

DOI: <https://doi.org/10.46296/ig.v7i13.0154>

TENDENCIAS DE LA INTELIGENCIA ARTIFICIAL EXPLICABLE EN EL ÁREA DE PSICOLOGÍA

TRENDS IN EXPLAINABLE ARTIFICIAL INTELLIGENCE IN THE AREA OF PSYCHOLOGY

Ibarra-Pincay Melanie ¹; Alcívar-Cevallos Roberth ²

¹ Facultad de Ciencias Informáticas, Universidad Técnica de Manabí. Portoviejo, Ecuador.
Correo: mibarra5622@utm.edu.ec. ORCID ID: <https://orcid.org/0000-0002-7872-678X>

² Facultad de Ciencias Informáticas, Universidad Técnica de Manabí. Portoviejo, Ecuador.
Correo: roberth.alcivar@utm.edu.ec. ORCID ID: <https://orcid.org/0000-0001-6282-8493>

Resumen

En la actualidad las herramientas de Inteligencia Artificial (IA) tienen un amplio uso en el área de psicología, pero al tratarse del área de la salud los procedimientos y decisiones que toma la IA se convierten en algo sensible de tratar, ahí entra la Inteligencia Artificial Explicable (XAI) que brinda una gran variedad de procesos que hacen posible el uso de IA en esta área de salud. Esta revisión sistemática de la literatura (SLR) demuestra que la XAI es una herramienta valiosa para esta área, ya que ofrece una amplia variedad de procesos y métodos útiles para mejorar la precisión de los diagnósticos, evaluar la respuesta al tratamiento, entender la conducta y patrones de comportamiento humano logrando acelerar el proceso de atención. Este estudio da respuesta a 3 preguntas de investigación que logran dar pie a la aceptación y uso de las herramientas que brinda la XAI, dando también un resumen de los estudios seleccionados para esta SLR.

Palabras clave: XAI, Psicología, IA, SLR, Métodos.

Abstract

Currently, Artificial Intelligence (AI) tools have wide use in the area of psychology, but when it comes to the area of health, the procedures and decisions made by AI become something sensitive to deal with, that is where Artificial Intelligence comes in. Explainable (XAI) that provides a wide variety of processes that make the use of AI possible in this health area. This systematic literature review (SLR) demonstrates that XAI is a valuable tool for this area, offering a wide variety of processes and methods useful for improving diagnostic accuracy, assessing treatment response, understanding behavior, and human behavior patterns, managing to accelerate the attention process. This study answers 3 research questions that lead to the acceptance and use of the tools provided by the XAI, also giving a summary of the studies selected for this SLR.

Keywords: XAI, Psychology, AI, SLR, methods.

Información del manuscrito:

Fecha de recepción: 18 de octubre de 2023.

Fecha de aceptación: 19 de diciembre de 2023.

Fecha de publicación: 10 de enero de 2024.



1. Introducción

La Inteligencia Artificial (IA) es una disciplina con extensa aplicación en el área de la psicología, que va desde la comprensión del comportamiento humano hasta la mejora del diagnóstico y la aplicación de tecnología en la práctica clínica. En los últimos años, el creciente uso de la IA en distintas áreas ha ido desembocando la creación de otras grandes ramas dentro de este campo, lo que nos lleva a la IA explicable o XAI por sus siglas en inglés (eXplainable Artificial Intelligence), que persigue la creación de algoritmos que permitan una mayor comprensión por parte del usuario sobre cómo funciona el sistema. Según el estudio de Tai et al. (2020), los modelos XAI ayudan a los profesionales de la salud mental a entender mejor los resultados o decisiones de los algoritmos de IA, permitiendo una mejor aplicación de la tecnología en la práctica clínica.

La Inteligencia Artificial Explicable (XAI), es una herramienta que permite la transparencia en la explicación de procesos llevados a cabo por la Inteligencia Artificial (IA), por tanto, un sistema de IA debe proporcionar evidencia, soporte; o

razonamiento para cada decisión realizada por el sistema. (Angelov et al. 2021).

