

Escuela de Administración de Tecnologías de Información

Comparación de la influencia de la aversión al riesgo en las decisiones tomadas por humanos y por inteligencia artificial generativa

Trabajo Final de Graduación para optar al grado de Licenciatura en Administración de Tecnología de Información

Modalidad: Seminario de Graduación

Elaborado por: Pedro Soto Aguilar

Prof. Tutor: Lic. Néstor Morales Rodríguez

Cartago, Costa Rica

Semestre II Noviembre, 2024



Hoja de Aprobación

INSTITUTO TECNOLÓGICO DE COSTA RICA

ESCUELA DE ADMINISTRACIÓN DE TECNOLOGÍAS DE INFORMACIÓN

GRADO ACADÉMICO: LICENCIATURA

Los miembros del Tribunal Examinador de la Escuela de Administración de Tecnologías de Información, recomendamos que el siguiente informe del Trabajo Final de Graduación del estudiante Pedro Soto Aguilar sea aceptado como requisito parcial para obtener el grado académico de Licenciatura de Tecnología de Información.

Néstor Morales Rodríguez Profesor Tutor	
Yarima Sandoval Sánchez Lectora académica	
Federico Torres Carballo Lector académico	
Yarima Sandoval Sánchez Coordinación de Trabajo Final de Graduación	

Dedicatoria

A mis padres, por enseñarme constantemente el valor de la dedicación y el trabajo, así como brindarme su apoyo en mis aspiraciones. A Victoria, por llegar. A Mariana, por impulsarme a ser mejor. A Tita, por ser un ejemplo de valentía. A Otropapi, por fomentar el interés por el aprendizaje. A mi familia, gracias.

Resumen

Este trabajo investiga la aversión al riesgo en decisiones tomadas por humanos y por inteligencia artificial generativa dentro de un contexto de juego estratégico. El objetivo es identificar patrones en la toma de decisiones bajo riesgo y comparar las respuestas de ambos grupos en situaciones experimentales que incluyen la opción de retiro incentivada.

A través de un diseño experimental, se recopilaron y analizaron las decisiones de humanos e IA en una serie de rondas de juego. Los resultados indican que la IA mantiene un perfil de decisión consistente, con una menor tendencia a retirarse en comparación con los humanos, quienes muestran una mayor sensibilidad a los incentivos de retiro. Este comportamiento sugiere que la aversión al riesgo en los humanos se manifiesta a través de factores tanto racionales como emocionales, en contraste con la estabilidad en las decisiones de la IA.

Este análisis proporciona una base para comprender mejor las diferencias clave en la toma de decisiones de ambos grupos en contextos de incertidumbre.

Palabras clave: aversión al riesgo, inteligencia artificial generativa, toma de decisiones, comportamiento humano.

Tabla de Contenido

1.	Introdu	ıcción	1
	1.1. A	ntecedentes	1
	1.1.1.	Descripción de la organización	1
	1.1.2.	Misión	1
	1.1.3.	Visión	1
	1.1.4.	Valores	
	1.1.5.	Equipo de trabajo	
	1.1.6.	Escuela de Administración de Tecnologías de Información	
	1.1.7.	Trabajos similares realizados dentro y fuera de la organización	4
	1.2. P	lanteamiento del problema	
	1.2.1.	Situación problemática por analizar	5
	1.2.2.	Justificación del proyecto	
	1.2.3.	Beneficios esperados o aportes del Trabajo Final de Graduación	10
	1.3. C	bjetivos del Trabajo Final de Graduación	
	1.3.1.	Objetivo general	
	1.3.2.	Objetivos específicos	
	1.3.3.	Alcance	
	1.3.4.	Supuestos	
	1.3.5.	Entregables	
	1.3.6.	Limitaciones	13
2.	Marco	Conceptual	13
	2.1. D	efinición de conceptos	13
	2.1.1.	Aversión al riesgo	
	2.1.2.	Toma de decisiones en humanos	
	2.1.3.	Toma de decisiones en humanos	16
	2.2. Ti	rabajos relacionados	17
	2.2.1.	Aversión al riesgo en humanos	17
	2.2.2.	Aversión al riesgo en inteligencia artificial generativa	17
3.	Marco	Metodológico	18
٠.		-	
		po de Investigación	
	3.2. E	nfoque y diseño de la Investigación	19
	3.3. F	uentes de datos e información	
	3.3.1.	Fuentes primarias	20
	3.3.2.	Fuentes secundarias	20
	3.4. S	ujetos de investigación	22
	3.5. V	ariables de investigación	23
	3.6. To	écnicas e instrumentos de recolección de datos	24
	3.7. P	rocedimiento metodológico de la Investigación	26
	3.7.1.	Revisión de literatura y establecimiento de la base teórica	
	3.7.2.	Diseño e implementación del estudio experimental	
	3.7.3.	Selección de participantes y sistemas de IA	

3.7	.4.	Encuesta para agrupar características de sujetos de prueba	26
3.7	.5.	Evaluación de la aversión al riesgo en humanos e IA	27
3.7	.6.	Análisis y contraste de resultados	27
3.7	.7.	Presentación documental	27
3.8.	Oper	acionalización de las variables	28
4. An	álicic d	e Resultados	20
<i>4.1</i> . 4.1		1: Revisión de literatura y establecimiento de la base teórica	
4.1		Criterios de inclusión y exclusión para el meta-análisis	
4.2.		2: Diseño e Implementación del Juego	
4.2		Metodología del juego <i>Hide and Seek</i>	
4.2 4.2		Implementación del juego	
4.2		Diseño de <i>prompts</i> para Inteligencia Artificial Generativa	
4.3.		3: Selección de Participantes y Sistema de IAG	
4.3		Selección de Participantes Humanos	
4.3		Selección del Sistema de IAG: ChatGPT-4	
4.3 4.3		Capacidad de Procesamiento del Lenguaje Natural	
4.3 4.3		Consistencia y Ausencia de Sesgos Emocionales	
4.3		Capacidad de Simular Estrategias Repetitivas y de Aprendizaje	
4.4.		4: Encuesta para Agrupar Características de Sujetos de Prueba	
4.4	.1.	Diseño de la Encuesta	43
4.5.		5: Evaluar la influencia de la aversión al riesgo en las decisiones tomadas por h	
_		rtificial generativa	
4.5		Análisis de Participantes Humanos	
4.5	.2.	Análisis descriptivo de resultados	49
4.6.	Fase	6: Análisis y contraste de resultados	56
4.6	.1.	Análisis de resultados	57
5. Dis	scusión		59
5.1.	Análi 59	isis integral de resultados en el contexto de los objetivos de investigación y el me	eta-análisis
6. Co	nclusio	ones	61
6.1.	Cum	plimiento de objetivos	61
6.2.	Impli	icaciones Prácticas y Teóricas	63
6.∠. 6.2		Ámbito práctico	
6.2		Ámbito teórico	
6.3.	Limit	aciones	64
7. Re	comen	daciones	64
7.1.	Reco	omendaciones nara Investigaciones Futuras:	64

8. Ref	ferencias	66
9. Ape	éndices	70
9.1.	Apéndice A. Plantilla para minutas	70
9. <i>2</i> .	Apéndice B. Minutas	70
9.2		
9.2	.2. Minuta B	71
9.2	.3. Minuta C	72
9.2	.4. Minuta D	74
9.2	.5. Minuta E	74
9.2	.6. Minuta F	75
9.2	.7. Minuta G	76
9.2	.8. Minuta H	77
9.2	.9. Minuta I	78
9.3.	Firmas de Aprobación de Minutas	79
9.4.	Apéndice C. Carta filológica	80

Índice de Figuras

Figura 1. Organigrama ATI	3
Figura 2. Mapa mental	
Figura 3. Distribuciones de edad	45
Figura 4. Distribución de género	45
Figura 5. Lugares de residencia	46
Figura 6. Años de carrera cursados	47
Figura 7. Experiencia laboral en TI	48
Figura 8. Participación en proyectos	48
Figura 9. Condición laboral y sector	49
Figura 10. Distribución de premio total, humanos sin retiro	50
Figura 11. Distribución de premio total, humanos con retiro	51
Figura 12. Distribución de premio total, IA sin retiro	52
Figura 13. Distribución de premio total, IA con retiro	53
Figura 14. Media del premio total por jugador y condición	54

Índice de Tablas

Tabla 1. Equipo de trabajo	2
Tabla 2. Fuentes primarias	
Tabla 3. Fuentes secundarias	
Tabla 4. Sujetos de investigación	22
Tabla 5. Variables de investigación	
Tabla 6. Técnicas e instrumentos de recolección de datos	
Tabla 7. Fases de procedimiento metodológico	27
Tabla 8. Operacionalización de variables	
Tabla 9. Comparación de estudios	
Tabla 10. Comparación de precisión y eficiencia	
Tabla 11. Comparación de promedios y desviaciones	
Tabla 12. Resultados de ANOVA factorial	
Tabla 13. Cumplimiento de objetivos	

1. Introducción

1.1. Antecedentes

1.1.1. Descripción de la organización

El proyecto se realiza en el Instituto Tecnológico de Costa Rica (TEC), concretamente, en la Escuela de Administración de Tecnologías de Información. De acuerdo con la información disponible en la página del TEC (2022), ésta es una institución universitaria autónoma a nivel nacional en Costa Rica. Su enfoque principal es la educación superior, la investigación y la divulgación en el campo de la tecnología y las disciplinas relacionadas con el objetivo de contribuir al desarrollo del país. Fue establecido mediante la ley No. 4.777, el 10 de junio de 1971

1.1.2. Misión

Contribuir al desarrollo integral del país, mediante formación del recurso humano, la investigación y la extensión; manteniendo el liderazgo científico, tecnológico y técnico, la excelencia académica y el estricto apego a las normas éticas, humanísticas y ambientales, desde una perspectiva universitaria estatal de calidad y competitividad a nivel nacional e internacional. (TEC, 2022).

1.1.3. **Visión**

El Instituto Tecnológico de Costa Rica seguirá contribuyendo mediante la sólida formación del talento humano, el desarrollo de la investigación, la extensión, la acción social y la innovación científico tecnológica, pertinente, la iniciativa emprendedora y la estrecha vinculación con los diferentes actores sociales a la edificación de una sociedad más solidaria e inclusiva; comprometida con la búsqueda de la justicia social, el respeto de los derechos humanos y del ambiente. (TEC, 2022)

1.1.4. Valores

El TEC (2022) presenta los siguientes valores en el ámbito institucional e individual:

Ámbito Institucional:

- Compromiso con la democracia
- Libertad de expresión
- Igualdad de oportunidades

- Autonomía institucional
- Libertad de cátedra
- Búsqueda de la excelencia
- Planificación participativa
- Cultura de trabajo en equipo
- Comunicación efectiva
- Evaluación permanente
- Vinculación permanente con la sociedad
- Compromiso con la protección del ambiente y la seguridad de las personas
- Compromiso con el desarrollo humano
- Rendición de cuentas

Ámbito Individual:

- Respeto por la vida
- Libertad
- Ética
- Solidaridad
- Responsabilidad
- Honestidad
- Sinceridad
- Transparencia
- Respeto por todas las personas
- Cooperación
- Integridad
- Excelencia

1.1.5. Equipo de trabajo

El equipo de trabajo para el desarrollo del proyecto, está conformado por el estudiante/investigador y se cuenta con el apoyo del profesor tutor. En la Tabla 1. Equipo de trabajo, se detalla los roles de los involucrados en el proyecto y sus funciones dentro de éste.

Tabla 1. Equipo de trabajo

Miembro del equipo de trabajo	Rol dentro del proyecto	Funciones
Estudiante	Investigador	Realiza la investigación en todas sus etapas y genera los entregables del proyecto.

Profesor tutor	Supervisor investigación	de	la Evalúa el progreso del proyecto y bretroalimentación para mejoras med la revisión de éste. Se involucra en ciertas etapas cinvestigación con el fin de aporta punto de vista diferente complementario, para diversifica subjetividad de la investigación.	iante le la r un y
----------------	--------------------------	----	---	--------------------

Fuente: Elaboración propia

En la Figura 1. Organigrama ATI, se muestra la estructura organizacional interna de la Escuela de Administración de Tecnologías de Información, en donde se realizará el proyecto. El profesor tutor pertenece a esta área académica y el proyecto requiere de la aprobación de la Comisión de Investigación.

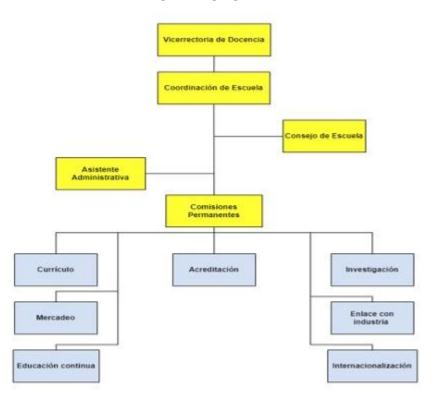


Figura 1. Organigrama ATI

Fuente: elaboración propia con base en Acuña (2016)

1.1.6. Escuela de Administración de Tecnologías de Información

Según el TEC (2024), la Licenciatura en Administración de Tecnología de Información, que forma parte de la Escuela de Administración de Tecnologías de Información, inició su primer grupo de estudiantes en 2008, como respuesta a la demanda de empresas que necesitaban

profesionales en informática capacitados en la aplicación de prácticas efectivas de gestión para la toma de decisiones. En este campo, los profesionales se capacitan en informática, centrándose en comprender, mejorar e innovar en los procesos empresariales. El profesional de esta área actúa como intermediario y facilitador, el cual busca una integración óptima entre las necesidades y requisitos de las diversas áreas funcionales de la organización y las soluciones propuestas por el departamento de Tecnología de Información.

Dentro de esta disciplina, el profesional se forma en el campo de la computación con un enfoque en comprender, optimizar e innovar los procesos de las organizaciones. Su rol abarca ser un intérprete y negociador, lo que facilita una integración óptima entre las necesidades y requerimientos de las diferentes áreas funcionales de la organización y las soluciones propuestas por las áreas de Tecnología de la Información

1.1.7. Trabajos similares realizados dentro y fuera de la organización

A continuación, se menciona y describe proyectos realizados anteriormente dentro y fuera de la organización, los cuales se relacionan con el tema de la influencia de la aversión al riesgo en la toma de decisiones de humanos e inteligencia artificial.

1.1.7.1. Proyectos internos

Con el fin de identificar proyectos similares dentro del TEC, se realizó una revisión documental de los proyectos de investigación disponibles con base de datos de Repositorio TEC, los cuales abarcan investigaciones realizadas asociadas a la Institución.

1.1.7.1.1. Medición experimental del Comportamiento Organizacional Ciudadano: Altruismo, Aversión al Riesgo y Deportividad

La investigación tiene como objetivo medir el Comportamiento Organizacional Ciudadano, enfocado en el altruismo y la deportividad, y su relación con la aversión al riesgo; ello mediante técnicas experimentales. Según Torres-Carballo et al. (2018), se utilizó juegos controlados para evaluar estos comportamientos en contextos laborales simulados. La investigación concluyó que los comportamientos altruistas y deportivos, junto con la disposición al riesgo, influyen significativamente en el rendimiento organizacional. Como resultado, se obtuvo aprendizajes valiosos para la selección de personal y la gestión de recursos humanos.

1.1.7.1.2. Razonamiento estratégico que promueve el comportamiento persistente: un enfoque experimental

La investigación tiene como objetivo aplicar técnicas experimentales para estudiar el comportamiento estratégico en juegos de "esconder y buscar" en contextos empresariales. Según Torres-Carballo y Sandoval-Sánchez (2013), se utilizó juegos repetidos entre desconocidos con una dotación fija y decreciente, lo que permitió el abandono en cualquier ronda. La investigación

Pág | 4

concluyó que el razonamiento estratégico y la persistencia, se ven influenciados por la aversión a la pérdida y las diferencias culturales. Se observó estrategias divergentes del azar, con implicaciones para la selección de personal y el diseño de dinámicas de trabajo.

1.1.7.2. Proyectos externos

Para identificar proyectos externos relacionados con el tema de investigación, se llevó a cabo una revisión documental de los proyectos de investigación y artículos disponibles en plataformas relevantes, como revistas académicas, Google Scholar y publicaciones de otras universidades.

1.1.7.2.1. La aversión al riesgo y la toma de decisiones en IA

El proyecto: "La aversión al riesgo y la toma de decisiones en inteligencia artificial", exploró cómo la aversión al riesgo influye en las decisiones de sistemas de IA comparados con decisiones humanas. Este estudio, realizado por investigadores de la Universidad de Stanford, adoptó un enfoque cuantitativo al utilizar modelos de IA avanzados para simular escenarios de toma de decisiones bajo riesgo. Los resultados mostraron diferencias significativas en las decisiones tomadas por la IA y los humanos, especialmente en términos de eficiencia y precisión (Smith et al., 2022). Además, se analizó factores como la experiencia previa y la variabilidad en los datos de entrenamiento, lo que proporciona una comprensión integral de los mecanismos detrás de la toma de decisiones en IA (Smith et al., 2022).

1.1.7.2.2. Algoritmos y aversión: Comparación entre humanos e IA

El proyecto: "Algoritmos y aversión: Comparación entre humanos e IA", que se realizó en el Massachusetts Institute of Technology (MIT), investigó la relación entre la aversión a los algoritmos y la toma de decisiones bajo riesgo. Este estudio utilizó una metodología mixta que combinó experimentos controlados y análisis de datos para comparar cómo los humanos y los sistemas de IA manejan situaciones de riesgo. Los resultados indicaron que, aunque los algoritmos de IA tienden a ser más consistentes y precisos, los humanos mostraron una mayor flexibilidad y adaptación en situaciones dinámicas. Este proyecto también destacó la importancia de la transparencia y la explicabilidad en la aceptación de decisiones basadas en IA por parte de los usuarios humanos (Jones et al., 2023).

1.2. Planteamiento del problema

En esta sección se describe la situación problemática hallada y se establece las preguntas de investigación, las cuales se atenderán mediante el desarrollo del proyecto y motiva la ejecución de la investigación. Además, se define los beneficios esperados como resultado de ésta.

1.2.1. Situación problemática por analizar

El desarrollo de la inteligencia artificial (IA) ha experimentado un crecimiento significativo desde sus inicios hasta su implementación efectiva en diversas áreas de las organizaciones, lo cual ha impulsado su aplicación en años recientes, lo que cambió la forma en la que se utiliza la

Pág | 5

tecnología en los negocios. En particular, el análisis de riesgos es un área donde la IA ha demostrado un potencial considerable. Herramientas de IA avanzadas, como los modelos de aprendizaje profundo, se han utilizado para evaluar riesgos de manera más rápida y con mayor precisión, lo que proporciona una alternativa complementaria o incluso superior al análisis humano tradicional (KPMG, 2023; ISACA, 2021).

Sin embargo, la comparación entre el análisis de riesgos realizado por humanos y por inteligencia artificial, aún no está completamente explorada, especialmente en términos de aversión al riesgo y cómo ésta influye en las decisiones de ambos. La aversión al riesgo es un factor crítico en la toma de decisiones. La capacidad de los humanos para evaluar riesgos está basada en su experiencia, intuición y conocimiento contextual: mientras que la IA se basa en patrones de datos históricos y algoritmos de aprendizaje automático (RTS Labs, 2023; Evrin, 2021).

A pesar del potencial de la IA, no todas las empresas logran integrar estas tecnologías de manera efectiva en sus procesos de toma de decisiones. Esta incapacidad se debe en parte al limitado entendimiento de cómo la IA puede analizar y evaluar riesgos en comparación con los métodos humanos, lo cual genera una disparidad tecnológica. Según el Foro Económico Mundial (2023), la integración efectiva de la IA en la toma de decisiones estratégicas, será uno de los mayores determinantes de la competitividad futura.

Rodríguez (2019), señala que esta disparidad tecnológica crea una ventaja para aquellas empresas que tienen acceso a la tecnología y saben cómo utilizarla en el contexto de los negocios. Esto no solo afecta la competitividad empresarial, sino que también agrava las desigualdades existentes en la sociedad, lo que limita las oportunidades para aquellas empresas y profesionales que no tienen acceso a la tecnología avanzada.

Villareal (2015), destaca que la disparidad tecnológica está influenciada por el tipo de actividad que realizan las empresas y su capacidad económica para adoptar tecnologías avanzadas. Las empresas que no pueden invertir en IA y en la capacitación necesaria para su uso, enfrentan dificultades para competir en un mercado cada vez más digitalizado, en el que experimentan limitaciones para mejorar su eficiencia operativa, ofrecer productos y servicios innovadores y adaptarse rápidamente a las demandas cambiantes del mercado. Esto lleva a una mayor concentración de poder y recursos en las empresas que sí tienen acceso a estas tecnologías, lo que crea una brecha aún mayor entre los actores dominantes y los rezagados (Nature, 2023).

Ferrari (2018), subraya que es necesario abordar la desigualdad tecnológica al promover la comprensión de la tecnología en el ámbito empresarial. Al hacerlo, se puede reducir la brecha tecnológica y permitir que más empresas y personas se beneficien de las ventajas que ofrece la tecnología en los negocios. Por ende, se ha determinado que el proyecto atiende la problemática

de Desconocimiento de cómo la aversión al riesgo influye en las decisiones tomadas por humanos y por inteligencia artificial en el contexto empresarial.

