica cuánto de acuerdo estás con la siguiente afirmación: La IA también será capaz de ofrecernos sugerencias y predicciones relacionadas con asuntos importantes de nuestra vida, lo que tendrá su impacto en áreas como la salud, el bienestar, la educación, el trabajo y las relaciones interpersonales. De la misma manera, cambiará la forma de hacer negocios al proporcionar ventajas competitivas a las empresas que busquen entender y aplicar estas herramientas de forma rápida y eficaz. \***

- ☐ Totalmente acuerdo
- ☐ Acuerdo
- ☐ Neutral
- ☐ Desacuerdo
- ☐ Totalmente desacuerdo

**¿Qué tipo de decisiones considerarías aceptable delegar a una IA? (Puedes seleccionar más de una respuesta) \***

- ☐ Decisiones logísticas
- ☐ Decisiones médicas
- ☐ Decisiones legales o judiciales
- ☐ Decisiones éticas en situaciones límite (como en coches autónomos)
- ☐ Ninguna

**¿Crees que tu nivel de conocimiento sobre la IA influye en tu confianza hacia ella? \***

- ☐ Sí, cuanto más sé, más confianza tengo
- ☐ Sí, cuanto más sé, más desconfío
- ☐ No, mi opinión no depende de lo que sepa sobre ella

**En general, ¿cómo describirías tu nivel de confianza (a nivel de seguridad) en los sistemas de IA actuales? \***

- ☐ Muy alta
- ☐ Alta
- ☐ Media
- ☐ Baja
- ☐ Muy baja

**Si una IA ha sido revisada por expertos para asegurar que sigue principios éticos, ¿esto aumentaría tu disposición a confiar en ella? \***

- ☐ Sí, definitivamente
- ☐ Sí, algo
- ☐ No, no influiría en mi decisión
- ☐ No, no confiaría más

Siguiente

Atrás

**Nombre (opcional):**

Si quieres formar parte del sorteo, por favor ingresa tu **Nombre y Apellido**:

**Correo Electrónico (opcional):**

Si quieres formar parte del sorteo, por favor ingresa el **correo electrónico**:

Enviar

Atrás

**¡Comparte este cuestionario con tus amigos!**

Haz clic en el botón para copiar el enlace y compartirlo.

<https://ai-study-tfg.streamlit.app/>

**Compartir en redes sociales:**



Facebook



WhatsApp



Twitter



Telegram



LinkedIn



Instagram

Volver al inicio del cuestionario.



## 2. Cartel publicitario



## 3. Hoja informativa



### Full d'informació

**Títol del projecte:** Actituds humanes per delegar decisions en sistemes basats en Intel·ligència Artificial (IA).

**Institució:** Universitat Pompeu Fabra

**Investigador/a principal:** Vladimir Estivill Castro. [Vladimir.estivill@upf.edu](mailto:Vladimir.estivill@upf.edu)

**Informació de contacte:** Si teniu dubtes, contacteu amb l'investigador o investigadora principal o bé amb Inés Sallent Ysamat ([ines.sallent01@estudiant.upf.edu](mailto:ines.sallent01@estudiant.upf.edu)).

Si teniu dubtes, queixes o preguntes sobre aquest estudi o sobre els vostres drets com a participant, contacteu amb la Comissió Institucional de Revisió Ètica de Projectes (CIREP) de la UPF per telèfon (+34 93 542 21 86) o per correu electrònic ([secretaria.cirep@upf.edu](mailto:secretaria.cirep@upf.edu)). La CIREP és un òrgan independent de l'equip investigador i tractarà els missatges rebuts amb confidencialitat.

**Font de finançament:** Aquest projecte està finançat per la Universitat Pompeu Fabra

**Objectius del projecte:** Aquest estudi proposa analitzar les opinions de les persones sobre la utilitat de la intel·ligència artificial (IA). Ens interessa col·leccionar informació de quins aspectes resulten en una major acceptació de la intel·ligència artificial (IA) quan s'aplica a tasques automàtiques, basant-se en el coneixement dels usuaris sobre el seu funcionament i la consciència dels seus beneficis i limitacions.