Los avances tecnológicos en el área de la IA explicable permiten a los profesionales de la salud mental identificar mejor los patrones de comportamiento, realizar un diagnóstico más temprano y aplicar conocimientos de forma eficaz. Convirtiéndose en una tendencia importante en el área de la psicología, donde la incorporación exitosa de la IA ha abierto puertas a nuevos modelos clínicos. Según un estudio de Bhowmik et al. (2019), el uso de IA explicable permite mejorar el diagnóstico de trastornos psicológicos, lo que fomenta el desarrollo de sistemas más precisos para ayudar a los profesionales a tomar decisiones clínicas informadas.

En Ecuador no se cuenta con implementaciones ni con recursos bibliográficos que abalen los beneficios que puede traer la XAI en el área de psicología, también cabe mencionar que la psicología ayudaría notablemente a la disciplina que aplica la IA desarrollando nuevos modelos explicativos, siendo recíproco un enlace entre ambas

disciplinas, la XAI ayudando a tomar decisiones informadas a los profesionales de la salud mental y los profesionales de la salud mental ayudando a la XAI a mejorar o crear nuevos modelos explicativos que podrían ayudar a los mismos en un futuro no muy lejano. Por ello se crea este artículo, el cual contendrá una Revisión Sistemática de la Literatura o por sus siglas en inglés Systematic Literature Review (SLR) que permitirá saber cuáles son las tendencias de XAI en psicología, basándonos en una investigación documental.

2. Materiales y métodos

La presente investigación es una Revisión Sistemática de la Literatura (SLR), la cual se basa en el proceso y los lineamientos realizados, en el estudio realizado por (Loor Zambrano, Tello Salvador, Alcívar Cevallos, & Vaca Cardenas, 2020). Este SLR consiguió seguir lo planteado a través de las siguientes etapas: Planificación, Revisión y Análisis.

2.1. Planificación

Esta revisión sistemática de literatura (SLR) fue realizada de

forma empírica. La información necesaria para la investigación se recopiló utilizando la herramienta informática Parsifal (Vázquez-Lozano et al., 2020). Como gestor de búsqueda también se utilizó Google Académico, Zapmeta, Metacrawler, el método de búsqueda en estos gestores de bases de datos fue empírico y se basó en una cadena de búsqueda corta que permitía el hallazgo de varios documentos encontrados paulatinamente según los términos de búsqueda. Se utilizaron bases de datos académicas como Scopus, IEEE Digital Library, Springer Link, Scielo, Science@Direct. Utilizando diversos términos de búsqueda (Cadena de Búsqueda) relacionados con el tema de la investigación, que se combinaron entre sí con los operadores Booleanos AND, OR y NOT, cabe destacar que las cadenas de búsquedas utilizadas variaban según el gestor o base de datos usado, ya que cada uno contiene una infinidad de documentación y para hacerlo de forma precisa se crearon cadenas de búsquedas para cada base de datos (ver en tabla I). Se indagó en los estudios seleccionados y se seleccionaron de

acuerdo con los criterios de inclusión y exclusión (ver en tabla II).

Esta búsqueda incluyó estudios primarios publicados entre 2016 y 2023. Se recopilaron los resultados de los estudios primarios, como los marcadores de inteligencia artificial explicable, los algoritmos de XAI que se usan en psicología y sus tendencias en el ámbito de la psicología. Además, se realizaron preguntas de investigación que ayudaron a establecer el problema objetivo, el alcance de la indagación y el enfoque necesario para abordar el tema. Las preguntas de investigación también ayudan a identificar los recursos relevantes para la búsqueda, junto con los datos necesarios para obtener una conclusión confiable. Al formular estas preguntas relevantes y específicas, se puede mejorar la calidad y profundidad del artículo académico, lo que resulta en un

producto final más informativo y convincente. Esto de acuerdo con el profesor de la Universidad de Duke, Robert J. Sternberg (2020)

A. Preguntas de investigación:

RQ1. ¿Para qué sirve y como se usa la Inteligencia Artificial Explicable en el área de la psicología?