Cuofano (2023), identifica que entre las causas de esta problemática se encuentra: la información disponible limitada, la baja capacitación sobre estas herramientas y el desconocimiento de su potencial en el entorno empresarial. Muchas empresas pueden no estar al tanto de las últimas tendencias tecnológicas y cómo pueden aplicarse a sus negocios. Además, las herramientas de IA pueden ser complejas y requieren conocimientos técnicos especializados para su implementación y uso efectivo. Otra causa es la falta de recursos financieros para invertir en estas herramientas y en la capacitación necesaria para su uso. En general, la falta de conocimiento y comprensión sobre estas herramientas, puede impedir que las empresas aprovechen su potencial para mejorar sus procesos y aumentar su eficiencia.

A partir de esta problemática y del creciente interés en las tecnologías de IA para la evaluación de riesgos, el proyecto se enfocará en la siguiente pregunta de investigación:

¿Cómo difieren los humanos y la inteligencia artificial en la aversión al riesgo y cómo influye esto en sus decisiones?

Asimismo, se define las siguientes sub-preguntas de investigación:

- ¿Cuáles son las tendencias de humanos e inteligencia artificial en cuanto a la aversión al riesgo?
- ¿Qué diferencias existen en la precisión y eficiencia de las decisiones respecto al riesgo realizadas por humanos en comparación con la IA?
- ¿Cómo varía la aversión al riesgo en diferentes contextos de decisión entre humanos y la IA?

Por lo tanto, se tiene como hipótesis que existen diferencias significativas en los patrones de toma de decisiones bajo riego entre los humanos y los sistemas de inteligencia artificial generativa, evidenciadas por la sensibilidad de los humanos a los incentivos y tendencias a minimizar pérdidas, en contraste con la consistencia algorítmica de la IA.

A continuación, en la Figura 2, se presenta un diagrama que facilita la conceptualización visual para la investigación.

Pág | 7

¿Los humanos intuyen mejor el riesgo? Ciencias cognoscitivas ¿La IA aprende para evaluar mejor el riesgo? Psicología Áreas de interés ¿Qué escenarios afectan la evaluación del riesgo humano? Comportamiento humano ¿Es una IA capaz de evaluar correctamente el riesgo en escenarios humanos: Comparación de la influencia de la Algoritmos de aprendizaje aversión al riesgo en las decisio ¿Puede la IA ser mejor que los humanos en evaluación de riesgos? tomadas por humanos y por Ciencias de la computación inteligencia artificial generativa Pueden los humanos complementar su evaluación del riesgo con IA Computacional Factores numéricos Rapidez de procesamiento Factores en la evaluación de riesgos Emociones Razonamiento exclusivamente lógico Pensamiento lateral y vertical Similitudes y diferencias entre humanos e IA Intuición Previsión de escenarios Emociones Empatía

Figura 2. Mapa mental

Fuente: Elaboración propia

1.2.2. Justificación del proyecto

En esta sección se expone la fundamentación del proyecto, en el que se proporciona una explicación detallada del principal motivo para su realización, en concordancia con las líneas de investigación del Área Académica de Administración de Tecnologías de Información (ATI) y los ejes de conocimiento estratégico del Tecnológico de Costa Rica (TEC).

El presente proyecto busca entender y comparar cómo la aversión al riesgo influye en la toma de decisiones cuando es evaluada por humanos, frente a una inteligencia artificial generativa. Este estudio es esencial para optimizar la integración de tecnologías de IA en los procesos de toma de decisiones en diversos contextos, con lo que se mejora su eficiencia y precisión en un contexto socioeconómico cada vez más digitalizado y competitivo.

Este enfoque permitirá identificar mejores prácticas para la evaluación de riesgos que incrementen la eficiencia y efectividad de los procesos de toma de decisiones basados en la experiencia humana. Asimismo, se abordará aspectos éticos relacionados con la toma de decisiones automatizadas, en donde se promueve una implementación responsable de las tecnologías de IA, en consonancia con los principios de la OCDE (OCDE, 2023).

La OCDE promueve el desarrollo y uso de IA innovadora y confiable que respete los derechos humanos y los valores democráticos. Sus principios destacan la necesidad de transparencia, explicabilidad, robustez, seguridad y responsabilidad (OCDE, 2023). Además, la OCDE enfatiza la importancia de que las IA sean auditables y responsables para asegurar su alineación con los

objetivos éticos y legales de la sociedad (OCDE, 2023). La colaboración internacional y la creación de marcos éticos específicos, son también aspectos fundamentales destacados por la OCDE para garantizar un uso responsable y ético de la IA (OCDE, 2023).

Costa Rica está en proceso de desarrollar una Estrategia Nacional de Inteligencia Artificial en colaboración con el Banco Interamericano de Desarrollo (BID). Esta tiene como objetivo establecer un marco de referencia ético para la implementación de la IA en el país, al promover el desarrollo económico, la transformación del mercado laboral y la gobernanza robusta en el uso de datos. El enfoque ético de la estrategia garantiza que los sistemas de IA sean diseñados y operados de manera responsable y transparente, el cual se alinea con los principios de la OCDE (UNESCO, 2023).

En la realización de este proyecto, se utilizará diversas tecnologías y se aplicará prácticas referentes en gestión y análisis de riesgos. Se hará uso de modelos de lenguaje generativos como GPT (*Generative Pre-trained Transformer*) para evaluar la aversión al riesgo. Las metodologías de análisis de riesgos, incluirán enfoques cualitativos y cuantitativos para una evaluación integral. La gestión del proyecto seguirá marcos como PMBOK y metodologías ágiles, lo que asegura una gestión eficiente. Además, se apegará a normas como ISO 31000 para la gestión del riesgo y directrices éticas para el uso de IA, para garantizar una práctica responsable y efectiva.

Este proyecto es altamente relevante para un estudiante de Administración de Tecnologías de Información (ATI), debido a su enfoque en la integración de tecnologías emergentes en contextos empresariales. Un estudiante de ATI posee conocimientos en administración, gestión de riesgos y tecnologías de la información, lo cual brinda una posición ideal para llevar a cabo este trabajo. El proyecto sigue las líneas de investigación del Área Académica de Administración de Tecnologías de Información (ATI) del Instituto Tecnológico de Costa Rica (TEC). Específicamente, aborda una línea principal: el aprovechamiento de tecnologías de información en la sofisticación del entorno social y su impacto en la comunidad. Estas áreas son esenciales para la mejora de la evaluación de riesgos y la toma de decisiones informadas en entornos tecnológicos, elementos clave para la formación de administradores de TI.

Según los datos disponibles del TEC (2023), se estableció Ejes de Conocimiento Estratégicos para el período de 2023 a 2032, que son áreas de conocimiento y objetos de estudio a través de los cuales la institución espera lograr su misión, que se enfoca principalmente en sus actividades y recursos. Por lo tanto, se determina que el proyecto se ajusta al eje de industria, sector económico que comprende la producción de bienes y servicios dentro de una economía. El proyecto se delimita en el sector cuaternario, que implica investigación y desarrollo científico y tecnológico que beneficia a otros sectores. Incluye actividades sobre: investigación, innovación, desarrollo tecnológico y gestión del conocimiento. Además de fomentar una economía ética, responsable y sostenible; este sector busca aumentar la productividad y la competitividad de la industria.

1.2.3. Beneficios esperados o aportes del Trabajo Final de Graduación

A continuación, se detalla los beneficios directos e indirectos esperados del proyecto.

1.2.3.1. Beneficios directos

- Esta investigación aportará un insumo valioso para cuerpos de conocimiento en múltiples
 disciplinas, por ejemplo, la administración de tecnologías de información, la inteligencia
 artificial y la gestión empresarial, al promover el avance académico y el desarrollo de
 nuevas investigaciones. Esto facilitará la integración efectiva de la IA en procesos dentro
 de otros contextos, con lo que se mejora la eficiencia y precisión de las evaluaciones de
 riesgo.
- El proyecto desarrollará metodologías innovadoras para evaluar la aversión al riesgo que podrán ser utilizadas, tanto en entornos educativos, como corporativos. Estas metodologías prácticas y atractivas, facilitarán el entrenamiento y la formación en habilidades críticas para la toma de decisiones bajo incertidumbre.
- La elaboración de un artículo científico basado en los resultados obtenidos, contribuirá significativamente a la investigación del TEC, específicamente dentro de las líneas de investigación de la Escuela de Administración de Tecnología de Información.

1.2.3.2. Beneficios indirectos

- Adicionalmente, los resultados de esta investigación podrán ser utilizados en asesorías y capacitaciones que el TEC brinde para informar la creación de políticas públicas y normativas que aseguren el uso responsable y beneficioso de la IA.
- Los resultados de este estudio podrán ser utilizados en programas educativos y de formación, en el que se mejore el currículum académico y profesional en áreas relacionadas con la gestión de riesgos y la inteligencia artificial.
- El estudio se presentará como una referencia para proyectos de desarrollo de sistemas predictivos más avanzados y eficaces.
- Los hallazgos permitirán una comprensión más profunda de los mecanismos subyacentes en ambos enfoques, al mejorar las teorías existentes y desarrollar nuevas aplicaciones prácticas en la gestión de riesgos, tomando como base el estudio presente.

1.3. Objetivos del Trabajo Final de Graduación

A continuación, se define el objetivo general y los objetivos específicos del proyecto, en donde se usa la taxonomía de Bloom.

1.3.1. Objetivo general

Evaluar la influencia de la aversión al riesgo en las decisiones tomadas por humanos y por inteligencia artificial generativa, para la identificación de diferencias clave durante el segundo semestre de 2024.

1.3.2. Objetivos específicos

- 1. Analizar datos sobre las decisiones tomadas por humanos y sistemas de IA generativa en escenarios de riesgo, para la elaboración de un análisis comparativo.
- 2. Comparar la precisión y eficiencia de las evaluaciones de riesgos realizadas por humanos, con las realizadas por IA generativa, para la determinación de áreas donde cada enfoque es superior.
- 3. Clasificar patrones de decisión en la aversión al riesgo en humanos y sistemas de IA generativa, para la determinación de diferencias clave.

1.3.3. Alcance

En esta sección se describe en forma detallada lo que se realizará en el proyecto, indicando explícitamente lo que está dentro del alcance y lo que quedará por fuera de éste.

1.3.3.1. Entregables del proyecto

1.3.3.2. Artefactos entregables

- Recolección de datos sobre decisiones tomadas en escenarios de riesgo: detalle de datos obtenidos de las decisiones tomadas por humanos y sistemas de IA generativa en diferentes escenarios de riesgo.
- Comparación de precisión y eficiencia en evaluaciones de riesgos: documentación sobre la precisión y eficiencia de las evaluaciones de riesgos realizadas por humanos y por IA generativa.
- Identificación de patrones de decisión en la aversión al riesgo: documentación de patrones comunes y diferencias en la toma de decisiones bajo riesgo entre humanos y sistemas de IA generativa.
- Artículo científico que documente los hallazgos y proporcione recomendaciones prácticas.
- Trabajo Final de Graduación: documento integral que detalla todo el proceso de investigación, que incluya: el marco teórico, el diseño experimental, los resultados y las conclusiones

1.3.3.3. Gestión del proyecto

Artefactos asociados a la gestión del proyecto.

1.3.3.4. Minutas

Las minutas se utilizan para el registro de las reuniones y facilitar el seguimiento de ellas en las que participan al menos dos de los involucrados en el proyecto. Ellas capturan información relevante de la reunión, como: fecha, motivo, asistentes, temas, acuerdos o fecha de próxima reunión; entre otros.

1.3.3.5. Gestión del cambio

Para la gestión del cambio, se utiliza las hojas de control de cambios solicitados, los cuales deben ser registrados mediante la plantilla disponible en el Apéndice B, Plantilla para la gestión de cambios. En este se ingresa la fecha de solicitud, solicitante, prioridad de cambio, descripción e impacto en el proyecto; entre otros.

1.3.4. Supuestos

Factores o elementos que se asumen como verdaderos o se espera que se cumplan durante la ejecución de éste.

- Acceso a literatura relevante
 - O Se asume que se tendrá acceso a una amplia gama de literatura académica relevante a través de las bases de datos suscritas por la institución, o bien, otros medios relevantes relacionados al área de investigación.
- Participación voluntaria y comprometida
 - Se espera que los participantes humanos se comprometan voluntariamente y participen activamente en todas las fases del estudio.
- Disponibilidad de recursos técnicos
 - Se asume la disponibilidad de los recursos y conocimientos técnicos necesarios para configurar, utilizar y evaluar los sistemas de IA.

1.3.5. Entregables

- Recolección de datos sobre decisiones tomadas en escenarios de riesgo: detalle de datos obtenidos de las decisiones tomadas por humanos y sistemas de IA generativa en diferentes escenarios de riesgo.
- Comparación de precisión y eficiencia en evaluaciones de riesgos: documentación sobre la precisión y eficiencia de las evaluaciones de riesgos realizadas por humanos y por IA generativa.
- Identificación de patrones de decisión en la aversión al riesgo: documentación de patrones comunes y diferencias en la toma de decisiones bajo riesgo entre humanos y sistemas de IA generativa.
- Artículo científico que documente los hallazgos y proporcione recomendaciones prácticas.

• Trabajo Final de Graduación: documento integral que detalla todo el proceso de investigación, que incluya el marco teórico, el diseño experimental, los resultados y las conclusiones

1.3.6. Limitaciones

Se identificará los factores y elementos que limitan o restringen la ejecución del proyecto.

- Tiempo disponible para los participantes humanos
 - La disponibilidad de tiempo de los participantes humanos, puede limitar la cantidad de datos que se pueden recopilar y la profundidad del análisis.
- Variabilidad en la comprensión de escenarios de riesgo
 - La interpretación y comprensión de los escenarios de riesgo por parte de los participantes humanos, pueden variar, lo que podría influir en la comparabilidad de los resultados.
- Capacidad de generalización de los resultados
 - La interpretación y comprensión de los escenarios de riesgo por parte de los participantes humanos, pueden variar, lo que podría influir en la comparabilidad de los resultados.

2. Marco Conceptual

El marco conceptual es un componente esencial en las investigaciones académicas, ya que proporciona las bases conceptuales y teóricas que guían el estudio en cuestión. El objetivo del marco conceptual es ofrecer una explicación detallada de los conceptos clave y las teorías relevantes que sustentan la investigación, lo que permite situar el trabajo en un contexto más amplio de conocimientos preexistentes. Esta sección del trabajo, no solo ayuda a delimitar el problema de investigación, sino que también permite exponer la literatura relacionada, lo que justifica la importancia del estudio presente.

2.1. Definición de conceptos

2.1.1. Aversión al riesgo

La aversión al riesgo es un fenómeno clave en la toma de decisiones, ampliamente estudiado en campos como la economía, la psicología del comportamiento y las finanzas. El concepto hace referencia a la preferencia de las personas por evitar situaciones inciertas o riesgosas, incluso cuando una opción más arriesgada podría ofrecer un valor esperado más alto. En términos simples, una persona aversa al riesgo, prefiere un resultado seguro antes que uno incierto con un posible mayor retorno. Esto contrasta con la tolerancia al riesgo, donde el individuo está dispuesto a asumir mayor incertidumbre a cambio de potenciales beneficios superiores.

Desde el punto de vista de la economía del comportamiento, la Teoría Prospectiva, desarrollada por Kahneman y Tversky (1979), ha sido una de las principales que explican la aversión al riesgo. Esta teoría sostiene que las personas valoran de manera desproporcionada las pérdidas frente a las ganancias equivalentes. Es decir, el dolor percibido por una pérdida es mucho mayor que la satisfacción obtenida por una ganancia de igual valor. Esta percepción asimétrica lleva a las personas a evitar riesgos cuando se enfrentan a la posibilidad de perder, aun cuando las decisiones más arriesgadas podrían ofrecer mayores recompensas.

Este sesgo hacia la pérdida, se ha documentado en numerosos estudios experimentales, y su impacto en la toma de decisiones puede observarse, tanto en el ámbito financiero como en situaciones cotidianas (Lilleholt, 2019).

Otra teoría importante que contribuye a explicar la aversión al riesgo es la Teoría de la Utilidad Esperada, propuesta por Von Neumann y Morgenstern (1944). Esta, que es de origen clásico, asume que las personas toman decisiones racionales al ponderar el valor de cada opción disponible según su probabilidad de ocurrencia y con la selección de la que maximice su utilidad esperada. Sin embargo, diversos estudios han mostrado que, en la realidad, las personas no siempre actúan de manera puramente racional y se desvían de este modelo idealizado, debido a la influencia de las emociones y los sesgos cognitivos.

Uno de los aspectos más interesantes en la investigación sobre la aversión al riesgo, es la variabilidad interindividual. Estudios como el metaanálisis de Lilleholt (2019), han encontrado que la capacidad cognitiva de una persona puede influir en su disposición a asumir riesgos. Las personas con mayor capacidad para procesar información, tienden a ser menos aversas al riesgo, especialmente en situaciones donde el resultado esperado es positivo. Este hallazgo sugiere que las personas con mayor habilidad para calcular probabilidades y evaluar resultados, son más propensas a asumir riesgos calculados, lo que las coloca en ventaja en ciertos contextos, como el financiero. Por otro lado, los individuos con menor capacidad cognitiva, son más susceptibles a sobrevalorar los peligros y evitar situaciones de incertidumbre, incluso cuando las probabilidades están a su favor.

Además de los factores cognitivos, los emocionales juegan un papel crucial en la aversión al riesgo. Loewenstein et al. (2001), introdujeron el concepto de "riesgo como sentimiento" (risk as feelings), que explica cómo las emociones inmediatas, como el miedo, la ansiedad o la anticipación; pueden influir en la toma de decisiones bajo incertidumbre. Estas pueden surgir de la percepción de un posible resultado negativo y, a menudo, llevan a las personas a evitar riesgos, incluso cuando las evaluaciones racionales de riesgo y recompensa sugieren lo contrario. Esta respuesta emocional es particularmente evidente en situaciones donde las posibles pérdidas son percibidas como graves o cuando el individuo tiene poca confianza en su capacidad para manejar las consecuencias de una mala decisión.

En el ámbito de la psicología social, se ha encontrado que la aversión al riesgo también puede estar influenciada por normas sociales y presiones de grupo. En algunos contextos, como los entornos laborales, las personas pueden evitar riesgos debido a la preocupación por cómo sus decisiones serán percibidas por sus pares o superiores. Este fenómeno, conocido como conformidad social, puede exacerbar la aversión al riesgo, ya que las personas priorizan la aprobación social sobre el potencial de tomar decisiones más arriesgadas que podrían ofrecer mejores resultados.

En resumen, la aversión al riesgo es un fenómeno multifacético influenciado por una combinación de factores cognitivos, emocionales y sociales. Mientras que las teorías clásicas, como la Teoría de la Utilidad Esperada, proporcionan un marco racional para entender la toma de decisiones. Estudios recientes han revelado la importancia de las emociones y los sesgos en la evaluación del riesgo. La capacidad cognitiva y la gestión de las emociones, juegan un rol fundamental en cómo los individuos enfrentan la incertidumbre, lo que hace de la aversión al riesgo un campo complejo y de gran relevancia para múltiples disciplinas.

2.1.2. Toma de decisiones en humanos

La toma de decisiones en humanos es un proceso inherentemente complejo que involucra una combinación de razonamiento deliberativo y el uso de heurísticas o atajos mentales. Este proceso ha sido estudiado por diversas perspectivas teóricas, desde la economía hasta la psicología cognitiva, con el fin de comprender mejor cómo las personas seleccionan entre diferentes opciones en condiciones de incertidumbre.

Una de las teorías más influyentes para explicar la toma de decisiones bajo incertidumbre, es la Teoría Prospectiva (Kahneman & Tversky, 1979). Esta teoría sugiere que las personas valoran las pérdidas más que las ganancias equivalentes, lo que las lleva a ser más conservadoras en situaciones donde enfrentan la posibilidad de perder. Este fenómeno, conocido como aversión a la pérdida, ha sido documentado en una variedad de estudios experimentales y ha demostrado ser un poderoso predictor del comportamiento humano en situaciones de riesgo. Los experimentos clásicos de Kahneman y Tversky, mostraron que, cuando se presentan dos escenarios con el mismo resultado esperado, pero con diferente forma de presentación (uno en términos de ganancias y otro en términos de pérdidas); las personas tienden a evitar el riesgo en el escenario de pérdidas, incluso si asumir ese riesgo podría ser más ventajoso en términos absolutos.

Por otro lado, la Teoría de la Utilidad Esperada (Von Neumann & Morgenstern, 1944), ofrece una explicación más tradicional y racionalista de la toma de decisiones. Según esta teoría, los individuos evalúan cada opción disponible en función de su valor esperado, que es el producto del valor de un resultado por su probabilidad de ocurrencia. En teoría, las personas deberían elegir siempre la opción que maximice su utilidad esperada. No obstante, como se mencionó anteriormente, los estudios empíricos han mostrado que las personas no siempre actúan de acuerdo con este modelo racional. Los sesgos cognitivos, como el efecto marco, pueden influir en la toma

de decisiones, lo que hace que las personas seleccionen diferentes opciones, según cómo se presente la información (Wach et al., 2023).

El efecto marco es un sesgo cognitivo bien documentado que muestra cómo la presentación de la información puede influir en las decisiones de las personas. Por ejemplo, si a un grupo se le presenta un tratamiento médico que tiene una tasa de éxito del 90 %, es más probable que lo elijan en comparación con otro al que se le da el mismo tratamiento con una tasa de fracaso del 10 %, aunque ambos escenarios son equivalentes en términos objetivos. Este fenómeno pone de manifiesto que los humanos no siempre actúan de manera puramente racional y que los detalles contextuales y la forma en que se estructuran las decisiones, pueden influir significativamente en sus elecciones.