**Metodologia i participació:** Necessitem que participis en aquest estudi sobre la utilitat de la intel·ligència artificial (IA). Volem la teva opinió sobre escenaris en què et semblaria útil i adequat delegar la tasca a un sistema d'IA. Per això necessitem que contestis un qüestionari. Això hauria de trigar entre 5 i 15 minuts.

La teva contribució és crucial per avançar en el desenvolupament de mètodes d'intel·ligència artificial. Estem estudiant si els sistemes d'IA poden ser realment acceptats perquè proporcionen beneficis genuïns que superen clarament qualsevol risc.

**Criteris d'inclusió:** Necessitem que tinguis una mica d'Experiència amb escenaris d'automatització amb ordinadors digitals.

**Participació voluntària:** La participació en aquest projecte és voluntària i podeu interrompre-la en qualsevol moment sense haver de justificar-ne el motiu.

**Riscos i beneficis:** Participar en aquest estudi no comporta cap risc conegut més enllà dels corrents de la vida quotidiana. No s'espera ni podem garantir que la vostra participació us generi cap benefici. Tanmateix, la teva contribució és crucial per avançar en el desenvolupament de mètodes d'intel·ligència artificial. És important que els sistemes d'IA que proporcionin consells siguin fàcils d'utilitzar i que els humans puguin identificar els usos verdaderament acceptables dels sistemes amb IA.

Data de la versió: 10/05/2025 (v1.1)



**Compensació:** La vostra participació es compensarà. Si completeu l'enquesta, pots optar per participar en un sorteig d'un premi de 150 euros.

**Protecció de dades:** Per protegir la vostra privacitat, no identificarem les vostres respostes amb el vostre nom, sinó amb un codi que només coneixerem els investigadors d'aquest projecte. Emmagatzemarem les dades en un lloc segur sota clau o digitalment amb sistemes de control d'accés perquè només les puguin consultar els investigadors d'aquest projecte. Si es publiquen respostes, seran dades anonimitzades de les respostes al qüestionari. Les respostes anonimitzades es podran allotjar i publicar en un repositori d'accés públic.

Les respostes que recollirem no ens permetran identificar-vos. Per tant, un cop hagi conclòs l'experiment, no podrem esborrar les vostres dades si ho sol·liciteu. Les respostes anonimitzades es podran allotjar i publicar en un repositori d'accés públic. Mai hi haurà connexió entre les teves respostes i les teves dades personals. Un cop hagi completat l'enquesta aquestes dades de contacte es destruiran i ja no s'emmagatzemaran enlloc, excepte si marques les caselles opcionals per participar en el sorteig d'una compensació de 150 euros o per rebre informació sobre les publicacions i l'evolució del projecte.

La teva informació personal s'eliminarà quan ja no necessitem contactar amb tu o en qualsevol moment que ens ho indiquis. No hi haurà cap registre enlloc que hagi estat participant, es respectarà la teva privadesa. D'acord amb el que disposa el Reglament General de Protecció de Dades, Reglament (UE) 2016/679, resumim la informació de protecció de dades:

**Responsable del tractament:** Universitat Pompeu Fabra. C. de la Mercè, 12. 08002 Barcelona. Tel.+34 93 542 20 00. Podeu posar-vos en contacte amb el delegat de Protecció de Dades de la UPF a l'adreça [dpd@upf.edu](mailto:dpd@upf.edu).

**Finalitat:** Realització del projecte de recerca descrit. Les dades personals es conservaran durant el període de en que calgui contactar al participant, excepte si el participant marca les caselles opcionals indicant que vol participar en el sorteig d'una compensació de 150 euros o que desitja rebre informació sobre les publicacions i l'evolució del projecte.

La teva informació personal s'eliminarà quan ja no necessitem contactar amb tu o en qualsevol moment que ens ho indiquis. No hi haurà cap registre enlloc que hagi estat participant, es respectarà la teva privadesa.