RQ2. ¿Qué tendencias existen en el área de psicología aplicando la Inteligencia Artificial Explicable?

RQ3. ¿Qué procesos o métodos de la Inteligencia Artificial Explicable son aplicables en el área de la psicología?

B. Cadena de Búsqueda

("Inteligencia Artificial Explicable" OR "Models" OR "Inteligencia Artificial application" OR "algoritmos" AND "psychology")

Inteligencia Artificial Explicable, Psicología, Inteligencia Artificial aplicable, modelos, algoritmos.

C. Cadenas de búsqueda según el gestor de base de datos académicas

Tabla 1. Cadena de búsqueda usada para cada Base de Datos Académica.

Cadena de búsqueda	Bases de Datos Académicas
("XAI" AND "psychology")	Scielo
("Inteligencia Artificial Explicable" OR "XAI" AND "psychology")	Scopus
("Inteligencia Artificial Explicable" OR "XAI" OR "algoritmos" AND "psychology")	IEEE Digital Library
("XAI" AND "psychology")	Springer Link

("Inteligencia Artificial Explicable" OR "XAI" OR "algoritmos" AND "psychology")	Science@Direct
("XAI" AND "psychology")	Otros

Fuente: Elaboración propia.

D. Criterios de selección:

Tabla 2. Criterios de exclusión e inclusión.

Criterios de Inclusión	Criterios de Exclusión
Publicaciones Académicas de los últimos 7 años	Estudios Duplicados
Publicaciones relevantes para el tema	Estudios sin discusión de resultados
Estudios primarios	Informes sin rigor científico
Investigación que describan su metodología y sus resultados de forma precisa	Tesis sin rigor científico
	Estudios que no contengan descripción de su metodología ni sus resultados

Fuente: Elaboración propia.

2.2. Revisión

En la realización de la investigación se empezó con la búsqueda y selección de los estudios, estos teniendo que pasar primero por los criterios planteados para inclusión y exclusión, posteriormente ser evaluados para ver si responden a las preguntas de investigación desarrolladas.

Se procede a realizar las indagaciones en las respectivas bases de datos a través de los gestores mencionados, en los primeros hallazgos se obtenían cantidades grandes en lo que refiere a exploración con la cadena de búsqueda base, filtrando los resultados y mejorando la cadena de búsqueda se consigue menos resultados, por ende más

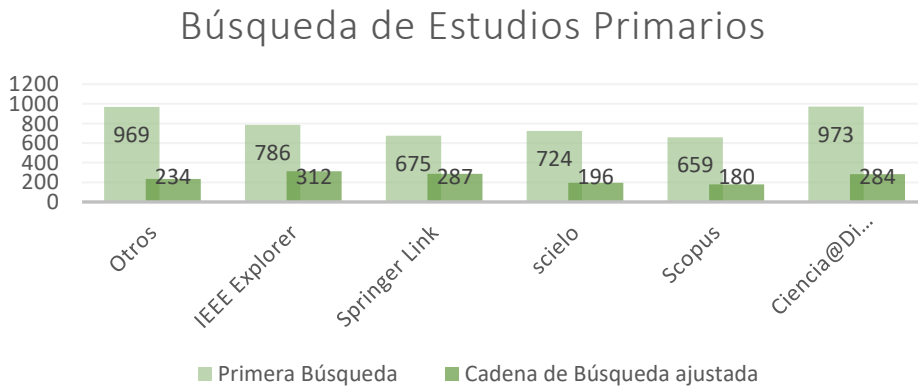
específicos y acordes al tema, una muestra de los datos que se obtuvieron se encuentra planteado en el gráfico (grafico I), la mejora de la cadena de búsqueda permitió una adaptación optima a lo que se precisa de información y posteriormente dando paso a la exploración de los estudios seleccionados, procesos similares se dieron con las demás bases de datos utilizadas en esta investigación para la búsquedas de los estudios, lo que dio como resultado el hallazgo de 55 estudios primarios (grafico II) que permitieron la ejecución de este estudio.