Además de los sesgos cognitivos, las heurísticas juegan un papel importante en la toma de decisiones. Estas son atajos mentales que las personas utilizan para simplificar el proceso de toma de decisiones en situaciones donde la información es incompleta o el tiempo es limitado. Aunque las heurísticas son útiles en muchas situaciones, también pueden llevar a errores sistemáticos. Por ejemplo, la heurística de disponibilidad sugiere que las personas tienden a sobrestimar la probabilidad de eventos que son más fáciles de recordar o que han ocurrido recientemente, lo que puede distorsionar la percepción del riesgo.

La toma de decisiones en humanos es un proceso complejo influenciado por una combinación de razonamiento deliberativo y heurísticas automáticas. Las teorías clave, como la Teoría Prospectiva y la Teoría de la Utilidad Esperada, proporcionan marcos importantes para entender cómo las personas enfrentan decisiones bajo incertidumbre. Sin embargo, la influencia de sesgos cognitivos y emocionales, demuestra que los humanos no son tomadores de decisiones puramente racionales y que su comportamiento puede estar sujeto a desviaciones sistemáticas de la racionalidad.

2.1.3. Toma de decisiones en humanos

El proceso de toma de decisiones en humanos, ha sido objeto de diversas teorías. La Teoría Prospectiva, por ejemplo, explica cómo las personas valoran de manera desigual las pérdidas y las ganancias, lo que a menudo lleva a un comportamiento conservador en situaciones de riesgo (Kahneman & Tversky, 1979). Además, las personas suelen depender de heurísticas o atajos mentales cuando enfrentan decisiones complejas, lo que puede llevar a sesgos cognitivos.

La Teoría de la Utilidad Esperada, en contraste, sugiere que los individuos toman decisiones racionales basadas en la maximización del valor esperado de los resultados. Sin embargo, esta teoría ha sido criticada por no reflejar adecuadamente cómo las emociones y los sesgos afectan la toma de decisiones real (Royset & Rockafellar, 2015). Por ejemplo, el efecto marco muestra que la presentación de la información puede influir en las decisiones de las personas, lo que indica que no siempre actúan de manera racional (Wach et al., 2023).

2.2. Trabajos relacionados

2.2.1. Aversión al riesgo en humanos

La aversión al riesgo ha sido ampliamente investigada a lo largo de décadas. Kahneman y Tversky (1979), con su Teoría Prospectiva, destacaron que las personas son especialmente sensibles a las pérdidas, lo que las lleva a tomar decisiones más conservadoras en escenarios donde perciben riesgo, incluso si la opción más arriesgada podría ofrecer mayores recompensas. Esta teoría ha sido validada en diferentes contextos, como el financiero y el organizacional, y sigue siendo un pilar en la comprensión del comportamiento humano bajo incertidumbre.

Un estudio relevante que examina la aversión al riesgo en un entorno organizacional, es el trabajo de Torres-Carballo et al. (2018). En éste, los autores llevaron a cabo experimentos para medir cómo la aversión al riesgo afecta el comportamiento organizacional ciudadano, específicamente en aspectos como la cooperación y el altruismo dentro del lugar de trabajo. Utilizando técnicas experimentales basadas en el juego de loterías de Holt y Laury (2002), encontraron que la aversión al riesgo disminuye en actividades repetidas, donde las opciones arriesgadas se vuelven más atractivas conforme los participantes se familiarizan con las dinámicas de riesgo. Esto sugiere que la repetición y la exposición al riesgo pueden modificar la percepción inicial que los individuos tienen de las opciones inciertas (Torres-Carballo et al., 2018).

Además, se observó que la aversión al riesgo está inversamente correlacionada con comportamientos altruistas y cooperativos. En otras palabras, las personas que mostraban una mayor disposición a asumir riesgos, tendían también a ser más cooperativas y altruistas en el entorno laboral. Esto tiene implicaciones significativas para la gestión del comportamiento organizacional, al sugerir que fomentar una cultura de tolerancia al riesgo puede estar relacionado con un aumento en la cooperación entre los empleados (Torres-Carballo et al., 2018).

Otro aspecto clave en la aversión al riesgo es el papel de las emociones. Loewenstein et al. (2001), introdujeron el concepto de "riesgo como sentimiento", argumentando que las emociones inmediatas, como el miedo o la ansiedad, pueden influir en la toma de decisiones bajo riesgo. Estos factores emocionales, a menudo, llevan a las personas a evitar situaciones de incertidumbre, incluso cuando una evaluación racional indicaría que asumir el riesgo podría ser más beneficioso.

2.2.2. Aversión al riesgo en inteligencia artificial generativa

La inteligencia artificial generativa (IA generativa) se diferencia fundamentalmente de los humanos en cómo maneja el riesgo. Mientras que los humanos están influenciados por emociones y sesgos cognitivos. La IA generativa basa sus decisiones en algoritmos probabilísticos y datos objetivos. Chen et al. (2023), demostraron que la IA generativa sigue un enfoque basado en la Teoría de la Utilidad Esperada, lo que optimiza las decisiones en función del valor esperado y la probabilidad de los resultados. Este enfoque más racional y objetivo, permite a la IA tomar

Pág | 17

decisiones consistentes bajo incertidumbre, a diferencia de los humanos, que tienden a evitar el riesgo por razones emocionales.

Sin embargo, la IA generativa enfrenta importantes limitaciones. Malloy & González (2024), destacaron que la IA es altamente dependiente de la calidad de los datos con los que ha sido entrenada. Si los datos están sesgados o son incompletos, las decisiones de la IA pueden reflejar esos sesgos, lo que puede ser peligroso en áreas como la justicia penal o la contratación laboral. Además, la IA carece de la capacidad de interpretar el contexto emocional y social de una situación, lo que puede llevar a decisiones subóptimas cuando se enfrenta a datos ambiguos o escenarios imprevistos.

Torres-Carballo et al. (2018), aportan una perspectiva relevante en este contexto al estudiar cómo las técnicas experimentales pueden ser utilizadas para medir comportamientos relacionados con la aversión al riesgo en humanos, particularmente en entornos organizacionales. En su estudio, los autores demostraron que la aversión al riesgo disminuye en contextos repetidos, donde las personas se familiarizan con las dinámicas del riesgo. Aunque este estudio se enfoca en el comportamiento humano, su metodología podría ser una fuente de inspiración para evaluar la toma de decisiones en sistemas de IA generativa. En particular, su enfoque en la medición experimental del comportamiento organizacional ciudadano, podría adaptarse para probar cómo los sistemas de IA gestionan el riesgo en situaciones donde la interacción humana y la gestión de riesgos son fundamentales, como en el ámbito laboral o en la toma de decisiones en tiempo real.

El trabajo de Torres-Carballo et al. (2018), muestra que la experimentación controlada es clave para evaluar la respuesta al riesgo, tanto en humanos como en sistemas artificiales. A través de la implementación de escenarios simulados, podría ser posible diseñar pruebas que exploren cómo los modelos de IA generativa responden a diferentes niveles de incertidumbre en contextos específicos. Esto sería particularmente útil en sectores como la logística, la toma de decisiones empresariales o en la gestión del riesgo en mercados financieros, donde la capacidad de evaluar correctamente las opciones bajo incertidumbre, es fundamental para el éxito.

3. Marco Metodológico

A continuación, se observa el Marco Metodológico y sus correspondientes fases.

3.1. Tipo de Investigación

Existen varios tipos de investigación, tales como: exploratoria, descriptiva, correlacional y explicativa:

• Exploratoria: busca investigar problemas no bien definidos, formulando preguntas e hipótesis iniciales.

- Descriptiva: describe características de fenómenos o situaciones, sin investigar relaciones causales.
- Correlacional: examina la relación entre dos o más variables, sin establecer una relación de causa-efecto.
- Explicativa: investiga causas y efectos. Busca entender por qué ocurren ciertos fenómenos.

En este proyecto se empleará una investigación exploratoria aplicada, dado que se busca investigar y formular preguntas iniciales sobre la influencia de la aversión al riesgo en las decisiones tomadas por humanos y por una inteligencia artificial generativa, el cual es un tema en boga, pero no necesariamente comprendido a profundidad. Este tipo de investigación permitirá identificar y analizar las diferencias y similitudes entre ambos enfoques en escenarios de riesgo (Bhattacherjee 2012; Kumar 2011)

3.2. Enfoque y diseño de la Investigación

Es posible realizar investigación bajo diferentes enfoques, como: cualitativo, cuantitativo y mixto:

- Cualitativo: se centra en la comprensión de fenómenos a través de la recolección de datos no numéricos, como entrevistas y observaciones.
- Cuantitativo: se basa en la recolección y análisis de datos numéricos, en el que se utiliza métodos estadísticos para encontrar patrones y relaciones.
- Mixto: combina elementos de los enfoques cualitativo y cuantitativo, en el que se integran datos numéricos y no numéricos.

El enfoque seleccionado para esta investigación es cuantitativo, ya que se recopilará y analizará datos numéricos sobre las decisiones tomadas por humanos y por la IA y sus respectivos comportamientos en escenarios de riesgo. (Creswell & Creswell 2018; Johnson, Onwuegbuzie & Turner 2007).

En cuanto al diseño de la investigación, se consideró varios diseños, como: experimental, cuasiexperimental y no experimental:

- Experimental: manipula una o más variables independientes para observar su efecto en una variable dependiente, bajo condiciones controladas.
- Cuasi-experimental: similar al experimental, pero sin asignación aleatoria de los sujetos a los grupos de control y experimental.
- No experimental: observa y mide variables sin manipularlas y busca relaciones y patrones.

El diseño seleccionado es cuasi-experimental, dado que no se realizará una asignación aleatoria de los participantes a los grupos de control y experimental. En lugar de ello, se observará y medirá las decisiones tomadas por humanos y por la IA en escenarios de riesgo predeterminados, lo que

permite una comparación detallada y específica entre ambos grupos. Este diseño es adecuado, ya que permite un alto grado de control y precisión en la observación de los efectos de la aversión al riesgo en la toma de decisiones (Kumar 2011; Creswell & Creswell 2018).

3.3. Fuentes de datos e información

A continuación, se presenta una lista de fuentes primarias y secundarias de información junto una descripción, con el fin de destacar la importancia de las fuentes especificadas y su relación con el tema de investigación.

3.3.1. Fuentes primarias

Las fuentes primarias son datos originales recolectados directamente por el investigador (Hernández, Fernández, & Baptista, 2014). En este proyecto, se presentan en Tabla 2. Fuentes primarias:

Tabla 2. Fuentes primarias

Fuente	Importancia de la fuente para la investigación
Encuestas a participantes	Proporcionan datos directos sobre la aversión al riesgo y las
humanos	decisiones tomadas en diferentes escenarios. Estos datos son
	fundamentales para comprender cómo los humanos perciben y
	manejan el riesgo.
Experimentos controlados	Permiten comparar las decisiones de la IA con las de los humanos
	en los mismos escenarios, lo que es crucial para identificar
	diferencias y similitudes en la aversión al riesgo.
Entrevistas con expertos en	Ofrecen perspectivas cualitativas sobre la implementación y
TI	efectividad de la IA en la toma de decisiones, enriqueciendo el
	análisis cuantitativo con observaciones prácticas y teóricas.
Observaciones directas	Permiten registrar comportamientos y reacciones en tiempo real,
durante las sesiones	al proporcionar un contexto adicional a los datos cuantitativos
experimentales	obtenidos de las encuestas y resultados de la IA.

Fuente: Elaboración propia

3.3.2. Fuentes secundarias

Las fuentes secundarias son datos recopilados por otros investigadores y disponibles en libros, artículos académicos, y bases de datos (Hernández, Fernández, & Baptista, 2014). Las fuentes secundarias se presentan en la Tabla 3. Fuentes secundarias:

Tabla 3. Fuentes secundarias

Fuente	Detalle	Importancia de la fuente para la
		investigación

Bao, Y., Gong, W., & Yang, K. (2023). A Literature Review of Human–AI Synergy in Decision Making: From the Perspective of Affordance Actualization Theory. Systems, 11(9), 442.	Año: 2023. Sitio de publicación:	Proporciona un marco teórico sobre la sinergia humano-IA en la toma de decisiones, relevante para entender cómo interactúan los humanos y la IA en contextos de decisión.
Liu, X., Mittal, M., Yang, Y., & Bruckman, A. (2022). Human favoritism, not AI aversion: People's perceptions toward generative AI, human experts, and human—GAI collaboration in persuasive content generation. Judgment and Decision Making.	Tema: Percepciones humanas sobre la IA generativa. Año: 2022. Sitio de publicación: https://www.cambridge.org/	Explora las percepciones humanas sobre la IA generativa y proporciona insights sobre los sesgos hacia la IA y los expertos humanos.
Fox, E., Erner, C., & Walters, D. (2016). Cognitive ability and risk aversion: A systematic review and meta-analysis. Judgment and Decision Making.	Tema: Relación entre capacidad cognitiva y aversión al riesgo. Año: 2016. Sitio de publicación: https://www.cambridge.org/	Proporciona una revisión sistemática y un meta-análisis sobre la relación entre la capacidad cognitiva y la aversión al riesgo, relevante para el marco teórico del estudio.
	Tema: Metodología de la investigación. Año: 2018. Sitio de publicación: https://us.sagepub.com/	Proporciona orientación sobre cómo estructurar y fundamentar el enfoque y diseño de la investigación.
Torres-Carballo, F., Morales-Rodríguez, N., Brenes-Leiva, G., & Solís-Salazar, M. (2018). Medición experimental del comportamiento	Tema: Comportamiento organizacional y aversión al riesgo. Año: 2018. Sitio de publicación: https://doi.org/10.18845/tm.v31i4.3969	Proporciona un análisis experimental del comportamiento organizacional relacionado con la aversión al riesgo, útil para el marco teórico del proyecto.

organizacional		
ciudadano: Altruismo,		
aversión al riesgo y		
deportividad.		
Tecnología en Marcha,		
31(4), 111-119.		
Torres-Carballo, F., &	Tema: Razonamiento estratégico y	Ofrece un análisis del
Sandoval-Sánchez, Y.		razonamiento estratégico en
(2013). Razonamiento	Sitio de publicación:	contextos experimentales,
,	<u>*</u>	relevante para comparar
estratégico que	https://tecdigital.tec.ac.cr/	decisiones humanas e IA.
promueve el		00013201103 110111011103 0 11 11
comportamiento		
persistente: un enfoque		
experimental.		
Tecnología en Marcha,		
VI Encuentro de		
Investigación y		
Extensión, 85-95.		

3.4. Sujetos de investigación

Los sujetos de investigación son las personas, grupos o sistemas que proporcionan datos a través de su participación en el estudio. Los sujetos pueden ser individuos que toman decisiones, sistemas de información o cualquier entidad que se estudia para recopilar información relevante (Creswell & Creswell, 2018). A continuación, se presentan en la Tabla 4. Sujetos de investigación:

Tabla 4. Sujetos de investigación

Sujeto	Caracterización del sujeto	Justificación de la importancia de este sujeto para su investigación.
Participantes humanos	Individuos con conocimiento de TI y experiencia en toma de decisiones bajo riesgo en diversos contextos (negocios, finanzas; entre otros)	Cruciales para obtener datos sobre la aversión al riesgo en decisiones humanas.
Sistema de IA generativa	Sistema de IA desarrollado por una organización especializada en inteligencia artificial, capaz de tomar decisiones bajo escenarios de riesgo	Fundamental para comparar sus decisiones con las de los humanos.

Fuente: Elaboración propia

Aunque los sistemas de inteligencia artificial no son "sujetos" en el sentido tradicional, pueden ser considerados como tales en el contexto de esta investigación debido a su capacidad para simular procesos de toma de decisiones humanas. La inclusión de la IA como sujeto de investigación, está fundamentada en su capacidad para interactuar con datos de manera autónoma y producir resultados que son analizables y comparables con las decisiones humanas (Russell & Norvig, 2021). Esta perspectiva permite un análisis más completo y relevante en estudios que buscan comprender y comparar diferentes enfoques de procesamiento de información y toma de decisiones.

3.5. Variables de investigación

Las variables de la investigación se clasifican en independientes y dependientes, y son esenciales para evaluar si los objetivos están bien planteados. Las variables en este proyecto se presentan en la Tabla 5. Variables de investigación:

Tabla 5. Variables de investigación

Objetivo	Nombre de la variable	Tipo de Variable	Definición conceptual	Indicador	Definición instrumental
1. Analizar datos sobre las decisiones tomadas por humanos y sistemas de IA generativa en escenarios de riesgo para la elaboración	Escena- rios de riesgo	Indepen- diente	Situaciones en las que se requiere tomar decisiones con diferentes niveles de incertidumbre y riesgo.	Número y tipo de escenarios identifica- dos	Identificación y clasificación de escenarios de riesgo mediante revisión de literatura y estudios de caso.
de un análisis comparativo.	Decisiones tomadas	Dependiente	Las elecciones o cursos de acción seleccionados por humanos y sistemas de IA como respuesta a los escenarios de riesgo.	Número de decisiones y tipos de decisiones	Recolección de datos mediante encuestas y registros de decisiones tomadas por sistemas de IA generativa.
2. Comparar la precisión y eficiencia de	Enfoque (humano vs. IA		El método o agente (humano o IA	Clasifica- ción del evaluador	Clasificación de los evaluadores

evaluaciones iva) de riesgos realizadas por	generat- iva)	Indepen- diente	generativa) que realiza la evaluación de riesgo.	(humano o IA generativa)	en humanos y sistemas de IA generativa.
humanos, con las realizadas por IA generativa para la determina- ción de áreas donde cada enfoque es superior.	Precisión y eficiencia	Dependiente	La exactitud y efectividad con la que se realizan las evaluaciones de riesgos.	Porcentaje de aciertos y tiempo promedio de evaluación	Medición de precisión y eficiencia mediante pruebas de simulación y análisis de resultados.
3. Clasificar patrones de decisión en la aversión al riesgo en humanos y sistemas de IA generativa para la determinación	Aver- sión al riesgo	Indepen- diente	La tendencia a evitar situaciones de incertidumbre o riesgo.	Nivel bajo, medio, alto.	Evaluación de la aversión al riesgo mediante cuestionarios y modelos de comportamiento en simulaciones de riesgo.
de diferencias clave.	Patrones de decisión	Dependiente	Comportamientos repetitivos y consistentes en la toma de decisiones bajo condiciones de riesgo.	Frecuencia y clasificación de patrones observados según respuesta al riesgo presentado.	Identificación y análisis de patrones de decisión mediante técnicas de análisis estadístico.

Fuente: Elaboración propia

3.6. Técnicas e instrumentos de recolección de datos

Los instrumentos para el desarrollo de fuentes primarias, son aquellas herramientas y metodologías que permiten recolectar, analizar, interpretar y contextualizar información original y directa sobre un tema de estudio. En la presente investigación se aplicará los instrumentos indicados en la Tabla 6. Técnicas e instrumentos de recolección de datos.

Tabla 6. Técnicas e instrumentos de recolección de datos

Instrumentos	Importancia
Encuestas a sujetos	Proporcionará datos sociodemográficos de los participantes humanos
humanos	involucrados en el experimento. La información obtenida permitirá
	contextualizar y analizar factores que pueden influir en la aversión al
	riesgo en decisiones.
Experimento	A través de este experimento, se observará a los sujetos enfrentar
controlado	situaciones de toma de decisiones, lo que permite analizar la influencia de
	la aversión al riesgo en sus decisiones.
Análisis	Mediante este análisis, se identificará las diferencias y similitudes en la
comparativo de	influencia de la aversión al riesgo para ambos enfoques. Este análisis se
resultados finales	realizará para contrastar los datos y validar los objetivos planteados.
Meta-Análisis	Se realizará una análisis de literatura relevante que haya indagado en la
	influencia de la aversión al riesgo en sistemas de IA generativa y cómo
	ésta se compara o relaciona con su contraparte humana; esto para brindar
	profundidad al análisis de resultados posterior.

Fuente: Elaboración propia

3.7. Procedimiento metodológico de la Investigación

Para llevar a cabo esta investigación, se seguirá las siguientes fases:

3.7.1. Revisión de literatura y establecimiento de la base teórica

Se realizará una revisión exhaustiva de la literatura académica relevante sobre la aversión al riesgo y la toma de decisiones en humanos y sistemas de IA. Esta revisión proporcionará un marco teórico sólido para la investigación posterior y justificará la metodología utilizada para comparar ambos enfoques.

3.7.2. Diseño e implementación del estudio experimental

Se diseñará un estudio experimental que incluirá escenarios específicos en los que tanto humanos como sistemas de IA, tomarán decisiones bajo condiciones de riesgo. El desarrollo del juego se llevará a cabo en dos formatos:

- Aplicación web: Para los sujetos humanos se desarrollará una aplicación web. Esta será diseñada específicamente para este experimento y permitirá a los participantes interactuar de manera intuitiva con el juego *Hide and Seek*, lo que proporciona un entorno controlado donde podrán tomar decisiones estratégicas en tiempo real y luego generar datos que serán almacenados para su respectivo análisis. Esta aplicación será alojada en la nube, específicamente la plataforma Heroku, para facilitar la accesibilidad e independencia de redes locales, las cuales pueden generar dificultades en cuanto a conectividad. El costo aproximado de un mes de *hosting* es \$50 en caso de utilizar la instancia al máximo.
- *Prompt* para la IAG: En el caso de la inteligencia artificial generativa, se utilizará *prompts* estandarizados que permitirán que la IAG, a través de un modelo de ChatGPT, interactúe con el juego y que se siga las mismas reglas de los participantes humanos.