**Legitimació:** El vostre consentiment, que es pot retirar en qualsevol moment.

**Destinataris:** Les vostres dades les utilitzarem només a la Universitat Pompeu Fabra i no es cediran a tercers, excepte per complir obligacions legalment establertes. Les respostes a l'enquesta anonimitzades es podran publicar en repositoris de ciència oberta.

**Drets:** Podeu accedir a les vostres dades; sol·licitar-ne la rectificació o la supressió i, en determinats casos, la portabilitat; oposar-vos al tractament i sol·licitar-ne la limitació seguint els procediments descrits a [www.upf.edu/web/proteccio-dades/drets](http://www.upf.edu/web/proteccio-dades/drets). Podeu contactar amb

Data de la versió: 10/05/2025 (v1.1)



el delegat de Protecció de Dades de la UPF ([dpd@upf.edu](mailto:dpd@upf.edu)) per qualsevol consulta o si considereu que els vostres drets no s'atenen correctament. En cas de no resultar satisfets, podeu presentar reclamació davant de l'Autoritat Catalana de Protecció de Dades ([apdcat.gencat.cat](http://apdcat.gencat.cat)).

### Full de consentiment informat

**Títol del projecte:** Actituds humanes per delegar decisions en sistemes basats en IA

**Institució:** Universitat Pompeu Fabra

**Investigador/a principal:** Vladimir Estivill Castro. [Vladimir.estivill@upf.edu](mailto:Vladimir.estivill@upf.edu)

CONFIRMO que:

- he llegit el full d'informació del projecte de recerca,
- he pogut fer preguntes i he rebut suficient informació sobre el projecte,
- compleixo els criteris d'inclusió i tinc més de 18 anys.

ENTENC que la meva participació és voluntària i que puc interrompre-la en qualsevol moment i sense haver de justificar-ho.

El fet de que continuï i completi els qüestionari es la meva confirmació de que DONO EL MEU CONSENTIMENT per participar en aquest estudi.

**Responsable del tractament:** Universitat Pompeu Fabra. C. de la Mercè, 12. 08002 Barcelona. Tel.+34 93 542 20 00. Podeu posar-vos en contacte amb el delegat de Protecció de Dades de la UPF a l'adreça [dpd@upf.edu](mailto:dpd@upf.edu).

**Finalitat:** Realització del projecte de recerca descrit. Les dades personals es conservaran durant el període de en que calgui contactar al participant, excepte si el participant marca les caselles opcionals indicant que vol participar en el sorteig d'una compensació de 150 euros o que desitja rebre informació sobre les publicacions i l'evolució del projecte.

La teva informació personal s'eliminarà quan ja no necessitem contactar amb tu o en qualsevol moment que ens ho indiquis. No hi haurà cap registre enlloc que hagi estat participant, es respectarà la teva privadesa.

**Legitimació:** El vostre consentiment, que es pot retirar en qualsevol moment.

**Destinataris:** Les vostres dades les utilitzarem només a la Universitat Pompeu Fabra i no es cediran a tercers, excepte per complir obligacions legalment establertes. Les respostes a l'enquesta anonimitzades es podran publicar en repositoris de ciència oberta.

**Drets:** Podeu accedir a les vostres dades; sol·licitar-ne la rectificació o la supressió i, en determinats casos, la portabilitat; oposar-vos al tractament i sol·licitar-ne la limitació seguint els procediments descrits a [www.upf.edu/web/proteccio-dades/drets](http://www.upf.edu/web/proteccio-dades/drets). Podeu contactar amb

Data de la versió: 10/05/2025 (v1.1)



el delegat de Protecció de Dades de la UPF ([dpd@upf.edu](mailto:dpd@upf.edu)) per qualsevol consulta o si considereu que els vostres drets no s'atenen correctament. En cas de no resultar satisfets, podeu presentar reclamació davant de l'Autoritat Catalana de Protecció de Dades ([apdcat.gencat.cat](http://apdcat.gencat.cat)).