En el siguiente grafico (grafico III) se observa la tendencia en incremento de publicaciones en los últimos 4 años, por lo que se estima que la XAI

ha ido ganando popularidad, notando que en los últimos 3 años se ha incrementado la publicación de estudios de este tema. Relacionando los temas con los estudios más

comúnmente asociados con el ámbito de la psicología y la medicina, tal como lo menciona Miler al et (2018) en el estudio que realizó.

Gráfico 1. Comparativa entre búsqueda normal y búsqueda con cadena de búsqueda personalizada.



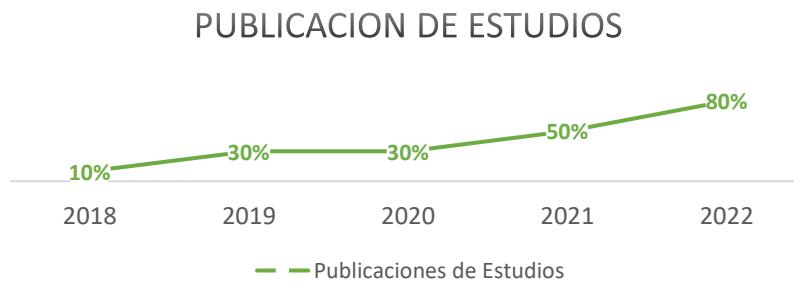
Fuente: Elaboración propia.

Gráfico 2. Cantidad de Estudios Seleccionados según su origen o Base de Datos.



Fuente: Elaboración propia.

Gráfico 3. Porcentaje de publicaciones encontradas según el año de publicación.



Fuente: Elaboración propia.

2.3. Análisis

Una vez seleccionados los estudios primarios se empezó con el análisis de cada uno de ellos, seccionando, clasificando y filtrando los estudios obtenidos, para organizar de manera óptima los resultados de esta investigación.

3. Resultados y discusión

Los resultados obtenidos se clasificaron según las preguntas de investigación a las que responden, (RQ1, RQ2, RQ3), posteriormente se realizó un análisis acerca de cuáles son los algoritmos o métodos más usados en el área de Psicología. Las consultas realizadas a los estudios primarios seleccionados se realizaron meticulosamente para dar con resultados precisos que muestren el inicio del camino para la SLR. Los beneficios, como se usa y cuáles son las tendencias al aplicar los métodos XAI se responden en las RQ1 y RQ2. Mientras que para medir cuales son los métodos de XAI más usados se tomara la RQ3.

RQ1. ¿Para qué sirve y como se usa la Inteligencia Artificial Explicable en el área de la psicología?

Con base a los estudios primarios recolectados se analizó a profundidad lo referente al tema a tratar, como resultado se hallaron diversas fuentes de información que presentan varios puntos de vistas, así como un conjunto de información que coincide en que, la XAI es una disciplina (anexo) que basa su eje en la explicación del porqué de la toma de decisiones que realiza la IA en cualquier área de conocimiento en la que se la utilice, permitiendo así un apoyo a los profesionales que utilizan herramientas de IA para realizar cualquier tipo de proceso; ahora bien, centrando la atención en lo que refiere a sus usos en el área de psicología según (Milda Poceviciute, Eliertsen, & Lundstrom, 2020) la XAI ofrece una variedad de herramientas que pueden ser aplicadas en la práctica clínica y científica, así como en la comprensión de los procesos de la mente humana. En el área de la psicología, la XAI se ha empleado para mejorar la precisión de los diagnósticos, la evaluación de la respuesta al tratamiento, el entendimiento de la conducta y el análisis de los patrones de comportamiento humano.