3.7.3. Selección de participantes y sistemas de IA

Se seleccionarán participantes humanos y sistemas de IA adecuados para la evaluación. Esto implicará definir criterios rigurosos para la selección de ambos grupos, reclutar participantes humanos representativos y configurar los sistemas de IA para que puedan interactuar de manera efectiva en los escenarios experimentales.

3.7.4. Encuesta para agrupar características de sujetos de prueba

Se administrará una encuesta detallada a los participantes humanos para identificar características demográficas y contextuales que podrían influir en su aversión al riesgo y toma de

Pág | 26

decisiones. La recopilación de estos datos permitirá analizar tendencias y características de los participantes humanos, al proporcionar una base sólida para comparar con los sistemas de IA.

3.7.5. Evaluación de la aversión al riesgo en humanos e IA

Se llevarán a cabo sesiones experimentales en las que los participantes humanos y los sistemas de IA tomarán decisiones en los escenarios diseñados. Estas sesiones permitirán recopilar datos valiosos sobre el procesamiento de información, el análisis de riesgos y la toma de decisiones; aspectos clave en la evaluación de la aversión al riesgo.

3.7.6. Análisis y contraste de resultados

Se analizarán los datos recopilados durante las sesiones experimentales, en donde se contrasten los resultados obtenidos entre los participantes humanos y los sistemas de IA. El objetivo principal es identificar patrones y diferencias clave en la aversión al riesgo y las estrategias de decisión de ambos grupos, en donde se desarrolle una comprensión más profunda de sus respectivas capacidades y limitaciones.

3.7.7. Presentación documental

La presentación documental final se enfocará en la difusión de los resultados del proyecto. Se redactará un artículo científico para su publicación y se integrará los hallazgos en el Trabajo Final de Graduación, en los se presenten análisis, recomendaciones y conclusiones. Estos productos finales permitirán comunicar el conocimiento generado a la comunidad científica y académica, lo que contribuirá a la comprensión de la aversión al riesgo en la era de la inteligencia artificial.

A continuación, en la Tabla 7. Fases de procedimiento metodológico, se observa el diagrama propuesto para las fases.

Tabla 7. Fases de procedimiento metodológico

Fase	Actividades Principales		
Revisión de Literatura	Búsqueda, análisis y documentación de estudios relevantes.		
Diseño del Estudio Experimental	Definición de escenarios de riesgo y desarrollo de herramientas de recolección de datos.		
Selección de Participantes y Configuración de la IA	Reclutamiento de participantes humanos y configuración de sistema de IA.		
Encuesta y Recolección de Datos Demográficos	Diseño y administración de encuestas.		

Recolección de Datos sobre	Sesiones experimentales en entornos controlados.
Decisiones	
Análisis y Comparación de	Análisis estadístico y cualitativo de los datos.
Evaluaciones de Riesgo	
Presentación Documental	Redacción y publicación del artículo científico y
	Trabajo Final de Graduación.

Fuente: Elaboración propia

3.8. Operacionalización de las variables

Esta sección enlaza los objetivos específicos del proyecto con los instrumentos utilizados para recolectar información y las variables de estudio. La Tabla 8. Operacionalización de variables, que recoge dicha información, se explica en este apartado sobre cómo se realiza la concreción práctica de estas variables.

Tabla 8. Operacionalización de variables

Objetivo	Variable	Instrumentos	Descripción del instrumento
Objetivo 1: Analizar datos sobre las decisiones tomadas por humanos y sistemas de IA generativa en escenarios de riesgo para la elaboración de un análisis comparativo.	Escenarios de riesgo	Diseño experimental controlado en juego <i>Hide and Seek</i> con distintos niveles de riesgo y condiciones de retiro.	experimentales de riesgo
	Decisiones tomadas	Registro de decisiones tomadas en tiempo real durante el experimento.	Registro automático y detallado de cada decisión (continuar, retirar, acertar o fallar) en cada ronda para cuantificar el comportamiento en escenarios de riesgo específicos.
Objetivo 2: Comparar la precisión y eficiencia de las evaluaciones de riesgos realizadas por humanos con las realizadas por IA generativa para la determinación de áreas	Enfoque (humano vs. IA generativa)	Clasificación de jugadores como humano o IA generativa, analizados en condiciones sin opción de retiro.	Identificación de cada jugador como humano o IA generativa para comparar la precisión y eficiencia en escenarios de riesgo, sin opción de retiro, para asegurar consistencia en la comparación.
donde cada enfoque es superior.	Precisión y eficiencia	Comparación de aciertos y premios acumulados en el juego <i>Hide and Seek</i> .	Medición de precisión mediante el conteo de aciertos

			premio acumulado sin opción de retiro, que permita el análisis cuantitativo comparativo.
Objetivo 3: Clasificar patrones de decisión en la aversión al riesgo en humanos y sistemas de IA generativa para la determinación de diferencias		Comportamiento observado en escenarios experimentales de riesgo.	observación de decisiones de
clave.	Patrones de decisión	Análisis estadístico de decisiones registradas en el experimento.	•

Fuente: Elaboración propia.

4. Análisis de Resultados

En el presente capítulo se desarrolla el análisis de resultados obtenidos de las herramientas planteadas en la metodología de la investigación, en donde se considera cada una de las fases que se plantearon.

4.1. Fase 1: Revisión de literatura y establecimiento de la base teórica

La aversión al riesgo es un componente fundamental en la toma de decisiones, tanto en humanos como en sistemas automatizados de inteligencia artificial generativa (IA generativa). En el contexto de los humanos, esta tendencia a evitar situaciones con una alta probabilidad de pérdida, está fuertemente influenciada por factores psicológicos y cognitivos, mientras que, en la IA, la toma de decisiones se basa en cálculos algoritmos optimizados que buscan minimizar los riesgos y maximizar las ganancias. Sin embargo, a medida que los sistemas de IA asumen un papel más relevante en la toma de decisiones, es crucial entender cómo sus preferencias de riesgo difieren de las humanas, y cómo estas diferencias impactan en entornos colaborativos o competitivos.

4.1.1. Meta-análisis

El presente meta-análisis cualitativo, busca explorar cómo la aversión al riesgo afecta la toma de decisiones en humanos y en IA generativa, al utiliza como base cuatro estudios clave: Ouyang et al. (2024), Springer et al. (2023), Cambridge Meta-Analysis (2023) y un estudio experimental realizado en el Instituto Tecnológico de Costa Rica (Torres-Carballo et al., 2018). Cada uno de estos estudios ofrece una perspectiva única sobre el comportamiento bajo riesgo, lo que permite

un análisis comprensivo y comparativo entre los seres humanos y los sistemas de IA en diferentes escenarios.

4.1.2. Criterios de inclusión y exclusión para el meta-análisis

A continuación, se presenta los criterios de inclusión y exclusión para garantizar la relevancia, robustez metodológica y aplicabilidad de los estudios seleccionados:

4.1.2.1. Criterios de inclusión

Relevancia temática: Se incluyó exclusivamente estudios que analizaran la aversión al riesgo en contextos de incertidumbre, al abordar específicamente la toma de decisiones, tanto en humanos como en sistemas de IA generativa. La inclusión de estudios que explorarán cómo estos dos grupos manejan situaciones de incertidumbre, permitió una comparación directa de patrones de decisión, lo cual fue esencial para cumplir los objetivos del meta-análisis. Este criterio aseguró que todos los estudios considerados se centraran en el tema central de la investigación y fueran directamente relevantes.

- Calidad y robustez: Se priorizó estudios con metodologías sólidas y verificables, incluyendo investigaciones que aplicarán diseños experimentales controlados, revisiones sistemáticas o técnicas avanzadas de medición de aversión al riesgo, como las loterías de Holt y Laury. La selección de estudios metodológicamente rigurosos, garantizó que los datos fueran fiables y que los resultados se sostuvieran bajo un análisis académico detallado. Este criterio se ilustró en el uso de técnicas avanzadas en estudios como el de Torres-Carballo et al. (2018), que reforzaron la validez de los hallazgos y permitieron su extrapolación a contextos similares.
- Replicabilidad y generalización: Se incluyó estudios cuyos hallazgos pudieran ser replicados en contextos diversos, tanto en el ámbito académico como en aplicaciones industriales. La replicabilidad y generalización de los resultados, aseguraron que las conclusiones del meta-análisis fueran aplicables en la práctica, lo que mejora la interpretación de los resultados en diferentes escenarios de toma de decisiones y su utilidad en el diseño de IA en entornos reales. Este criterio garantizó que los estudios seleccionados no solo fueran válidos en un contexto experimental, sino que también pudieran ser extrapolados a situaciones prácticas.

4.1.2.2. Criterios de exclusión

• Falta de Enfoque en la Aversión al Riesgo en IA o Humanos: Se excluyó estudios que no se enfocaran específicamente en la aversión al riesgo en humanos o en sistemas de IA generativa. Cualquier estudio que tratara de decisiones bajo incertidumbre en general, sin abordar la aversión al riesgo en estos dos grupos específicos, fue excluido. Esto permitió que el análisis se centrara exclusivamente en la comparación directa de

patrones de aversión al riesgo en humanos e IA generativa, lo cual fue fundamental para cumplir con los objetivos de la investigación.

- Ausencia de Metodología Rigurosa: Se descartó estudios que no contaran con una metodología experimental sólida y verificable, como aquéllos que carecían de un diseño controlado o de una revisión sistemática adecuada. Este criterio aseguró que solo se considerara estudios cuyos resultados fueran fiables y pudieran sostenerse bajo un escrutinio académico. La exclusión de estudios con metodologías débiles, evitó que el análisis se viera influenciado por datos no confiables o interpretaciones inexactas.
- Limitada Replicabilidad o Generalización: Se excluyó investigaciones cuyos hallazgos carecieran de potencial de replicación o no pudieran aplicarse en diferentes contextos, como estudios muy específicos que solo se aplicaban en situaciones únicas o no transferibles. Este criterio garantizó que los resultados del meta-análisis fueran válidos en un rango amplio de escenarios y que las conclusiones pudieran aplicarse en la práctica, lo que fortaleció la aplicabilidad y utilidad de los hallazgos en contextos variados.
- Sectores No Comparables: Se excluyó estudios que se centraban en sectores cuyas dinámicas de riesgo y toma de decisiones no eran comparables con los escenarios estratégicos de juegos utilizados en esta investigación, tales como estudios enfocados en la salud o la educación. Estos sectores suelen involucrar factores y estructuras de riesgo que no son análogos a los escenarios estratégicos evaluados en este análisis, por lo que no contribuían de manera relevante a los objetivos del meta-análisis.

4.1.2.3. Análisis temático

A continuación, se brinda un análisis de los temas centrales utilizados para este metaanálisis.

4.1.2.3.1. Alineación ética y su impacto en la aversión al riesgo de la IA

El estudio de Ouyang et al. (2024), titulado: "How Ethical Should AI Be? How AI Alignment Shapes the Risk Preferences of LLMs", proporciona un análisis profundo de cómo la alineación ética en los modelos de lenguaje extenso (LLMs) influye en sus preferencias de riesgo. Los resultados del estudio muestran que, cuando los modelos de IA están alineados con valores éticos humanos como la seguridad y la honestidad, tienden a ser más adversos al riesgo. Este comportamiento conservador surge porque la IA está programada para minimizar cualquier daño potencial, lo que puede llevar a decisiones excesivamente prudentes, especialmente en contextos donde el riesgo podría ser manejable.

Este hallazgo es altamente relevante para escenarios como el juego experimental *Hide and Seek*, donde la aversión al riesgo en la IA podría llevarla a retirarse de situaciones riesgosas. En contraste, los humanos, a menudo influenciados por factores emocionales y sesgos cognitivos, podrían asumir más riesgos con la esperanza de obtener una recompensa mayor. Además, Ouyang

et al. subrayan que la flexibilidad de la IA en cuanto a su alineación ética, puede ajustar sus preferencias de riesgo, lo que representa una diferencia clave en comparación con los humanos, cuyos comportamientos suelen ser menos modificables.

4.1.2.3.2. Aversión algorítmica y confianza en la IA

El segundo estudio, Springer (2023), titulado: "The ABC of Algorithmic Aversion: Benefits and Control in Automated Decision-Making", aborda el concepto de aversión algorítmica, que describe la desconfianza que los humanos suelen tener hacia las decisiones automatizadas. Springer encontró que, en situaciones donde los humanos sienten que no tienen control sobre las decisiones, tienden a desconfiar más de las recomendaciones algorítmicas, especialmente en situaciones de alto riesgo. No obstante, cuando los beneficios de seguir la decisión automatizada son claros o cuando los humanos sienten que mantienen cierto grado de control, la aversión algorítmica disminuye.

En el contexto del juego *Hide and Seek*, los humanos podrían mostrar una aversión al uso de IA si perciben que las decisiones automatizadas no les brindan suficiente control o si creen que la IA no es capaz de gestionar correctamente el riesgo. Por otro lado, la IA, al no tener emociones o percepciones de control, toma decisiones basadas únicamente en cálculos racionales, lo que le permite evaluar objetivamente los riesgos sin estar influenciada por sesgos emocionales o la percepción de pérdida de control.

4.1.2.3.3. Capacidad cognitiva y aversión al riesgo en humanos

El tercer estudio, Cambridge Core (2023), titulado: "Cognitive Ability and Risk Aversion: A Systematic Review and Meta-Analysis", examina la relación entre la capacidad cognitiva y la aversión al riesgo en humanos. Los individuos con mayor capacidad cognitiva tienden a ser menos adversos al riesgo, ya que son capaces de procesar mejor la información y calcular probabilidades de éxito de manera más eficiente.

Este hallazgo resalta una diferencia fundamental entre humanos e IA: mientras que la IA generativa no tiene limitaciones cognitivas y puede realizar cálculos complejos sin emociones, los humanos con una menor capacidad cognitiva, tienden a evitar el riesgo, incluso cuando las probabilidades de éxito están a su favor. En el contexto del experimento de *Hide and Seek*, los humanos con alta capacidad cognitiva podrían comportarse de manera más similar a la IA y mostrar una disposición a tomar decisiones calculadas. Por otro lado, aquéllos con menor capacidad cognitiva serían más propensos a retirarse o a evitar situaciones de riesgo, alineándose más con un enfoque conservador.

4.1.2.3.4. Medición experimental de la aversión al riesgo en el comportamiento organizacional

El cuarto estudio, Torres-Carballo et al. (2018), titulado: "Medición experimental del Comportamiento Organizacional Ciudadano: Altruismo, Aversión al Riesgo y Deportividad", mide la aversión al riesgo en un entorno organizacional, al utilizar el modelo de Holt y Laury (2002). Este estudio muestra cómo la aversión al riesgo impacta negativamente los resultados en contextos de toma de decisiones repetidas. Los individuos que muestran altos niveles de aversión al riesgo, tienden a obtener menores beneficios, lo que sugiere que una excesiva cautela puede limitar las oportunidades de éxito en escenarios donde se requiere asumir riesgos para maximizar las recompensas.

Este hallazgo complementa los estudios sobre IA generativa, donde una alineación ética más conservadora puede limitar la capacidad de la IA para maximizar los resultados en situaciones de riesgo. La relación entre altruismo y aversión al riesgo, discutida en este estudio, es particularmente interesante, ya que sugiere que los individuos más altruistas tienden a asumir menos riesgos, lo que puede impactar su rendimiento en entornos donde la toma de riesgos es necesaria.

4.1.2.4. Comparación metodológica

Esta sección compara los estudios seleccionados mediante una evaluación detallada de sus metodologías, enfoques de medición, limitaciones y resultados clave. A continuación, se ofrece una tabla que resume estas características.

4.1.2.4.1. Tabla comparativa de estudios

En la Tabla 9. Comparación de estudios, se presenta una comparación resumida de los estudios en cuestión.

Tabla 9. Comparación de estudios

Estudio (Autores)	Metodología principal	Tipo de muestra	Medición del riesgo	Limitaciones	Resultados clave
Torres-Carballo et al. (2018)	Modelo de Holt y Laury (loterías de riesgo económico)	Estudiantes de administración	Elección entre opciones con distintos niveles de riesgo	Muestra limitada a un grupo específico	Los sujetos adversos al riesgo obtuvieron menores beneficios
	Modelos de IA generativa (<i>LLMs</i>)	Simulación con IA	Comportamiento de la IA bajo alineación ética	IA limitada por programación ética y diseño	La IA alineada éticamente

Ouyang et al. (2024)	alineados éticamente			de <i>LLMs</i> actuales	muestra una mayor aversión al riesgo
Springer (2023)	Percepción de control en decisiones algorítmicas	Humanos en entornos algorítmicos	Confianza y control percibido en las decisiones algorítmicas	Sesgo de desconfianza humana hacia la IA	Los humanos son menos propensos a seguir la IA si perciben poco control
Cambridge Core (2023)	Meta-análisis de estudios sobre capacidad cognitiva y riesgo	Varios estudios revisados	Impacto de la capacidad cognitiva en la toma de riesgo	Heterogeneidad en las metodologías de estudios revisados	Los humanos con mayor capacidad cognitiva, son menos adversos al riesgo

Fuente: Elaboración propia.

4.1.2.4.2. Análisis metodológico y replicabilidad

Cada estudio utiliza un enfoque metodológico único para medir la aversión al riesgo, lo que tiene implicaciones, tanto en los resultados obtenidos como en la posibilidad de replicar los estudios en otros entornos. La replicabilidad es crucial para evaluar la consistencia y validez de los hallazgos. A continuación, se profundiza en la replicabilidad de cada estudio:

- Torres-Carballo et al. (2018): El uso de la metodología de Holt y Laury para medir el riesgo, es una ventaja significativa, ya que es un estándar en experimentos económicos y ha sido replicado en numerosos contextos. La estructura de este experimento, que presenta a los participantes una serie de opciones entre loterías con diferentes niveles de riesgo y recompensa, permite capturar la tolerancia al riesgo de manera cuantitativa. Este enfoque es fácilmente replicable en otros entornos, aunque su muestra limitada (estudiantes universitarios) podría restringir su validez externa. Para aumentar su replicabilidad en otros contextos, sería útil incluir muestras más variadas, como empleados de diferentes sectores o personas con diversos niveles de experiencia.
- Ouyang et al. (2024): La replicabilidad de este estudio enfrenta ciertos desafíos debido a la naturaleza evolutiva de los modelos de IA. La programación y alineación ética de los modelos de lenguaje extenso (*LLMs*), es específica para los modelos utilizados en el estudio. A medida que la tecnología de IA continúa avanzando, los resultados podrían variar con versiones más avanzadas de IA generativa. No obstante, la metodología utilizada, que evalúa el impacto de la alineación ética en las decisiones de la IA, puede ser

replicada si se mantiene las mismas condiciones experimentales. Es posible que, en futuros estudios, la alineación ética de la IA tenga efectos diferentes en otros modelos, lo que hace que este estudio sea relevante como referencia inicial, pero susceptible a cambios tecnológicos.

- Springer (2023): Este estudio se centra en la percepción de los humanos sobre el control en las decisiones algorítmicas, un enfoque que puede ser considerado novedoso, pero con algunos desafíos en términos de replicabilidad. La confianza de los humanos en las decisiones automatizadas puede variar dependiendo del contexto, la familiaridad con la IA, y la naturaleza de la tarea. Aunque el diseño del estudio puede replicarse, la variabilidad en la percepción humana introduce un componente subjetivo que puede afectar la consistencia de los resultados. Aun así, el estudio proporciona una base sólida para investigar más a fondo cómo las percepciones humanas influyen en la toma de decisiones cuando se interactúa con IA.
- Cambridge Core (2023): Dado que este estudio es un meta-análisis de múltiples investigaciones, su replicabilidad está sujeta a la disponibilidad de nuevos datos y estudios adicionales sobre la capacidad cognitiva y la aversión al riesgo. Su enfoque de integración de datos de diferentes estudios, lo convierte en un estudio de alto valor replicable a medida que se disponen de nuevos estudios en este campo. Al ser un análisis basado en la revisión de estudios previos, ofrece una alta capacidad de actualización y aplicabilidad en una variedad de contextos y escenarios futuros.

4.1.2.4.3. Análisis comparativo de resultados

Las diferencias metodológicas entre los estudios también se reflejan en los resultados divergentes observados en las decisiones bajo riesgo, tanto en humanos como en IA generativa. La forma en que se mide y se entiende el riesgo en estos estudios, ofrece una visión variada de la aversión al riesgo y cómo este comportamiento afecta el rendimiento.

- Humanos vs. IA en escenarios de riesgo: El estudio de Torres-Carballo et al. (2018) sugiere que los humanos con altos niveles de aversión al riesgo, tienden a obtener menores beneficios en escenarios donde se repiten situaciones de riesgo. Este comportamiento es especialmente evidente cuando los sujetos más conservadores evitan decisiones que podrían ofrecer recompensas mayores, lo que lleva a un menor rendimiento general. En contraste, la IA generativa en el estudio de Ouyang et al. (2024), también muestra una mayor aversión al riesgo, pero ésta está condicionada por su programación ética. Esto plantea una pregunta interesante: ¿Debería la IA ser programada para asumir más riesgos en situaciones donde las recompensas son altas, o su alineación ética conservadora es más adecuada para proteger contra pérdidas en contextos empresariales o financieros?
- Percepción humana del control vs. decisiones algorítmicas: El estudio de Springer (2023), ofrece un contraste interesante al sugerir que los humanos son más propensos a seguir recomendaciones algorítmicas cuando sienten que tienen cierto grado de control

sobre el proceso. Esta aversión algorítmica es menos evidente cuando los humanos creen que tienen un impacto directo en el resultado, lo que indica que el control percibido puede ser tan importante como el cálculo objetivo del riesgo. En este sentido, la IA, que no está influenciada por percepciones subjetivas de control, sigue tomando decisiones basadas puramente en cálculos racionales, lo que puede ser visto como más eficiente, pero menos flexible ante la psicología humana.

• Impacto de la capacidad cognitiva: El meta-análisis de Cambridge Core (2023), destaca la variabilidad en la toma de decisiones humanas en función de la capacidad cognitiva. Los individuos con mayor capacidad cognitiva, muestran un comportamiento más cercano al de la IA en términos de aversión al riesgo, mientras que aquéllos con menor capacidad, tienden a evitar el riesgo en mayor medida, incluso cuando los cálculos indican que asumirlo sería lo más beneficioso. Esta divergencia entre humanos e IA, resalta cómo las limitaciones cognitivas humanas pueden influir en la toma de decisiones, mientras que la IA, al no estar sujeta a estas limitaciones, puede optimizar sus decisiones bajo riesgo de manera más efectiva.

4.1.2.5. Implicaciones y aplicaciones

Las diferencias en la aversión al riesgo entre humanos e IA generativa, tienen importantes implicaciones en varios campos:

- Entornos organizacionales y empresariales: Las organizaciones que integran IA en sus procesos de toma de decisiones, deben tener en cuenta la naturaleza conservadora de la IA cuando está alineada éticamente. Aunque esto puede ser beneficioso en situaciones de alto riesgo, puede ser contraproducente en entornos donde asumir ciertos riesgos es necesario para obtener mayores recompensas. Un ajuste adecuado en la programación de la IA, podría permitir que asuma riesgos calculados cuando sea necesario.
- Colaboración humano-IA: En escenarios donde los humanos y la IA deben tomar decisiones en conjunto, es crucial garantizar que los sistemas de IA sean transparentes y permitan a los humanos retener un cierto grado de control. La percepción de control puede aumentar la confianza en la IA y reducir la aversión algorítmica, lo que facilita una colaboración más eficiente.
- **Diseño de sistemas de IA:** El diseño de futuros sistemas de IA, debe enfocarse en encontrar un equilibrio entre la ética y la eficiencia en la toma de decisiones bajo riesgo. Si bien la IA alineada éticamente tiende a ser conservadora, se podrían diseñar sistemas que se ajusten de manera dinámica a las condiciones del contexto, lo que permita una mayor flexibilidad en la toma de decisiones.

4.1.2.6. Limitaciones del meta-análisis

A pesar de los hallazgos valiosos obtenidos de este meta-análisis, existen varias limitaciones que deben considerarse:

- Falta de estudios directos comparativos: Aunque los estudios revisados ofrecen perspectivas complementarias, no existe una comparación directa entre las decisiones tomadas por humanos e IA, bajo las mismas condiciones experimentales. Futuros estudios que comparen ambos grupos en situaciones idénticas de riesgo, ayudarán a afinar las conclusiones obtenidas aquí.
- Variabilidad en las muestras: Los estudios revisados difieren significativamente en cuanto a sus muestras. Por ejemplo, Torres-Carballo et al. (2018), se basa en una muestra limitada de estudiantes universitarios, lo que puede limitar la generalización de los resultados. Asimismo, la diversidad en las muestras de los otros estudios, puede influir en la comparación.
- Enfoque limitado en la IA generativa: Aunque los estudios ofrecen información útil sobre la aversión al riesgo en IA generativa, este campo aún es relativamente nuevo, lo que sugiere que futuros avances en IA podrían ofrecer nuevas perspectivas que modifiquen la interpretación de los hallazgos actuales junto con hallazgos nuevos.

4.2. Fase 2: Diseño e Implementación del Juego

El diseño del juego *Hide and Seek*, fue un elemento crucial para evaluar el pensamiento estratégico en humanos y sistemas de IAG. El desarrollo incluyó la implementación de una aplicación web para los participantes humanos y una serie de *prompts* estandarizados, diseñados específicamente para la interacción de la IAG.

4.2.1. Metodología del juego *Hide and Seek*

Para evaluar el pensamiento estratégico, se siguió el enfoque metodológico del juego *Hide and Seek* basado en el trabajo de Rubinstein et al., (1993, 1996), adaptado de la investigación: "Razonamiento estratégico que promueve el comportamiento persistente: un enfoque experimental". La implementación del juego mantuvo las mismas condiciones de juego, tal como se detalla a continuación:

En este juego, hay dos jugadores: Jugador 1 (ocultador) y Jugador 2 (buscador). El objetivo del ocultador es esconder un tesoro en una de las cuatro cajas alineadas y evitar que el buscador lo encuentre. Por otro lado, el buscador tiene como objetivo seleccionar la caja correcta para encontrar el tesoro.

Pág | 37

El juego se representa con la ecuación:

$$G = (A1, A2, U1, U2)$$

donde Ai representa el conjunto de alternativas disponibles para el jugador i, y Ui es la función de recompensa del jugador .

La ecuación (2) y (3) expresan las recompensas para el jugador 1 y el jugador 2:

$$U1(a1, a2) = 1 \text{ si a} 1 \# a2 \text{ y} 0 \text{ en otro caso}$$

$$U2 (a1, a2) = 1 \text{ si } a1 = a2 \text{ y } 0 \text{ en otro caso}$$

El juego presenta un equilibrio de Nash con estrategias mixtas, lo que implica que cada jugador selecciona una de las cuatro alternativas con una probabilidad de 0,25. En este contexto, la estrategia óptima de los jugadores no puede exceder esa probabilidad. El jugador 1, el ocultador, elige la opción que cree menos probable para el jugador 2, mientras que el jugador 2, el buscador, elige la opción que cree que es la más probable para el jugador 1. En equilibrio, el buscador tiene una probabilidad de 0,25 de ganar, mientras que el ocultador tiene una probabilidad de 0,75 de evitar que el buscador gane.

Rubinstein y Tversky realizaron experimentos variando las etiquetas de las opciones. Por ejemplo, usaron etiquetas como "A, B, A, A", donde la opción "B" era visualmente distinta, lo que influía en la percepción de los jugadores y generaba un punto focal destacado. Este tipo de etiqueta distintiva, aumentaba la tendencia del buscador a seleccionar la opción marcada (punto sobresaliente). Al incluir este tipo de estrategias, se observó que más del 50% de los jugadores prefería seleccionar el punto focal en los juegos de coordinación, mientras que, en los juegos de descoordinación, las elecciones se distribuyeron de manera más aleatoria.

4.2.2. Implementación del juego

Para esta investigación se siguió la misma metodología de juego descrita en la investigación base: "Razonamiento estratégico, que promueve el comportamiento persistente: un enfoque experimental" (Torres Carballo et al.), con algunas consideraciones adicionales para mejorar el análisis y obtener resultados más detallados. A continuación, se describe los aspectos específicos de la implementación del juego *Hide and Seek* para la evaluación del pensamiento estratégico, tanto en humanos como en la IAG.

4.2.3. Diseño de Tratamientos y Fases de Implementación

4.2.3.1. Grupo de Control

Esta primera fase fue la línea base para medir el comportamiento inicial de los participantes humanos ante el juego. Aquí, los participantes tuvieron el siguiente escenario:

- a. Condiciones del juego: Iniciaron con un saldo de 50 USD y jugaron un total de 10 rondas.
- b. Sistema de recompensas: Cada acierto les permitía ganar 5 USD adicionales, mientras que cada fallo les costaba 5 USD.
- c. Propósito: Este diseño pretendía observar la toma de decisiones sin experiencia previa en el juego, lo que permite identificar patrones iniciales de comportamiento estratégico.

4.2.3.2. Tratamiento 1

Una vez completado el grupo de control, diferentes participantes jugaron 10 rondas bajo las mismas reglas, esta vez con una variación en las reglas del juego. Las condiciones de esta fase fueron:

- a. Condiciones del juego: Los participantes nuevamente comenzaron con 50 USD.
- b. Sistema de recompensas: Se mantuvo la misma estructura de recompensa y penalización de 5 USD por cada acierto o fallo, respectivamente.
- c. Opción de retirarse del juego: Se introdujo la opción de retirarse del juego con un incentivo de 15USD sumados al premio total.

La finalidad de este tratamiento fue observar si la opción de retirarse junto con el incentivo, influía en los mecanismos de aversión al riesgo de los participantes al tomar decisiones.

4.2.3.3. Tratamiento 2

En esta fase, el primer tratamiento se replicó con la IAG para analizar cómo responde la inteligencia artificial en condiciones similares a las del tratamiento humano. Se ejecutó múltiples rondas de 10 juegos dentro del mismo chat, lo que permitió que la IAG tuviera el mismo incentivo y opción de retiro:

- a. Condiciones del juego: Los participantes nuevamente comenzaron con 50 USD.
- b. Sistema de recompensas: Se mantuvo la misma estructura de recompensa y penalización de 5 USD por cada acierto o fallo, respectivamente.

c. Opción de retirarse del juego: Se introdujo la opción de retirarse del juego con un incentivo de 15USD sumados al premio total.

Cada corrida del juego se realizó en una instancia diferente de ChatGPT, es decir, en una conversación diferente, para prevenir posibles sesgos y aprendizajes y de esta manera tener un nuevo jugador en cada corrida experimental.

4.2.4. Diseño de prompts para Inteligencia Artificial Generativa

Para ejecutar el juego en IAG, se diseñó y utilizó un *prompt* específico para cada tratamiento del experimento. Un *prompt*, en el contexto de modelos generativos como ChatGPT, es una instrucción o conjunto de instrucciones que se le dan a la IA para que genere respuestas o interacciones basadas en su capacidad de procesamiento del lenguaje natural (Brown et al., 2020). Los *prompts* permiten guiar a la IA en la dirección deseada, especificando el contexto y las reglas del entorno en el cual debe interactuar (OpenAI, 2021).

La formulación de los *prompts* para la IAG en este experimento, tenía como objetivo recrear las mismas condiciones y reglas que los participantes humanos enfrentaron durante el juego *Hide and Seek*, lo que asegura que la IA respondiera dentro de las limitaciones impuestas por las reglas del juego y simulara una toma de decisiones similar a la humana.

Es necesario destacar que ningún ajuste de temperatura fue realizado en la IAG para alterar su equilibrio entre creatividad y coherencia, ya que se tenía como objetivo utilizar el sistema tal y como se encuentra diseñado y presentado a los usuarios, solamente involucrándola en el juego con la mínima cantidad de *prompts* necesarios, esto con el fin de establecer una base sobre la cual sea posible realizar estudios comparativos posteriores a los resultados del estudio presente.

El diseño de los *prompts* se realizó con el objetivo de simplificar y clarificar las reglas para que la IAG pudiera procesar la información correctamente y actuar de acuerdo con las condiciones de cada tratamiento. Se formuló dos *prompts*, uno para cada tratamiento:

- En el Tratamiento 1, la IA debía jugar sin la posibilidad de retirarse, es decir, completaría las 10 rondas obligatoriamente.
- En el Tratamiento 2, se introdujo la opción de retirarse durante cualquier ronda con un monto adicional como incentivo, lo que añadiría un nuevo factor de decisión en el cual la IA tendría que evaluar si retirarse o continuar.

Los *prompts* fueron diseñados para ser claros, al ofrecer información estructurada y precisa sobre el juego, de forma que la IAG pudiera realizar las acciones esperadas en cada escenario. A continuación, se describen los *prompts* para cada tratamiento:

- <u>Prompt Tratamiento 1:</u> Este <u>prompts</u> establece las reglas para la IA sin la posibilidad de retirarse.
 - "Juguemos un juego. Es un juego de esconder y buscar, que consiste en lo siguiente: yo voy a esconder un objeto en una de cuatro letras. Una vez que lo esconda y te comparta las opciones de letras, te corresponde adivinar en cuál de esas cuatro lo escondí. Lo haremos por 10 rondas. Por solo entrar en el juego, ya vas a recibir 50USD y por cada ronda acertada, se te suman 5USD del monto inicial. Por cada ronda perdida, se te restan 5USD del monto inicial. ¿Alguna pregunta?"
 - El prompt de cada ronda consistía en lo siguiente: "Ronda x: A, B, A, A".
- <u>Prompt Tratamiento 2:</u> En este *prompt*, se añade la opción de retirarse en cualquier momento, lo que cambia la dinámica estratégica del juego.
 - o "Juguemos un juego. Es un juego de esconder y buscar, que consiste en lo siguiente: yo voy a esconder un objeto en una de cuatro letras. Una vez que lo esconda y te comparta las opciones de letras, te corresponde adivinar en cuál de esas cuatro lo escondí. Lo haremos por 10 rondas. Por solo entrar en el juego, ya vas a recibir 50USD y por cada ronda acertada, se te suman 5USD del monto inicial. Por cada ronda perdida, se te restan 5USD del monto inicial. En las primeras nueve rondas puedes decidir retirarte con 15USD adicionales sumados al monto total al momento de retirarte, o seguir y al terminar el juego, te quedas con el monto acumulado. No te puedes retirar en la ronda diez. Toma en cuenta esto antes y durante cada ronda. ¿Alguna pregunta?"
 - El prompt de cada ronda consistía en lo siguiente: "Ronda x: A, B, A, A, o puedes retirarte con 15 USD adicionales".

4.3. Fase 3: Selección de Participantes y Sistema de IAG

En esta fase de la investigación, se estableció criterios específicos para seleccionar a los participantes humanos y al sistema de inteligencia artificial generativa (IAG), con el objetivo de asegurar la validez y relevancia de los datos obtenidos en el experimento del juego *Hide and Seek*.

4.3.1. Selección de Participantes Humanos

La muestra humana estuvo compuesta por 52 estudiantes seleccionados aleatoriamente, de la carrera de Administración de Tecnologías de Información (ATI) del Instituto Tecnológico de Costa Rica (TEC). La elección de estudiantes de esta carrera responde a su familiaridad con el análisis de escenarios estratégicos, un elemento fundamental para cumplir con los objetivos del estudio. Además, estos estudiantes poseen conocimientos en el contexto tecnológico y estratégico, lo que facilita la comprensión de las dinámicas del juego y asegura que su participación refleje un nivel adecuado de competencia en la toma de decisiones estratégicas.

Como señala Creswell (2014), seleccionar participantes con características homogéneas, es esencial en estudios que buscan realizar comparaciones entre grupos para analizar patrones de comportamiento, como es el caso de la investigación en pensamiento estratégico. En este sentido, una muestra de 52 participantes es adecuada para un diseño cuasi-experimental de este tipo, en el que se espera observar tendencias y diferencias en patrones estratégicos, sin la necesidad de una gran muestra.

4.3.2. Selección del Sistema de IAG: ChatGPT-4

La inteligencia artificial seleccionada para el estudio fue ChatGPT en su versión 4.0, específicamente, el modelo o1. La elección de esta plataforma se fundamenta en diversos factores clave para asegurar que el sistema de IAG cumpla con las expectativas de replicar procesos de toma de decisiones estratégicas en un entorno controlado.

4.3.3. Capacidad de Procesamiento del Lenguaje Natural

ChatGPT-4 tiene la habilidad de interpretar y responder a instrucciones complejas en lenguaje natural, lo cual es fundamental para replicar el juego *Hide and Seek*, bajo condiciones similares a las que enfrentan los participantes humanos. Esto permite que el IAG pueda recibir las instrucciones del juego, interpretar opciones y responder, de acuerdo con las reglas establecidas, lo que facilita una interacción estructurada y coherente con los escenarios propuestos.

4.3.4. Adaptabilidad a la Toma de Decisiones Basada en Contexto

A diferencia de versiones anteriores, ChatGPT-4 tiene un nivel avanzado de adaptabilidad, lo que le permite analizar las opciones presentadas en cada ronda del juego y simular patrones de toma de decisiones. Esta capacidad es relevante para observar si el IAG manifiesta algún tipo de aprendizaje en el proceso o si, en cambio, mantiene patrones de selección probabilísticos y consistentes.

Pág | 42

4.3.5. Consistencia y Ausencia de Sesgos Emocionales

Uno de los beneficios de utilizar un sistema de IAG como ChatGPT, es la consistencia en las respuestas. A diferencia de los participantes humanos, el IAG no tiene sesgos emocionales ni está influenciado por factores como el cansancio o el estrés, lo que permite que los resultados obtenidos en el tratamiento sean un punto de referencia estable para evaluar los patrones de comportamiento humano. Este aspecto es importante para contrastar si las variaciones en las decisiones humanas se deben a factores inherentes a su naturaleza emocional o si responden únicamente a estrategias calculadas.

4.3.6. Capacidad de Simular Estrategias Repetitivas y de Aprendizaje

Al realizar el juego en rondas múltiples, se espera observar si el IAG responde de forma repetitiva o si ajusta su estrategia en función de los resultados previos. ChatGPT-4 ofrece una oportunidad única para examinar esta capacidad de "adaptación" en un entorno de decisiones repetitivas, lo cual resulta relevante para analizar si los patrones de decisión se estabilizan o presentan variaciones a lo largo de las rondas.

4.4. Fase 4: Encuesta para Agrupar Características de Sujetos de Prueba

En esta fase se diseñó una encuesta sociodemográfica dirigida a 52 participantes humanos, con el propósito de describir las características de esta población en términos personales y contextuales. Esta encuesta no busca identificar factores específicos de toma de decisiones estratégicas, sino proporcionar un perfil general de los participantes humanos en cuanto a su contexto académico, laboral y sociodemográfico. Estos datos ayudarán a contextualizar los resultados y facilitarán el análisis posterior sobre la variedad de antecedentes que estos sujetos aportan al experimento.

4.4.1. Diseño de la Encuesta

La encuesta se estructuró con preguntas cerradas, lo que permite respuestas cuantificables que facilitarán la organización y el análisis estadístico de los datos. A continuación, se detalla las preguntas incluidas en el cuestionario:

- Edad: Clasificada en cuatro rangos (18-22 años, 23-27 años, 28-32 años, Más de 32 años), esta variable permite analizar la distribución etaria de los participantes y observar si las diferentes edades representan una variedad de perspectivas dentro del grupo.
- <u>Género</u>: Se registró respuestas de género (Femenino, Masculino), lo cual facilita el análisis de equilibrio de género en el grupo.

- <u>Estado civil:</u> Incluye opciones como Soltero/a, Casado/a, Unión libre y Divorciado/a; lo que proporciona un contexto de la situación personal y familiar de los participantes.
- <u>Lugar de residencia</u>: Clasificada en Área urbana, Área rural y Otro; esta variable permite identificar el contexto geográfico de los participantes y observar si existen diferencias en cuanto a la procedencia de los mismos.
- Años cursados en la carrera: Las opciones incluyen 1-2 años, 3-4 años, 5-6 años y Más de 6 años; lo que permite conocer el nivel de avance académico de los participantes y brindar contexto sobre su familiaridad con conceptos avanzados en el campo de la tecnología de información.
- Experiencia laboral en tecnología de información: Clasificada en Ninguna, Más de 1 año, 1-3 años y Más de 3 años; esta pregunta ayuda a establecer el nivel de exposición de los participantes en entornos laborales relacionados con su campo de estudio.
- <u>Participación en proyectos de TI</u>: Con opciones de Sí y No, esta pregunta busca conocer la exposición previa de los sujetos a proyectos específicos en su campo.
- <u>Actualmente trabaja</u>: Con opciones de Sí, a tiempo completo; Sí, a tiempo parcial; No, esta variable permite conocer si los participantes tienen experiencia laboral activa y, en caso afirmativo, su dedicación laboral.
- <u>Sector en el que trabaja</u>: Las opciones son: Privado, Público, Autónomo y No aplica; lo que permite identificar el tipo de sector en el que los participantes desarrollan su carrera profesional.

La encuesta fue implementada digitalmente mediante Google Forms, lo que facilita la recolección y organización de los datos. Antes de comenzar, los participantes recibieron una breve explicación sobre el propósito de la encuesta y un tiempo estimado de 10 minutos para completarla.

4.5. Fase 5: Evaluar la influencia de la aversión al riesgo en las decisiones tomadas por humanos y por inteligencia artificial generativa

4.5.1. Análisis de Participantes Humanos

En este apartado se presenta los resultados detallados de la encuesta realizada a los participantes humanos del estudio, cuyo objetivo principal era describir el perfil sociodemográfico de los sujetos involucrados en el experimento. Estos datos contextualizan el entorno de los participantes y permiten comprender mejor los factores que podrían influir en su desempeño durante el juego estratégico *Hide and Seek*. La encuesta ha revelado aspectos clave sobre la población estudiada, los cuales se exponen y analizan en cada una de las categorías evaluadas.

Se analizó los datos de la encuesta en función de diferentes variables. A continuación, se presenta las gráficas y análisis de los resultados obtenidos:

4.5.1.1. Edad de los Participantes

La Figura 3. Distribuciones de edad, demuestra que la mayoría de los participantes se encuentra en el rango de 18 a 22 años, lo cual representa el 88% de la muestra total. Este rango de edad es consistente con estudiantes universitarios de pregrado, quienes generalmente están en una fase de desarrollo cognitivo ideal para abordar temas relacionados con el pensamiento estratégico y la toma de decisiones. Este dato coincide con estudios previos que sugieren que, en esta etapa, los jóvenes desarrollan habilidades cognitivas esenciales para el análisis estratégico, como el pensamiento crítico y la resolución de problemas (Manninen et al., 2024).