Un ejemplo de aplicación de la XAI en el área de la psicología es la implementación de algoritmos para predecir la respuesta de los pacientes a los tratamientos. Permitiendo a los profesionales de la salud anticipar la efectividad de los tratamientos, lo que puede ayudar a acelerar el proceso de atención. Para muestra de ello, un estudio reciente realizado por el Departamento de Psicología de la Universidad de Stanford encontró que los algoritmos de XAI aplicados a los datos del tratamiento de pacientes con trastornos de ansiedad mostraron una mejoría significativa en la precisión de la predicción de los resultados del tratamiento, en comparación con los métodos tradicionales (Gill, Smith, & Smith, 2020). Empleado en el área de la psicología para el análisis de los patrones de comportamiento humano. Esto se realiza mediante el uso de algoritmos para detectar patrones ocultos en la conducta humana. Estos algoritmos pueden ser usados para identificar patrones en los datos recogidos durante las entrevistas clínicas, lo que puede ayudar en la identificación de trastornos psicológicos (Luo, Zhou, Shang, & Xu, 2020). Además, la XAI

se utiliza para extraer conocimiento de los datos recogidos durante los estudios científicos, lo que puede ayudar a los psicólogos a comprender mejor los procesos de la mente humana.

RQ2. ¿Qué tendencias existen en el área de psicología aplicando la Inteligencia Artificial Explicable?

En el campo de la psicología, hay una tendencia creciente de aplicar la Inteligencia Artificial Explicable para resolver problemas psicológicos. Un ejemplo claro de ello es la utilización de la XAI para predecir el estado emocional y el comportamiento de los usuarios. Esto se ha hecho utilizando la recopilación de datos de redes sociales para identificar el estado emocional de los usuarios. Permitiendo a los psicólogos comprender mejor los patrones de comportamiento de sus pacientes. Asimismo, el uso de XAI también se ha destinado para predecir trastornos mentales, como la depresión, utilizando modelos basados en datos y análisis de texto (Kumar, Subramanian, & Balasubramanian, 2018).

La inteligencia artificial explicable se ha manejado para crear nuevos modelos de diagnóstico basados en

datos médicos para enfermedades mentales como el trastorno bipolar o el trastorno de déficit de atención con hiperactividad (ADHD). Esto ha permitido que los profesionales médicos puedan realizar un diagnóstico más preciso y rápido de estas enfermedades mentales (Ghorbani, 2017).

Así como en los anteriores puntos la tendencia crece a medida que el área de aplicación de XAI se extiende, por ello se ha observado una tendencia creciente en el uso de la Inteligencia Artificial Explicable en el área de la psicología. Esto ha permitido el desarrollo de nuevas tecnologías que han contribuido a mejorar la calidad de vida de muchas personas (Bauer, von Zahn, & Hinz, 2021). Estas tecnologías incluyen el uso de XAI para predecir el estado emocional y el comportamiento de los usuarios, y para el diagnóstico de enfermedades mentales (Fellous, Sapiro, Rossi, Mayberg, & Ferrante, 2019).

RQ3. ¿Qué procesos o métodos de la Inteligencia Artificial Explicable son aplicables en el área de la psicología?

Uno de los principales procesos de XAI que puede aplicarse en el área de psicología es el modelado interpretable. Este se refiere al uso de técnicas de inteligencia artificial para que los modelos sean explicables a nivel de humano. Lo que significa que los modelos pueden interpretarse por los usuarios para comprender cómo se está llevando a cabo una tarea específica, ayudando así a los psicólogos a comprender mejor los procesos psicológicos subyacentes a determinados comportamientos. (Kumar, Subramanian, & Balasubbramanian, 2018)

Otro proceso de XAI que puede ser útil en el campo de la psicología es la visualización de la profundidad. Esta técnica refiere al uso de visualizaciones para explicar los procesos internos de un modelo de inteligencia artificial. Lo que permite a los usuarios comprender cómo se está llevando a cabo un proceso, lo que es útil para que los psicólogos entiendan mejor los procesos subyacentes a determinados comportamientos. (Kindermans & Al-Shedivat, 2017)

La explicabilidad en la XAI resulta sumamente beneficiosa para llevar a

cabo la evaluación de tratamientos psicológicos. La utilización de los algoritmos que presenta la inteligencia artificial explicable permite a los psicólogos una evaluación precisa de la eficacia que tienen los tratamientos en determinados pacientes, obteniendo una explicación detallada de los mismos. Permitiendo a los profesionales de la salud mental comprender las razones del por qué un determinado tratamiento no es efectivo y ajustar el tratamiento en consecuencia (Dazeley, y otros, 2021).