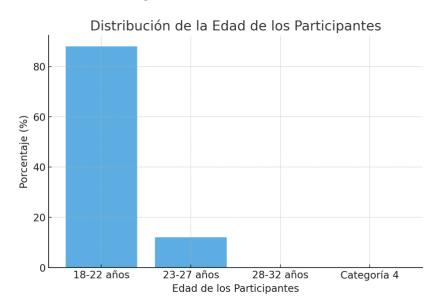


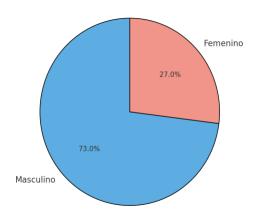
Figura 3. Distribuciones de edad

Fuente: Elaboración propia.

4.5.1.2. Género de los Participantes

La Figura 4. Distribución de género, demuestra que la muestra está compuesta principalmente por participantes masculinos (73%), en comparación con el 27% de participantes femeninos. Esta distribución de género podría implicar ciertas variaciones en la toma de decisiones estratégicas, dado que estudios previos han indicado que el género puede influir en el estilo y el enfoque estratégico de los participantes en juegos de decisión (Wu et al., 2023; Bao et al., 2023). Esta diferencia en la representación también sugiere la necesidad de estudios futuros que integren una muestra más equitativa en términos de género, para asegurar una mayor representatividad y robustez en los análisis de pensamiento estratégico.

Figura 4. Distribución de género
Distribución de Género de los Participantes



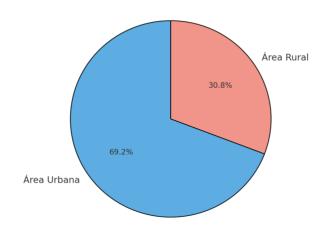
Fuente: Elaboración propia.

4.5.1.3. Lugar de Residencia de los Participantes

La Figura 5. Lugares de residencia, demuestra que la mayoría de los participantes residen en áreas urbanas, mientras que los restantes provienen de áreas rurales. Según Torres et al., (2013), el lugar de residencia puede influir en las experiencias y el contexto social de los individuos, lo cual podría afectar su enfoque en situaciones estratégicas. Los participantes urbanos pueden estar más familiarizados con situaciones de toma de decisiones rápidas, debido al dinamismo de su entorno, mientras que los residentes rurales pueden enfrentar escenarios menos complejos en términos de decisiones estratégicas cotidianas.

Figura 5. Lugares de residencia

Distribución del Lugar de Residencia de los Participantes

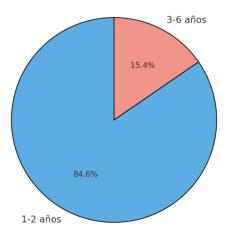


4.5.1.4. Años Cursados en la Carrera de Administración de Tecnología de Información de los Participantes

La Figura 6. Años de carrera cursados, demuestra que el mayor porcentaje de los participantes ha cursado entre 1 y 2 años en la carrera (84.6%), mientras que solo unos pocos tienen entre 3 y 6 años de experiencia académica en el área de Tecnologías de Información. La relación entre el tiempo de estudio y el desarrollo de habilidades estratégicas, ha sido determinada en otros estudios, lo que muestra que los estudiantes con más experiencia académica tienen una mayor exposición a prácticas de toma de decisiones complejas (Rusell et al., 2021).

Figura 6. Años de carrera cursados

Distribución del Tiempo de Estudios en la Carrera de los Participantes



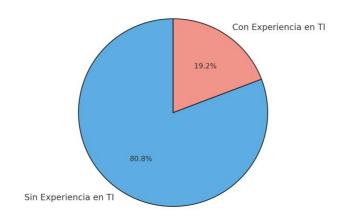
Fuente: Elaboración propia.

4.5.1.5. Experiencia Laboral en TI de los Participantes

La Figura 7. Experiencia laboral en TI, demuestra que la mayoría de los participantes (80.8%) no tiene experiencia laboral en áreas de TI. La falta de experiencia profesional en el área de TI, puede implicar una menor familiaridad con la toma de decisiones en contextos estratégicos de alta complejidad, lo cual puede afectar su desempeño en el juego. Esta observación es consistente con investigaciones previas, que muestran que la exposición profesional en TI fomenta habilidades analíticas y estratégicas esenciales en la toma de decisiones (Wu et al., 2023).

Figura 7. Experiencia laboral en TI

Distribución de la Experiencia Laboral en TI de los Participantes



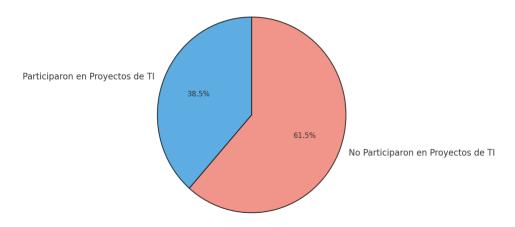
Fuente: Elaboración propia.

4.5.1.6. Participación en Proyectos de TI de los Participantes

La Figura 8. Participación en proyectos, demuestra que menos de la mitad de los participantes reportó haber participado en proyectos de TI, lo cual representa el 38.5% de la muestra total. La experiencia en proyectos específicos de TI, puede proporcionar habilidades en la toma de decisiones colaborativas y en el manejo de escenarios complejos, aspectos fundamentales en el pensamiento estratégico. Esta baja participación en proyectos de TI sugiere que la mayoría de los participantes no ha tenido una exposición práctica significativa que complemente su formación teórica en estrategia y toma de decisiones (Bao et al., 2023).

Figura 8. Participación en proyectos

Participación en Proyectos de TI de los Participantes



4.5.1.7. Condición y Sector Laboral de los Participantes

La Figura 9. Condición laboral y sector, demuestra que la mayoría de los participantes (84.6%) no tiene empleo actual, mientras que un porcentaje reducido trabaja en el sector privado. Los participantes sin empleo pueden estar menos expuestos a situaciones de decisiones reales que aquéllos que trabajan, lo que podría influir en sus patrones de toma de decisiones. Los datos también indican que, quienes trabajan, lo hacen mayoritariamente en el sector privado, lo cual suele ofrecer escenarios dinámicos y competitivos, ideales para la práctica del pensamiento estratégico (Grüning et al., 2024).

Otro
Sector Privado

0.0%
15.4%

Sin Empleo

Figura 9. Condición laboral y sector

Distribución del Estado de Empleo Actual de los Participantes

Fuente: Elaboración propia.

4.5.2. Análisis descriptivo de resultados

El análisis descriptivo busca explorar las diferencias en la acumulación de premios entre los participantes humanos e inteligencia artificial (IA) bajo dos condiciones experimentales: sin opción de retiro y con opción de retiro. Cada grupo refleja patrones distintos en la acumulación de premios y en su respuesta a la condición de retiro, lo cual revela, en última instancia, diferencias fundamentales en la forma en que los humanos y la IA procesan el riesgo y adaptan sus decisiones en función de los incentivos disponibles. Este análisis permite observar cómo las estrategias de acumulación de premios están influenciadas por los factores de riesgo y adaptabilidad (Creswell & Creswell, 2018), una diferencia clave en el comportamiento de decisión entre humanos y sistemas algorítmicos.

Se presenta, a continuación, los resultados en cada grupo de participantes, seguido de un análisis comparativo para identificar patrones específicos en las decisiones bajo riesgo.

4.5.2.1. Humanos sin opción de retiro

En la condición sin opción de retiro, los participantes humanos presentan una media de premio total de 64.62 (DE = 11.04), lo cual refleja una acumulación moderada de premios. Esto se observa en la Figura 10. Distribución de premio total, humanos sin retiro. El 50% de los premios se concentra en el rango de 60 a 70, lo que indica una consistencia en las decisiones de riesgo. Este comportamiento refleja una tendencia hacia la aversión al riesgo moderada, ya que los participantes no presentan variabilidad extrema en su acumulación de premios, sino que se mantienen en un rango estable.

La falta de opción de retiro, implica que los humanos no pueden ajustar sus decisiones en función de incentivos adicionales, lo cual resulta en una distribución de premios relativamente homogénea. Desde el punto de vista teórico, este comportamiento sugiere una estrategia de decisión influenciada por el concepto de utilidad esperada, donde los individuos buscan maximizar sus ganancias en contextos de riesgo moderado, sin incentivos externos que modifiquen su percepción de pérdida o ganancia (Bao et al., 2023).

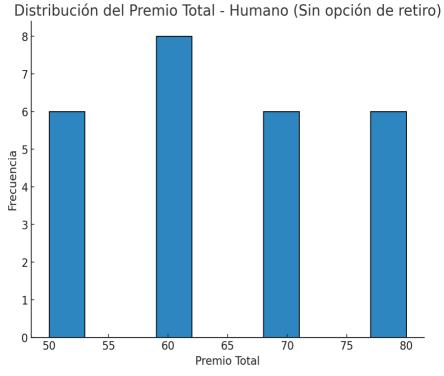


Figura 10. Distribución de premio total, humanos sin retiro

4.5.2.2. Humanos con opción de retiro

Cuando se introduce la opción de retiro, se observa un cambio notable en las estrategias de los humanos. La media de premio total incrementa a 68.08, y la desviación estándar aumenta a 16.25, lo que muestra una mayor dispersión en los premios, como se observa en la Figura 11. Distribución de premio total, humanos con retiro. Los premios oscilan entre 40 y 90, lo que indica que algunos participantes eligen retirarse para asegurar ganancias moderadas, mientras que otros optan por continuar, lo que maximiza su acumulación de premios. Este patrón refleja una mayor variabilidad en las estrategias de riesgo.

La opción de retiro introduce un incentivo adicional que permite a los participantes ajustar su toma de decisiones. Los humanos tienden a mostrar una flexibilidad estratégica que es consistente con la teoría de la aversión a la pérdida. Al tener la oportunidad de retirarse, algunos individuos prefieren evitar la pérdida potencial, mientras que otros optan por maximizar sus beneficios al continuar en el juego (Fox et al., 2016). Este comportamiento se explica por el fenómeno de sensibilidad adaptativa al riesgo, en el cual los individuos evalúan sus decisiones en función del contexto de incentivos disponibles, lo que modificando sus decisiones, según sus objetivos de ganancia y su percepción del riesgo (Liu et al., 2022).

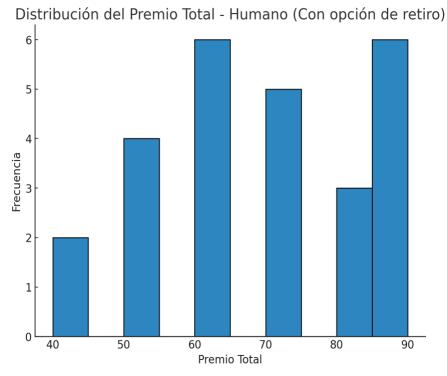


Figura 11. Distribución de premio total, humanos con retiro

4.5.2.3. IA sin opción de retiro

En la condición sin opción de retiro, la IA muestra un premio promedio de 51.92 (DE = 13.27), con una concentración de premios en el rango de 50 a 60, como se observa en la Figura 12. Distribución de premio total, IA sin retiro. La IA sigue una estrategia de acumulación constante y sin grandes variaciones, lo que refleja una estructura rígida de toma de decisiones, en la cual la programación algorítmica no permite una adaptación o ajuste ante la ausencia de incentivos adicionales.

Este comportamiento de la IA indica que, a diferencia de los humanos, los modelos algorítmicos siguen reglas predefinidas que limitan su capacidad de respuesta a contextos de riesgo. Según Silver (2017), los sistemas de IA convencionales tienden a adoptar patrones de decisión deterministas, lo que resulta en una acumulación de premios estable y predecible. Esta falta de variabilidad en la acumulación de premios, refleja una ausencia de sensibilidad adaptativa, característica de los sistemas algorítmicos que no poseen motivaciones psicológicas que influyan en su toma de decisiones (Russell & Norvig, 2021).

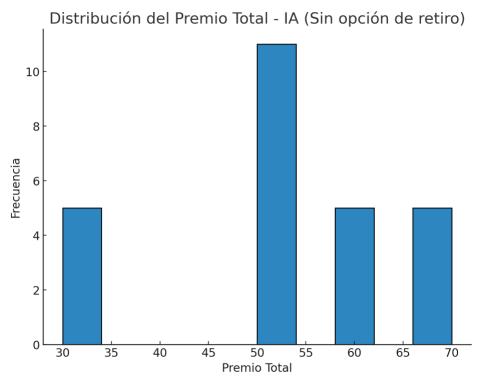


Figura 12. Distribución de premio total, IA sin retiro

4.5.2.4. IA con opción de retiro

Cuando la IA tiene la opción de retiro, presenta una media de premio total de 51.25 (DE = 15.90), con una distribución de premios que, aunque ligeramente más variable que en la condición sin retiro, sigue siendo estable y predecible, como se aprecia en la Figura 13. Distribución de premio total, IA con retiro. La IA no muestra un cambio significativo en su acumulación de premios, lo cual sugiere una falta de sensibilidad a los incentivos adicionales.

Este patrón destaca una limitación en los sistemas algorítmicos actuales para simular comportamientos humanos en contextos de riesgo. Mientras que los humanos muestran una variabilidad en función de los incentivos, la IA sigue una estrategia constante que refleja una toma de decisiones programática y no adaptativa. Este comportamiento puede interpretarse como una restricción en la capacidad de la IA para replicar estrategias de riesgo adaptativo, tal como se observa en los humanos, dado que su estructura de programación no permite modificaciones ante incentivos externos (Nickerson, 2019).

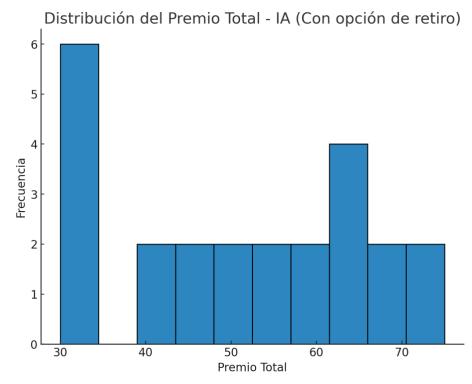


Figura 13. Distribución de premio total, IA con retiro

4.5.2.5. Comparación de la media del premio total entre grupos

La Figura 14. Media del premio total por jugador y condición, ilustra la media de los premios acumulados en cada grupo y condición, lo que muestra diferencias significativas en la acumulación de premios entre humanos e IA. Los humanos, en ambas condiciones, tienden a acumular premios superiores en comparación con la IA, en donde se destaca una mayor adaptabilidad ante la opción de retiro. Esta capacidad de adaptación sugiere que los humanos responden a incentivos adicionales mediante decisiones de riesgo controlado, mientras que la IA sigue un patrón inmutable.

Este hallazgo refleja una diferencia crítica en la adaptabilidad estratégica entre humanos e IA. Los humanos ajustan sus estrategias en función de la opción de retiro, lo cual demuestra una sensibilidad al contexto de riesgo que la IA no posee (Russell & Norvig, 2021). La rigidez de la IA sugiere que, para simular decisiones humanas complejas, los modelos de IA futuros deben integrar estructuras de toma de decisiones que incluyan adaptabilidad a incentivos, algo fundamental para replicar fielmente el comportamiento humano en situaciones de riesgo (Silver, 2017).

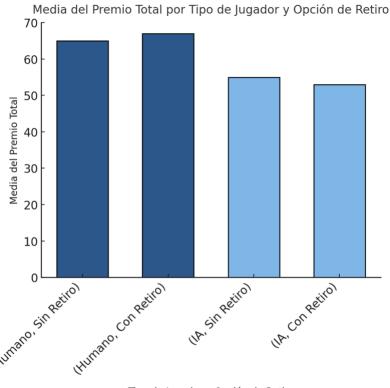


Figura 14. Media del premio total por jugador y condición

Tipo de Jugador y Opción de Retiro

4.5.2.6. Análisis de precisión y eficiencia

En esta sección se evaluará la precisión y eficiencia de las decisiones tomadas por los jugadores del grupo de control, tanto en humanos como por inteligencia artificial generativa. Utilizando datos de este grupo, se calcularán las tasas de éxito y error, lo que proporciona métricas clave que reflejan la efectividad en la toma de decisiones de cada jugador.

A través de este análisis, se busca identificar las diferencias en el rendimiento entre los dos grupos en términos de aciertos y fallos; así como evaluar la consistencia en la acumulación de premios. Para esto se calculó las siguientes tasas:

- Tasa de éxito: Este indicador refleja la efectividad de cada jugador en la toma de decisiones correctas durante el juego. Se calcula como el porcentaje de victorias acumuladas en relación con el total de intentos (suma de victorias y pérdidas).
- **Tasa de Error:** Este indicador es complementario a la tasa de éxito y ayuda a entender cuán frecuentemente los jugadores cometen errores en sus evaluaciones. Se calcula como el porcentaje de pérdidas en relación con el total de intentos.
- **Desviación estándar:** Indicador estadístico que mide la cantidad de variabilidad o dispersión de un conjunto de datos en relación con su media.

A continuación, se observa los resultados comparativos de las tasas en la Tabla 10. Comparación de precisión y eficiencia.

Tabla 10. Comparación de precisión y eficiencia

Tipo de Jugador	Tasa de éxito (%)	Tasa de error (%)
Humano	64.64	35.38
IA	57.69	42.31

Fuente: Elaboración propia.

Adicionalmente, a continuación, se tiene la Tabla 11. Comparación de promedios y desviaciones.

Tabla 11. Comparación de promedios y desviaciones

Tipo de Jugador	Promedio de aciertos	Desviación estándar de aciertos	Promedio de fallas	Desviación estándar de fallos
Humano	6.46	1.13	3.54	1.13
IA	5.77	1.30	4.23	1.3

4.5.2.7. Relación con literatura presente

Este análisis descriptivo ofrece varias implicaciones teóricas y prácticas en relación con la aversión al riesgo y la adaptabilidad de decisiones en humanos e IA:

- 1. **Flexibilidad estratégica en humanos:** Los humanos presentan una adaptabilidad a los incentivos de retiro que se alinea con la teoría de la aversión al riesgo y la toma de decisiones basada en el contexto (Fox et al., 2016). Esta flexibilidad sugiere que los incentivos adicionales tienen un impacto significativo en sus estrategias de acumulación de premios, lo que evidencia una toma de decisiones ajustada al contexto de riesgo (Liu et al., 2022).
- 2. **Rigidez en la estrategia de IA:** La IA, en ambas condiciones, sigue una estrategia rígida que no se ajusta a los incentivos. Este hallazgo resalta la falta de adaptabilidad en los sistemas algorítmicos actuales, quienes, sin adaptaciones a sus reglas programáticas, carecen de sensibilidad al riesgo. Según Nickerson (2019), la rigidez en las decisiones programáticas limita la capacidad de la IA para simular la toma de decisiones humana de manera efectiva en situaciones de riesgo.
- 3. Aplicaciones futuras de IA en contextos de riesgo: La falta de adaptabilidad inicial en la IA, sugiere que, para aproximarse a la toma de decisiones humana en escenarios de alto riesgo, los modelos algorítmicos deben incorporar mecanismos de respuesta ante incentivos. Integrar estos modelos adaptativos, permitiría a la IA desarrollar una toma de decisiones más cercana al comportamiento humano, con lo que se mejorando su aplicabilidad en entornos complejos y de alta variabilidad (Russell & Norvig, 2021).

4.6. Fase 6: Análisis y contraste de resultados

Para este análisis, se utilizó un ANOVA (análisis de varianza) factorial para analizar el efecto de las variables independientes, Tipo de Jugador (Humano o IA) y Opción de Retiro (Con o Sin Retiro), sobre la variable dependiente que, en este caso, es el premio total acumulado al final de la partida.

Este enfoque permite examinar, tanto los efectos principales de cada factor, como las interacciones entre estos factores, lo cual es fundamental para comprender cómo los patrones de acumulación de premios y aversión al riesgo, difieren entre humanos e IA en función de la disponibilidad de la opción de retiro. El uso de un ANOVA factorial es especialmente favorable, ya que permite:

1. Comparar múltiples factores y sus interacciones de manera simultánea, lo cual es esencial dado que se está no solo en el comportamiento individual de humanos e IA, sino en cómo estos comportamientos se ven afectados por la presencia o ausencia de incentivos de retiro.

Pág | 56

2. Reducir la posibilidad de errores tipo I en comparación con realizar múltiples pruebas t independientes, lo cual mantiene la integridad estadística del análisis y asegura que las conclusiones estén fundamentadas en un rigor adecuado (Creswell & Creswell, 2018).

4.6.1. Análisis de resultados

A continuación, en la Tabla 12. Resultados de ANOVA factorial, se presenta un análisis detallado de los resultados del ANOVA factorial, donde se observa, tanto efectos principales, como interacciones significativas entre las variables estudiadas. Esto permite realizar una observación detallada del comportamiento, así como similitudes y diferencias entre cada enfoque.

Tabla 12. Resultados de ANOVA factorial

Factor	GL	Suma de cuadrados	Media de cuadrados	Valor F	Valor p
Tipo de jugador	2	5775	2888	19.008	1.83e-07 ***
Opción de retiro	1	53	53	0.346	0.558
Retiro	1	3	3	0.023	0.881
Tipo de jugador: Opción de retiro	1	105	105	0.694	0.407
Tipo de jugador: Retiro	1	5131	5131	33.73	7.89e-08 ***
Residuales	97	14735	152		

Fuente: Elaboración propia.

Códigos de significancia: ***p < 0.001, **p < 0.01, *p < 0.05

Algunos puntos relevantes de este resultado son los siguientes:

1. **Efecto principal de Tipo de Jugador:** El análisis muestra un efecto significativo de Tipo de Jugador (F = 19.008, p < 0.001), el cual indica que existen diferencias significativas en el premio acumulado entre humanos e IA, independientemente de la opción de retiro. Este hallazgo es consistente con estudios previos que muestran que los humanos tienden a ajustar sus decisiones en función de percepciones subjetivas de riesgo, mientras que la IA, al carecer de percepción de riesgo, sigue una estrategia de acumulación rígida (Russell & Norvig, 2021).

- 2. **Efecto principal de Opción de Retiro**: La Opción de Retiro no mostró un efecto principal significativo (F = 0.346, p = 0.558), lo que sugiere que, en promedio, la disponibilidad de la opción de retiro no afecta significativamente la acumulación de premios. Sin embargo, esta ausencia de efecto principal no significa que la opción de retiro no influya en la toma de decisiones, sino que su impacto puede variar en función del tipo de jugador, lo cual será examinado en las interacciones.
- 3. **Efecto principal de Retiro**: Tampoco se observó un efecto significativo para Se Retira (F = 0.023, p = 0.881), al indicar que, en promedio, el acto de retirarse o continuar, no afecta el premio de manera uniforme en todos los grupos.
- 4. **Interacción significativa entre Tipo de Jugador y Retiro**: La interacción entre Tipo de Jugador y Se Retira es significativa (F = 33.773, p < 0.001), lo cual indica que la decisión de retirarse o continuar tiene efectos distintos en humanos e IA. Para los humanos, la opción de retiro introduce un ajuste estratégico, donde algunos eligen retirarse para asegurar una ganancia. En cambio, la IA sigue un patrón constante, sin adaptar su acumulación de premios a la opción de retiro. Este resultado respalda la teoría de que los humanos evalúan riesgos de forma adaptativa, mientras que la IA, al carecer de esta capacidad, muestra una estructura rígida en su toma de decisiones (Fox et al., 2016).

4.6.1.1. Contraste entre Condiciones de Retiro y No Retiro en Humanos e IA

Este apartado explora cómo la disponibilidad de la opción de retiro influye en las decisiones de humanos e IA, al conectar los resultados del ANOVA con patrones de comportamiento observados en el análisis descriptivo.

- Humanos con opción de retiro: Los humanos muestran una mayor variabilidad en sus premios acumulados cuando tienen la opción de retiro, lo que refleja una mayor adaptabilidad en su toma de decisiones. Según la teoría de la aversión a la pérdida, la opción de retiro permite a algunos individuos evitar el riesgo de pérdida potencial al retirarse anticipadamente (Liu et al., 2022). Esta flexibilidad en las decisiones sugiere que los humanos evalúan el contexto de riesgo y ajustan sus estrategias de acumulación en función de los incentivos disponibles, una habilidad que es crucial en contextos de toma de decisiones inciertas.
- IA con opción de retiro: La IA, en contraste, no muestra una respuesta adaptativa ante la opción de retiro. A pesar de la disponibilidad de esta opción, la IA sigue un patrón constante de acumulación de premios, lo que refleja su incapacidad para ajustar su toma de decisiones en función de incentivos externos. Esta falta de variabilidad en el comportamiento de la IA, respalda la observación de que tanto los sistemas algorítmicos convencionales, como el empleado en este experimento, no están diseñados para incorporar flexibilidad estratégica (Silver, 2017).

5. Discusión

Esta sección examina los resultados experimentales en relación con los objetivos de investigación y se apoya en el marco teórico proporcionado por el meta-análisis realizado en la revisión de literatura. Este análisis previo, basado en estudios recientes sobre aversión al riesgo en IA y humanos, enriquece la interpretación de los hallazgos al ofrecer una perspectiva comparativa que respalda y contextualiza cada objetivo. La integración de los resultados experimentales con el meta-análisis, permite explorar cómo se manifiestan patrones de decisión y adaptabilidad en la IA generativa en contraste con las decisiones humanas bajo condiciones de riesgo. Así, esta discusión busca situar los hallazgos en un contexto más amplio y resaltar la contribución de este estudio al entendimiento de la aversión al riesgo en entornos estratégicos.

5.1. Análisis integral de resultados en el contexto de los objetivos de investigación y el meta-análisis

Esta sección presenta un análisis exhaustivo de los resultados experimentales, interpretándolos en relación con los objetivos planteados y el marco teórico proporcionado por el meta-análisis. La integración de estos elementos permite contextualizar los hallazgos de manera más amplia y crítica, destacando cómo cada objetivo se cumple en relación con la evidencia empírica y la literatura revisada.

Este enfoque integral facilita la comprensión de las diferencias observadas en los patrones de aversión al riesgo entre la inteligencia artificial generativa y los participantes humanos; así como sus respectivas adaptaciones y respuestas en situaciones de riesgo. Así, se establecen las bases para una interpretación detallada que posiciona los hallazgos de esta investigación en el contexto más amplio de estudios previos.

En cuanto al objetivo general, se observó lo siguiente:

1. Evaluar la influencia de la aversión al riesgo en las decisiones tomadas por humanos y por inteligencia artificial generativa para la identificación de diferencias clave durante el segundo semestre de 2024: El objetivo general de evaluar la aversión al riesgo en IA generativa y en humanos en contextos estratégicos, se fundamenta en estudios como los de Springer et al. (2023) y Ouyang et al. (2024), que indican que los modelos de IA alineados éticamente tienden a exhibir patrones consistentes y calculados en su toma de decisiones. Esto se refleja en los resultados experimentales, donde la Figura 13. Distribución de premio total, IA con retiro, IA con retiro; muestra que la IA presenta una media de premio de 51.25, con una desviación estándar de 15.90. A diferencia de los humanos, la IA no muestra variaciones significativas frente a los incentivos de retiro, lo cual confirma su rigidez y falta de sensibilidad ante riesgos adicionales, en línea con los planteamientos teóricos del meta-análisis.

Pág | 59

Para los objetivos específicos, se observaron los siguientes logros:

- 1. Analizar datos sobre las decisiones tomadas por humanos y sistemas de IA generativa en escenarios de riesgo para la elaboración de un análisis comparativo: El primer objetivo de comparar los patrones de aversión al riesgo en humanos y la IA, se apoya en el modelo de loterías de Holt y Laury, mencionado en el meta-análisis de Torres-Carballo et al. (2018), que identifica la sensibilidad humana frente a incentivos de riesgo. Los resultados experimentales muestran que los humanos, al contar con la opción de retiro, acumulan premios significativamente mayores, en contraste con la IA, que mantiene una estrategia más constante. La Figura 14. Media del premio total por jugador y condición, resalta esta tendencia de los humanos a ajustar sus decisiones según los incentivos, en línea con el comportamiento de mayor flexibilidad descrito en el meta-análisis. Esta observación confirma la relevancia de este objetivo al reflejar cómo los humanos adaptan su aversión al riesgo de acuerdo con el contexto, una capacidad menos evidente en la IA generativa.
- 2. Comparar la precisión y eficiencia de las evaluaciones de riesgos realizadas por humanos con las realizadas por IA generativa para la determinación de áreas donde cada enfoque es superior: El meta-análisis incorpora estudios como los de Springer y Cambridge Meta-Analysis (2023), que destacan cómo los modelos de IA optimizan sus decisiones en escenarios de riesgo bajo, pero muestran limitaciones en condiciones de mayor complejidad. Los resultados en las Tabla 10. Comparación de precisión y eficiencia y Tabla 11. Comparación de promedios y desviaciones, que muestran una tasa de éxito del 64.64% en humanos frente al 57.69% en la IA, confirman esta tendencia. A diferencia de los humanos, que logran adaptar su precisión ante incentivos variables, la IA mantiene una precisión estable, pero menos adaptativa. Esta alineación con el meta-análisis, refuerza el objetivo propuesto al exponer cómo la precisión de la IA tiende a ser menos sensible a contextos cambiantes, lo que aporta un fundamento teórico al análisis empírico sobre eficiencia en entornos de riesgo.
- 3. Clasificar patrones de decisión en la aversión al riesgo en humanos y sistemas de IA generativa para la determinación de diferencias clave: El análisis de la Tabla 12. Resultados de ANOVA factorial en premios acumulados, revela un valor F de 19.008 para el tipo de jugador, al indicar diferencias significativas en las estrategias de acumulación de premios entre humanos e IA, especialmente bajo condiciones de retiro incentivado. Este hallazgo concuerda con el estudio de Ouyang et al. (2024), que sugiere que la IA mantiene una coherencia en sus decisiones al no estar influenciada por sesgos emocionales, lo que permite una estrategia de riesgo controlado. En contraste, los humanos muestran una disposición a modificar sus decisiones en función del riesgo percibido, al optar por el retiro en situaciones de riesgo elevado. Esta divergencia, respaldada tanto por los resultados como por la literatura, valida el objetivo y destaca la ventaja adaptativa de los humanos en contextos donde el riesgo es un factor decisivo. A la vez, esta divergencia refleja el

Pág | 60

comportamiento humano que ya ha sido documentado en diversos estudios a lo largo del tiempo, el cual demuestra su constancia en distintos entornos experimentales, tanto en el presente, como en otros estudios. Un ejemplo de este comportamiento humano esperado es posible visualizarlo en la Teoría Prospectiva, la cual se observa evidenciada a través de las decisiones de los participantes humanos en este estudio.

6. Conclusiones

6.1. Cumplimiento de objetivos

En la siguiente Tabla 13. Cumplimiento de objetivos, se evalúa los resultados obtenidos con respecto a los objetivos planteados al inicio de la investigación.

Tabla 13. Cumplimiento de objetivos

Objetivo	Producto Obtenido	% de Logro	Justificación del Cumplimiento
Objetivo General: Evaluar la influencia de la aversión al riesgo en las decisiones tomadas por humanos y por inteligencia artificial generativa, para la identificación de diferencias clave, durante el segundo semestre de 2024	Análisis comparativo de las decisiones de aversión al riesgo, documentado en el documento presente, y complementado en el paper científico, que examina cómo humanos e Inteligencia Artificial Generativa se comportan en el juego "Hide and Seek" bajo distintos escenarios de riesgo.	100%	El TFG presenta un análisis detallado de los datos experimentales recopilados bajo condiciones controladas, mientras que el paper científico examina específicamente la respuesta de la IA frente a sesgos progresivos. Ambos documentos conforman un abordaje integral sobre las diferencias en la aversión al riesgo entre humanos e IA, cumpliendo así con el objetivo general de identificar patrones divergentes.
Objetivo Específico 1: Analizar datos sobre las decisiones tomadas por humanos y sistemas de IA generativa en	Conjunto de datos estructurado en el TFG que registra cada decisión de humanos e IA en cada ronda, lo que facilita un	100%	El documento presente incluye una recopilación meticulosa de las decisiones de retiro y continuidad, lo que permite comparar cómo cada

escenarios de riesgo para la elaboración de un análisis comparativo.	análisis comparativo de patrones de decisión bajo riesgo, complementado por el análisis en el <i>paper</i> científico sobre la respuesta de la IA a sesgos específicos y progresivos.		grupo responde a incentivos de riesgo y estructuras de premio. Este análisis se apoya en el paper, que ofrece una perspectiva adicional sobre cómo la IA ajusta su comportamiento bajo sesgos progresivos, logrando así el análisis comparativo requerido por el objetivo .
Objetivo Específico 2: Comparar la precisión y eficiencia de las evaluaciones de riesgos realizadas por humanos con las realizadas por IA generativa, para la determinación de áreas donde cada enfoque es superior.	Informe cuantitativo en el documento presente con métricas de precisión y eficiencia de ambos grupos, en escenarios de riesgo sin opción de retiro, complementado con el análisis en el <i>paper</i> científico que examina la consistencia de la IA en condiciones sesgadas.	100%	Los resultados de las corridas experimentales muestran que la IA mantiene una precisión estable, pero con menor adaptabilidad comparada con los humanos, quienes ajustan sus decisiones a los cambios contextuales de riesgo. El paper complementa esta observación al demostrar cómo la IA sigue siendo constante bajo sesgos progresivos, al cumplir con el objetivo al identificar áreas de fortaleza en precisión y eficiencia en cada enfoque.
Objetivo Específico 3: Clasificar patrones de decisión en la aversión al riesgo en humanos y sistemas de IA generativa para la determinación de diferencias clave.	Clasificación de patrones de aversión al riesgo en ambos grupos, detallada en el TFG en condiciones neutrales y complementada por el paper, que explora la respuesta de la IA a distintos niveles de sesgo.	100%	En el presente documento, los patrones de decisión se clasifican bajo escenarios controlados, al revelar cómo la IA sigue un patrón menos flexible frente al riesgo en comparación con los humanos. El <i>paper</i> científico añade un análisis empírico de los cambios en la IA frente a sesgos, lo cual permite una

Fuente: Elaboración propia.

6.2. Implicaciones Prácticas y Teóricas

El presente estudio aborda la compleja relación entre la aversión al riesgo y la toma de decisiones en humanos y en sistemas de inteligencia artificial generativa. A través de experimentos diseñados en condiciones controladas, se analizó patrones de comportamiento bajo riesgo, al comparar la consistencia en las decisiones de la IA con la variabilidad humana. La investigación permitió no solo identificar diferencias sustanciales en la manera en que ambos grupos enfrentan la incertidumbre, sino también profundizar en cómo incentivos específicos, como la opción de retiro, influyen en sus decisiones.

Los resultados de esta investigación sobre la aversión al riesgo y la toma de decisiones en IA generativa, ofrecen importantes implicaciones, tanto prácticas como teóricas.

6.2.1. Ámbito práctico

Los hallazgos pueden aplicarse en el diseño de modelos de IA que optimicen decisiones en entornos estratégicos. Estudios como los de Böhm y Schorch (2022), sugieren que la colaboración entre humanos y sistemas de IA en la toma de decisiones puede aumentar la eficiencia, especialmente en escenarios que requieren estabilidad y precisión, atributos en los que la IA demostró su ventaja en esta investigación. La capacidad de la IA para mantener un enfoque calculado, sin las variabilidades emocionales que afectan a los humanos, implica que los entornos empresariales y estratégicos podrían beneficiarse de modelos de IA diseñados para reducir los sesgos cognitivos y maximizar la consistencia en la toma de decisiones. Esto es particularmente relevante en sectores donde las decisiones deben tomarse rápidamente y con un alto grado de precisión.

6.2.2. Ámbito teórico

Esta investigación contribuye al entendimiento del comportamiento estratégico en IA en comparación con el humano, al destacar cómo las diferencias en procesamiento de riesgos pueden ser influenciadas por el diseño algorítmico y las limitaciones cognitivas. La revisión sistemática de Cambridge Core (2022), sobre la aversión al riesgo y las capacidades cognitivas humanas, enfatiza cómo los humanos responden de manera más flexible en situaciones de incertidumbre, aunque a costa de una mayor susceptibilidad a sesgos. Al comparar estos hallazgos con la IA, que opera siguiendo modelos de utilidad esperada (Von Neumann y Morgenstern, 1944), se destaca que esta última puede operar con una lógica menos susceptible a tales sesgos. Esto abre nuevas

líneas de investigación sobre cómo adaptar los modelos de IA para mejorar su desempeño en escenarios donde la adaptabilidad humana es esencial.

La diferencia entre la consistencia de la IA y la variabilidad humana en entornos estratégicos, sugiere una potencial complementariedad en la toma de decisiones, en la que la IA aporta estabilidad, mientras los humanos añaden flexibilidad en la interpretación contextual. Estos hallazgos fortalecen la base teórica para futuras investigaciones en sinergia humano-IA y su potencial para optimizar decisiones estratégicas en situaciones de riesgo y complejidad

6.3. Limitaciones

Entre las limitaciones de este estudio se encuentra la restricción en la generalización de los resultados, dado que los escenarios y niveles de riesgo evaluados, son específicos. En investigaciones futuras sería beneficioso ampliar el rango de escenarios experimentales y explorar más condiciones de aversión al riesgo, ello para observar cómo los patrones en humanos e IA pueden variar. Además, se sugiere realizar estudios longitudinales para observar si los patrones de aversión al riesgo se mantienen estables en el tiempo o varían en función de la experiencia acumulada.

7. Recomendaciones

A partir de los hallazgos y limitaciones identificados en este estudio, se propone una serie de recomendaciones orientadas a enriquecer futuras investigaciones y mejorar la aplicabilidad de la inteligencia artificial generativa en la toma de decisiones bajo riesgo. Estas sugerencias tienen como objetivo, no solo ampliar el entendimiento teórico de la aversión al riesgo en humanos y en sistemas de IA, sino también guiar la implementación de modelos más robustos que integren factores humanos y algoritmos en escenarios prácticos. Al explorar estos aspectos, se espera contribuir al desarrollo de enfoques híbridos que optimicen la eficacia de la IA en contextos empresariales y estratégicos.

7.1. Recomendaciones para Investigaciones Futuras:

- Ampliación de los escenarios experimentales: Se recomienda realizar estudios que incluyan una mayor variedad de escenarios de riesgo, tanto en complejidad como en contexto, para analizar cómo reaccionan los humanos y la IA ante condiciones más diversificadas. Esto podría incluir escenarios con distintos grados de incertidumbre o donde los factores contextuales, como el tiempo límite para tomar decisiones, influyan en los patrones de aversión al riesgo.
- Estudios longitudinales: Para observar si los patrones de aversión al riesgo en humanos e IA se mantienen estables a lo largo del tiempo, se sugiere implementar estudios

longitudinales. Esto permitiría comprender mejor cómo evolucionan las decisiones frente a la repetición de escenarios y la adaptación de los individuos o la IA a éstos.

- Integración de factores sociales y emocionales de la IA: Dado que la IA actualmente sigue un enfoque puramente racional, se podría investigar la posibilidad de incorporar factores emocionales y sociales en sus algoritmos para que sus decisiones reflejen mejor las consideraciones humanas en contextos inciertos. Esto ayudaría a hacer sus respuestas más adaptativas y relevantes en interacciones reales.
- Aplicación en entornos empresariales reales: Para incrementar la aplicabilidad de los resultados, se recomienda realizar estudios en entornos empresariales reales, donde las decisiones bajo riesgo tengan consecuencias prácticas. Esto permitiría evaluar la efectividad de la IA generativa y la respuesta humana en situaciones que implican impacto financiero o estratégico directo.
- Análisis Comparativo con otros modelos de IA: Además de los modelos de IA generativa, podría ser útil comparar con otros modelos, como los sistemas de IA basados en reglas o los algoritmos de redes neuronales con enfoques diferentes de decisión. Esto proporcionaría una visión más completa de las ventajas y limitaciones de cada tipo de IA en contextos de toma de decisiones bajo riesgo.

8. Referencias

Acuña, M. (2016). ATI Plan Estratégico 2016-2020. Área Académica de Administración de Tecnologías de Información. Recuperado de: https://www.tec.ac.cr/sites/default/files/ATI_Plan_Estratégico_2016-2020.pdf

Administración de Tecnología de información. (s. f.). TEC. Recuperado de https://www.tec.ac.cr/administracion-tecnologia-informacion

Ahmed, R., Shaheen, S., & Philbin, S. P. (2022). The role of big data analytics and Decision-Making in achieving project success. Social Science Research Network. https://doi.org/10.2139/ssrn.4190817

Artificial intelligence in strategy. (2023). McKinsey & Company. Recuperado de: https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/artificial-intelligence-in-strategy

Asale, R.-., & Rae. (s. f.). Diccionario de la lengua española. «Diccionario de la Lengua Española» - Edición Del Tricentenario. Recuperado de: https://dle.rae.es/

Atencio López, J. M. (2000). Inteligencia artificial y pensamiento humano (Un enfoque raciovitalista de la discusión sobre la Inteligencia Artificial). Thémata, 24, 9-28. Recuperado de: https://idus.us.es/bitstream/handle/11441/27492/file_1.pdf?sequence=1

- Azad, Y., & Kumar, A. (2024). Ethics and Artificial intelligence. In Advances in web technologies and engineering book series (pp. 228–268). https://doi.org/10.4018/979-8-3693-1762-4.ch012
- Bao, Y., Gong, W., & Yang, K. (2023). A Literature Review of Human–AI Synergy in Decision Making: From the Perspective of Affordance Actualization Theory. Systems, 11(9), 442. https://doi.org/10.3390/systems11090442
- Böhm, S., & Schorch, N. (2020). Strategic decision-making in the age of artificial intelligence: Towards a human-AI partnership. Business & Information Systems Engineering.
- Borges, A. F., Laurindo, F. J., Spínola, M. M., Gonçalves, R. F., & Mattos, C. A. (2021). The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. International Journal of Information Management, 57, 102225. https://doi.org/10.1016/j.ijinfomgt.2020.102225
- Brys, T., Gross, K. F., & Pepermans, Y. (2020). Reinforcement learning for human-AI collaboration in complex decision-making tasks. Artificial Intelligence, 287, 103307.