La XAI ofrecer una amplia gama de procesos y técnicas valiosas para el área de psicología. Estas incluyen el modelado interpretable, la visualización de la profundidad y la evaluación de la efectividad de los tratamientos psicológicos. Estas técnicas de XAI ofrecen una forma potencialmente útil de abordar preguntas clave en el área de la psicología.

4. Conclusiones

Los hallazgos de esta revisión sistemática de la literatura (SRL) concluyen con la revisión de 45

artículos que fueron relevantes para el tema, siendo que la tendencia que existe en la actualidad es que el uso de la XAI en el área de la psicología va en aumento, lo cual indica el gran auge que pueda surgir debido a los beneficios que brinda en esta área y en otras áreas en las que es utilizado. Durante el desarrollo de esta revisión se halló que algunos de los métodos más usados en la XAI son: Numérico, basado en reglas, textuales, visuales, mixtos, entre otros; los cuales son las bases para el desarrollo de una XAI,

En la tabla (tabla IV) se organiza la información encontrada de manera más detallada, referenciando los estudios primarios encontrados, considerando: el nombre de los estudios y su respectiva cita, el enfoque del estudio redactado en un breve resumen, el tipo de estudio y los Métodos de XAI que mencionan en el estudio. En la sección que se menciona a los métodos usados en los estudios seleccionados se plantea según (Vilone & Longo, Explainable Artificial Intelligence: a Systematic Review, 2020) que los métodos son:

Numérico: Métodos de atribución de características aditivas, básicamente

combinaciones lineales de las características de entrada, para construir un modelo que es una interpretación interpretable.

Reglas: Métodos producen explicaciones basadas en reglas mediante la explotación de varias técnicas de extracción de reglas, como los enfoques basados en el razonamiento automatizado.

Textual: Método que produce explicaciones en lenguaje natural.

Visual: Explicar el funcionamiento interno de un modelo a través de ayudas gráficas.

Mixto: Producen explicaciones numéricas junto con representaciones gráficas para hacerlas más interpretables.

Otros: Basados en otros métodos que dan como resultado la explicabilidad.

N/A: No usa ni menciona métodos en el estudio.

Tabla 4. Organización de estudios primarios seleccionados según el enfoque o descripción que contiene o sección que abarcan, con secciones que hacen referencia tipo de estudio y Métodos usados en el estudio.

	Nombre del Estudio/ Artículo y su referencia	Enfoque o descripción breve del Estudio/ Artículo	Tipo de estudio	Métodos usados en el estudio
1	xxAI – Beyond Explainable AI (Holzinger, y otros, 2020)	Habla sobre el éxito de la inteligencia artificial (IA) basada en el aprendizaje automático (ML) y la necesidad de desarrollar herramientas que sean tanto predictivas como interpretables para los humanos.	Libro.	Numérico, Reglas, Textual, Visual, Mixto, Otros.
2	Psychological Foundations of Explainability and Interpretability in Artificial Intelligence (Broniatowski, 2021)	Argumenta que la interpretabilidad y la explicabilidad son requisitos distintos para los sistemas de aprendizaje automático.	Artículo de Divulgación Científica.	Reglas, Numéricos, Visual.
3	Four Principles of Explainable Artificial Intelligence (Phillips, y otros, 2021)	Presenta cuatro principios para la inteligencia artificial explicada (IAE) que comprenden propiedades fundamentales para los sistemas de IAE explicados. Estos cuatro principios se denominan explicación, significativo, precisión en la explicación y límites del conocimiento.	Artículo de Divulgación Científica.	N/A
4	Mapping of machine learning approaches for description, prediction, and causal inference in the social and health sciences (Lesit, y otros, 2022)	Discute el potencial de la utilización de métodos de aprendizaje automático (machine learning o ML) en la investigación social y de la salud. Se mencionan sus ventajas, como la capacidad de procesar grandes cantidades de datos, detectar relaciones no lineales y mejorar la precisión en la predicción.	Artículo de Divulgación Científica.	N/A