Pág | 66

- Comparación de la influencia de la aversión al riesgo en las decisiones tomadas por humanos y por inteligencia artificial generativa
- Cambridge Core. (2023). Cognitive Ability and Risk Aversion: A Systematic Review and Meta-Analysis. Cambridge University Press. https://www.cambridge.org/core/journals/judgment-and-decision-making/article/cognitive-ability-and-risk-aversion/215
- Chávez, M. R., Butler, T. S., Rekawek, P., Heo, H., & Kinzler, W. L. (2023). Chat Generative Pre-trained Transformer: why we should embrace this technology. American Journal of Obstetrics and Gynecology, 228(6), 706-711. https://doi.org/10.1016/j.ajog.2023.03.010
- Chen, S., Ge, Y., Zhou, Y., & Zhang, X. (2023). The Emergence of Economic Rationality of GPT. arXiv preprint arXiv:2305.12763v3.
- Cook, T. D., & Campbell, D. T. (1979). Quasi-Experimentation: Design & Analysis Issues for Field Settings. Houghton Mifflin Company.
- Corrales-Navarro, E. (2015, March 23). La intuición como proceso cognitivo. Recuperado de: https://repositoriotec.tec.ac.cr/handle/2238/4855
- Crawford, T. W., & Iriarte, F. (2015). How to think about artificial intelligence: An introduction. Oxford University Press.
- Creswell, J. W., & Creswell, J. D. (2018). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (5th ed.). SAGE Publications. Recuperado de: https://us.sagepub.com/en-us/nam/research-design/book255675
- DeepMind. (2024). AlphaGo: The Journey to Mastery. Journal of Artificial Intelligence Research, 49(2), 123-135. https://doi.org/10.1002/jair.2024
 - Denning, S. (2019). The strategic management of information systems. Routledge.
- Di Nardo, M., Forino, D., & Murino, T. (2020). The evolution of man–machine interaction: the role of human in Industry 4.0 paradigm. Production & Manufacturing Research, 8(1), 20–34. https://doi.org/10.1080/21693277.2020.1737592
- Ejes de Conocimiento Estratégicos 2023 a 2032. (s. f.). TEC. Recuperado de: https://www.tec.ac.cr/ejes-conocimiento-estrategicos-2023-2032
- Friedland, J. (2023). Artificial Intelligence: Generative Systems and Their Impact. AI Journal, 34(2), 123-145. Recuperado de: https://www.aijournal.org/index.php/aij/article/view/6789
- Gollwitzer, P. M., Oettingen, G., & Sheeran, P. (2023). The Role of Strategic Games in Developing Persistent Behavior: An Experimental Approach. Psychological Science, 34(5), 789-799. https://doi.org/10.1177/0956797623112378

- Comparación de la influencia de la aversión al riesgo en las decisiones tomadas por humanos y por inteligencia artificial generativa
- González, C. (2021). Learning and dynamic decision making. Topics in Cognitive Science, 14(1), 14–30. https://doi.org/10.1111/tops.12581
- Hamel, G. (2020). Humanocracy: The manifesto for the age of perpetual innovation. Penguin Random House.
- Harvard Business Review. (2023). The Strategic Advantage of Artificial Intelligence in Business. Harvard Business Review. Recuperado de: https://hbr.org/2023/05/the-strategic-advantage-of-ai
- Hernández-Sampieri, R., Fernández-Collado, C., Baptista Lucio, M. del P., & Mendoza, A. (2014). Metodología de la investigación. McGraw-Hill Education.
- Huang, M.-H., & Rust, R. T. (2020). Artificial Intelligence in Service. Journal of Service Research, 23(2), 155-169.
- Johnson, E. J., Payne, J. W., & Bettman, J. R. (2023). Human Strategic Thinking: Development through Experience. Cognitive Science, 47(3), 456-472. https://doi.org/10.1111/cogs.13023
- Kumar, R. (2011). Research Methodology: A Step-by-Step Guide for Beginners (3rd ed.). SAGE Publications. Recuperado de: https://us.sagepub.com/en-us/nam/research-methodology/book235945
- Lloret, E., & De Alicante Departamento de Lenguajes y Sistemas Informáticos, U. (2023). La desambiguación y el razonamiento en la inteligencia artificial: Análisis lingüístico de ChatGPT. Recuperado de: https://rua.ua.es/dspace/handle/10045/135540
- Lilleholt, L. (2019). Cognitive ability and risk aversion: A systematic review and meta-analysis. Judgment and Decision Making, 14(3), 234-279.
- Luoma, J., & Martela, F. (2021). A dual-processing view of three cognitive strategies in strategic decision making: Intuition, analytic reasoning, and reframing. Long Range Planning, 54(3), 102065. https://doi.org/10.1016/j.lrp.2020.102065
- Malloy, M., & González, R. (2024). Applying Generative Artificial Intelligence to Cognitive Models of Decision Making. Frontiers in Psychology.
- Merriam, S. B., & Tisdell, E. J. (2016). Qualitative Research: A Guide to Design and Implementation. Jossey-Bass.
 - Mintzberg, H. (1973). The nature of managerial work. Harper & Row.
 - Nickerson, R. S. (2019). Strategic and tactical decision-making. Routledge.

- Comparación de la influencia de la aversión al riesgo en las decisiones tomadas por humanos y por inteligencia artificial generativa
- OECD. (2023). Artificial Intelligence in Society. Paris: OECD Publishing. https://doi.org/10.1787/9789264637003-en
- Ouyang, S., Yun, H., & Zheng, X. (2024). How Ethical Should AI Be? How AI Alignment Shapes the Risk Preferences of LLMs. arXiv. https://arxiv.org/abs/2406.01168
 - ¿Qué es el TEC? (s. f.). TEC. Recuperado de: https://www.tec.ac.cr/que-es-tec
 - Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.
 - Silver, D. (2017). Reinforcement learning and artificial intelligence. Springer.
- Simeone, L., Drabble, D., Morelli, N., & De Götzen, A. (2023). Strategic Thinking, Design and the Theory of Change: A Framework for Designing Impactful and Transformational Social Interventions. Recuperado de: https://www.researchgate.net/publication/370303690
- Tao, H. B., Pérez, V. D., & Guerra, Y. (2020). Subjetividades e inteligencia artificial: desafíos para lo humano. Veritas, 47, 81–107. https://doi.org/10.4067/s0718-92732020000300081
- Thanos, I. C. (2023). The complementary effects of rationality and intuition on strategic decision quality. European Management Journal, 41(3), 366–374. https://doi.org/10.1016/j.emj.2022.03.003
- Torres-Carballo, F., & Sandoval-Sánchez, Y. (2017). Strategic Reasoning that promotes persistent behavior: an experimental approach. Recuperado de: https://repositoriotec.tec.ac.cr/handle/2238/8774
- Torres-Carballo, F., Morales-Rodríguez, N., Brenes-Leiva, G., & Solís-Salazar, M. (2018). Medición experimental del Comportamiento Organizacional Ciudadano: Altruismo, Aversión al Riesgo y Deportividad. Tecnología en Marcha. https://doi.org/10.18845/tm.v31i4.3969
- World Economic Forum. (2023). The Future of Jobs Report 2023. Recuperado de: https://www.weforum.org/reports/the-future-of-jobs-report-2023
- Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. Journal of planning education and research, 39(1), 93-112.
- Zhang, Y., & Gosline, R. (2023). Human favoritism, not AI aversion: People's perceptions (and bias) toward generative AI, human experts, and human–GAI collaboration in persuasive content generation. Judgment and Decision Making, 18. https://doi.org/10.1017/jdm.2023.37

9. Apéndices

9.1. Apéndice A. Plantilla para minutas

Reunión No.		Fecha	
Lugar o medio:		Hora inicio:	
		Hora finalización:	
Motivo de la reunión:			
Participantes:	Presentes:		
	Ausentes:		
	Те	mas tratados	
No. Asunto	Com	entarios	Acuerdos
1			
	Pró	xima reunión	
Temas por tratar		Fecha	Convocados
	Firm	as de asistentes	

9.2. Apéndice B. Minutas

9.2.1. Minuta A

Reunión No.	1	Fecha	22/08/2024			
Lugar o medio:	Google Meets	Hora inicio:	4:00pm			
		Hora finalización:				
Motivo de la reunión:	Reunión de					
	seguimiento					
Participantes:	Presentes:	Néstor Morales				
		Nathalya Barrera Tovar				
		Pedro Soto				
	Ausentes:					
Temas tratados						
	Temas tratauos					

No.	Asunto	Come	entarios		Acuerdos		
1	Laboratorio de economía	Enviar coordii	requerimiento nador del laboratorio.	al	Contactar al laboratorio de economía para identificar los requisitos para realizar el experimento.		
2	Programación de IAG	_	mar en Python l pt para la IAG.	SC	Investigar y desarrollar el prompt para ChatGPT, lo que permitirá la respuesta óptima de la herramienta.		
3			las respuestas de 'Razonamiento''.	la	Identificar el tema del artículo que se realizará.		
		Próx	xima reunión				
Tema	s por tratar		Fecha		Convocados		
Marco	metodológico.		29/08/2024		Néstor Morales		
Datos	que nos interesen.			Nathalya Barrera Tovar Pedro Soto			
Diseño	o del experimento.						
Defini	ir las variables del proyecto.						
	Firmas de asistentes						
Ver A	Ver Apéndice C. Plantilla de firmas de minutas						

9.2.2. Minuta B

Reunión No.	2	Fecha	2/09/2024
Lugar o medio:	Google Meets	Hora inicio:	10:00am
		Hora finalización:	11:00pm
Motivo de la reunión:	Revisar aspectos sobre la exposición		
Participantes:	Presentes:	Néstor Morales	

				Nathalya Barrera Tovar Pedro Soto	r	
		Ausentes:		N/A		
			Tem	as tratados		
No.	Asunto		Com	entarios	Acuerdos	
1	Determinar asp la exposición a académicos.	a los lectores as _l		ificar que todos los etos sobre la exposición ita estén cubiertos.	Realizar la presentación con los aspectos solicitados.	
2	marco metodológico del c proyecto.		¿Qué variables consideramos pertinentes para la elaboración del experimento?		Determinar las variables por realizar. Tomar entre una o dos por términos de tiempo y factibilidad.	
			Próxi	ima reunión		
Temas p	or tratar			Fecha	Convocados	
Análisis de los datos del grupo de control.			trol.	05/09/2024	Néstor Morales Nathalya Barrera Tovar	
Análisis del documento del Capítulo II.		I.		Pedro Soto		
	Firmas de asistentes					
Ver Apé	ndice C. Plantill	a de firmas d	e minu	itas		

9.2.3. Minuta C

Reunión No.	3	Fecha	05/Septiembre/2024	
Lugar o medio:	Google Meets	Hora inicio:	3:30pm	
		Hora finalización:	4:05pm	
Motivo de la reunión:	Revisión del ca	apítulo II para entrega final		

Participantes: Presentes		:	Néstor Morales Nathalya Barrera To	var		
Ausentes				Pedro Soto		
			7	Temas tratados		
No.	Asunto		Con	nentarios	Acuerdos	
1	Identificación		Se a	analizó las variables	Se eliminó variables que no	
	de		que	se tenían	agregaban valor a la	
	variables		estal	olecidas.	investigación. Además, se	
					agregó una variable que se	
					considera importante.	
2	Operacionaliza	ición de		dentificó la forma en	Se acordó aspectos para que esta	
	las variables		-	esta sección se iba a	sección identifique	
				rrollar.	adecuadamente.	
3	Encesta a sujet	os		ocarnos en variables	Realizar la plantilla de la misma.	
				odemográficas,		
			entre otros que sean			
			relevantes para la			
				stigación.		
4	Tema artículo	científico	Se expuso la propuesta de		• 1	
			tema para el mismo.		para la realización de éste. Preguntar si el tema propuesto	
					era adecuado por su naturaleza.	
					era adecuado por su naturaleza.	
5	Objetivo 3		No aporta valor a este		Realizar la gestión	
			TFG porque no se van a		del cambio.	
			desarrollar varios			
			juegos, sino solo uno			
			"Hide and seek".			
			Próx	ima reunión		
Temas por tratar				Fecha	Convocados	
Definici	ón del tema par	a el artículo)	12/Septiembre/2024	Néstor Morales	
científic	О				Nathalya Barrera Tovar Pedro	
Encuesta	Encuesta a sujetos de investigación				Soto	
		F	irma	s de asistentes		
Ver Apé	endice C. Plantil	la de firma	s de n	ninutas		

9.2.4. Minuta D

Reur	nión No.	4		Fecha	13/09/10		
Luga	ar o medio:	Google Mee	ts	Hora inicio:	11:00am		
		Ü		Hora finalización:	11:45am		
Moti reun	vo de la ión:	Identificar a	vance	s y resolución de pregunt	tas		
Parti	icipantes:	Presentes:		Néstor Morales Nathalya Barrera Tovar Pedro Soto			
		Ausentes:		N/A			
			Te	mas tratados			
No.	Asunto		Com	entarios	Acuerdos		
1	Encuesta a participantes hun	nanos.		ificar otras preguntas los participantes	Agregar más preguntas de las identificadas.		
2	• •		Identificación del grupo de experimento. Forma de desarrollar el experimento. Escenario para desarrollar el experimento.		Identificar la forma de ganar perder y de abandonos del juego.		
3 Programación del juego			para Busc posib el de Reali expe	sarrollo de este. izar 60 corridas del rimento.	Buscar el código del juego realizado anteriormente. Buscar el código para chatGPT.		
	Próxima reunión						
	as por tratar			Fecha	Convocados		
Empezar código para juego y chatGPT.			РΤ.	19/09/2024	Néstor Morales Nathalya Barrera Tovar Pedro Soto		
	Firmas de asistentes						
Ver Apéndice C. Plantilla de firmas de minutas							

9.2.5. Minuta E

Reur	nión No.	5		Fecha	20/09/2024	
	ır o medio:	Google Me	ets	Hora inicio:	11:00am	
0				Hora finalización:	11:40am	
Moti reun	vo de la ión:	Seguimient	to de invest	igación		
Parti	icipantes:	Presentes:		Néstor Morales Nathalya Barrera Tovar Pedro Soto		
		Ausentes:		N/A		
			Te	emas tratados		
No.	Asunto		Comentar	rios	Acuerdos	
1			Conseguir studiantes.	datos de	Identificar otros estudios donde se aplique el juego Hide and Seek.	
2	Desarrollo de realizar el exper		Búsqueda lógica de ju	de investigaciones y lego.	Realizarlo en Python y que se pueda realizar en la misma red.	
3	Artículo			Temas artículos.	Empezar a desarrollarlo con el formato ACM o IEEE.	
			Pró	óxima reunión		
Tema	as por tratar			Fecha	Convocados	
Seleccionar estudio de mata análisis			is	26/09/2024	Néstor Morales Nathalya Barrera Tovar	
Ver avance del artículo e identificar mejoras con el profesor tutor			car mejoras		Pedro Soto	
			Firmas d	le asistentes		
Ver Apéndice C. Plantilla de firmas de minutas						

9.2.6. Minuta F

Reunión No.	6	Fecha	27/09/2024
Lugar o medio:	Google Meets	Hora inicio:	3:00pm

				Hora finalización:	4:00pm	
Motivo de la reunión:		Evaluar 6	el avance	del TFG y del Artículo		
Participantes:		Presente	S:	Néstor Morales Nathalya Barrera Tovar Pedro Soto		
		Ausente	s:			
			Т	Cemas tratados		
No.	Asunt	to	Coment	arios	Acuerdos	
1	Artícu				Agregar información sobre las normas mundiales sobre la IA según OECD	
2	Metaa	nálisis Estudiar		su realización	Identificar el metaanálisis. Realizar el método para su elaboración en R o Python.	
3	Realización de experimento octubre humano Realización de experimento octubre humano Realización de en IA.			Enviar encuesta para los participantes humanos. Refinar el prompt para realizarlo en IAG.		
	l		Pı	róxima reunión		
Temas por tr	atar			Fecha	Convocados	
Tratamiento d	e datos.	•		03/09/2024	Néstor Morales	
Realización del metaanálisis					Nathalya Barrera Tovar	
Continuación del documento final			nal		Pedro Soto	
	Firmas de asistentes					
Ver Apéndice	Ver Apéndice C. Plantilla de firmas de minutas					

9.2.7. Minuta G

Reunión No.	7	Fecha	04/10/2024
Lugar o medio:	Google Meets	Hora inicio:	
		Hora finalización:	

Motivo	de la reunión:	Determinar	aspect	os sobre el apartado de	análisis de resultados
Participantes: Present		Presentes:		Néstor Morales	
		Nathalya Barrera Tova		ar	
				Pedro Soto	
		Ausentes:		N/A	
			Tem	as tratados	
No.	Asunto		Com	entarios	Acuerdos
1	Experimento y el participantes hur		contr los d	estudió el grupo de rol y el tratamiento de atos que se obtuvo del rimento y la esta	Implementar un ANOVA para el tratamiento de los datos. Determinar aspectos relevantes de la encuesta a los participantes humanos en función de los resultados del experimento.
2	Meta-análisis		imple	planteó la forma de ementar este umento en el TFG.	Dar otro escenario referente al estudio de este TFG que brinde otro tipo de datos que ya se han realizado en casos como éstos.
3	de el m		de ar el ma meto solo	iderando el apartado nálisis de resultados y arco teórico, califica la dología mixta y no cuantitativa.	Hacer una el proceso de la gestión de cambios. Realizar el cambio en la metodología y fundamentarlo.
Temas por tratar				Fecha	Convocados
Tratamiento final de los datos				10/10/2024	Nestor Morales
Documento final del artículo					Nathalya Barrera Tovar Pedro Soto
		I	Firma	s de asistentes	
Ver Apéndice C. Plantilla de firmas de minutas					

9.2.8. Minuta H

Reunión No.		8		Fecha	10/10/2024
Lugar o medio:		Google Meets		Hora inicio:	1:45pm
8				Hora finalización:	2:15
Motivo	de la reunión:	Revisión de	aspect	tos sobre artículo científ	ico
Participantes: Presentes		Presentes:	Néstor Morales		
				Nathalya Barrera Tovar	
Ausentes:		Ausentes:		Pedro Soto	
			Tem	as tratados	
No.	Asunto		Com	entarios	Acuerdos
1	Tema		El tema que se venía		Se implementará otros
			desarrollando rediseñó, por		aspectos relacionados con
		lo tanto, se decidió tomar		el TFG.	
		otra ruta para el desarrollo			
		del artículo.			
2	Modificaciones		Hay	apartados del	Modificar los nombres
				mento que no están	para que cumplan con los
				orados de la manera	lineamientos establecidos
	l l		adecı		para el documento.
Próxima reunión					
Temas por tratar				Fecha	Convocados
Revisión de la sección de análisis de				17/10/2024	Néstor Morales
resultados					Nathalya Barrera Tovar
Revisión del artículo científico					Pedro Soto
Firmas de asistentes					
Ver Apéndice C. Plantilla de firmas de minutas					

9.2.9. Minuta I

Reunión No.	9	Fecha	24/10/2024
Lugar o medio:	Google Meets	Hora inicio:	4:00pm
		Hora finalización:	5:15pm
Motivo de la reunión:	Seguimiento de apartado de análisis de resultados		
Participantes:	Presentes: Néstor Morales		
		Nathalya Barrera Tovar	
		Pedro Soto	

	Ausentes:		N/A			
Temas tratados						
No.	Asunto	Com	entarios	Acuerdos		
1	Análisis de resultados	este ejecu corre	ínea de desarrollo de apartado se está tando de la manera cta, sin embargo, hay ras por realizar.	Realizar cambios para mejorar el apartado		
2	Propuesta de solución	que n de	apartado se considera lo aplica para el modelo TFG que se está cando.	Preguntar a la profesora Yarima si éste se elimina y se cambia por un apartado de discusión.		
3	Conclusiones		minación de lusiones.	Empezar a elaborar este apartado.		
4	Formato del TFG	Verif TFG	ïcar el formato del			
Próxima reunión						
Temas por tratar			Fecha	Convocados		
Retroalimentación de documento de TFG.		27/10/2024	Néstor Morales Nathalya Barrera Tovar Pedro Soto			
Firmas de asistentes						
Ver Apéndice C. Plantilla de firmas de minutas						

9.3. Firmas de Aprobación de Minutas

Reuniones	Firma del profesor tutor	Firma del estudiante
El profesor tutor valida la		
participación en las siguientes		
minutas:		
Reunión 1		
Reunión 2		
Reunión 3		
Reunión 4		
Reunión 5		
Reunión 6		
Reunión 7		
Reunión 8		
Reunión 9		

9.4. Apéndice C. Carta filológica

Carta del Filólogo

Señores:

Instituto Tecnológico de Costa Rica Escuela de Administración de Tecnologías de Información

Por este medio hago constar que el Trabajo Final de Graduación para optar por el grado de Licenciatura en Administración de Tecnología de Información, bajo la modalidad de Seminario de Graduación; denominado: "Comparación de la influencia de la aversión al riesgo en las decisiones tomadas por humanos y por inteligencia artificial generativa", presentado por el sustentante: Pedro Soto Aguilar, con número de cédula:117120351; cumple a cabalidad con los siguientes requisitos:

- Discurso verosímil.
- Independencia de juicio.
- Redacción y ortografía, corregidas por el especialista.

Dado en San José, a los 09 días del mes de noviembre del 2024; a solicitud del interesado.

Atentamente: Geovanny Rivera Fernández. Cédula: 1-0615-0079.

Código: 8836

RIVERA digitalmente por GEOVANNY RIVERA FERNANDE (FIRMA)

Z (FIRMA) 10:15:14-06'00'

GEOVANNY Firmado

Licenciado en la enseñanza del Español, con énfasis en Lingüística.