5	Explainable Artificial Intelligence Methods in Combating Pandemics: A Systematic Review (Giuste, y otros, 2021)	Analiza el uso de la inteligencia artificial explicativa (XAI) durante la pandemia de COVID-19 y cómo su uso puede superar las barreras para lograr un impacto clínico significativo	Revisión sistemática de la Literatura.	Visual, Mixto, otros.
6	Roadmap of Designing Cognitive Metrics for Explainable Artificial Intelligence (XAI) (H. HSIAO, HT NGAI, QIU, YANG, & Chen CAO, 2021)	Discute cómo los métodos utilizados en ciencias cognitivas y psicología para inferir los procesos cognitivos detrás del comportamiento humano pueden ser aplicados para examinar los mecanismos que subyacen el comportamiento de los sistemas de inteligencia artificial (IA) y proporcionar explicaciones.	Artículo de divulgación científica.	Mixto, Otros.
7	A mental models approach for defining explainable artificial intelligence (Merry, Pat, & Jim, 2021)	Discute las preocupaciones existentes sobre el uso de métodos de modelado de caja negra en contextos sensibles como la atención médica, y cómo la falta de explicabilidad de estos modelos obstaculiza la adopción de la inteligencia artificial (IA) en estos campos.	Revisión sistemática de la Literatura.	Numérico, visual, otros.
8	¿Qué queremos de la Inteligencia Artificial Explicable (XAI)? – Una perspectiva de las partes interesadas sobre la XAI y un modelo conceptual que guía la investigación interdisciplinaria de la XAI (Langer, y otros, 2021)	Perspectiva de las partes interesadas sobre la XAI y un modelo conceptual que guía la investigación interdisciplinaria de la XAI.	Artículo de Divulgación Científica.	N/A
9	Mirando dentro de la caja negra: una encuesta sobre Inteligencia Artificial Explicable (XAI) (Adadi & Berrada, 2018)	Información sobre la Inteligencia Artificial Explicable (XAI) y cómo se está desarrollando en la actualidad.	Survey (Encuesta).	N/A
10	Explanations of artificial intelligence: Author proposes model that highlights evidence of fairness (Explanations of artificial intelligence: Author proposes model that highlights evidence of fairness, 2023)	Información sobre la creciente demanda de la explicabilidad de la inteligencia artificial (IA) en contextos de alta importancia ética y regulatoria, como por ejemplo en la creación de nuevas formas de puntajes crediticios más allá del FICO tradicional.	Revista de congreso.	N/A
11	Expl(AI)ned: The Impact of Explainable Artificial Intelligence on Cognitive Processes (Bauer, von Zahn, & Hinz, 2021)	Explora el impacto de las explicaciones basadas en características en el procesamiento de información de los usuarios de sistemas de inteligencia artificial (IA) modernos.	Libro.	Numérico, reglas, textual, otros.
12	Explanation in artificial intelligence: Insights from the social sciences (Miller, 2019)	Aborda el resurgimiento de la inteligencia artificial explicada y cómo se puede mejorar el proceso de explicación, revisa estudios relevantes y destaca hallazgos importantes para informar la mejora en la explicación de la inteligencia artificial.	Revisión sistemática de la Literatura.	N/A
13	Aproximaciones a la Explicación de Decisiones algorítmicas: Inteligencia Artificial Explicable (Vaquero de Miguel, 2020)	Información sobre cómo se pueden explicar las decisiones tomadas por algoritmos de inteligencia artificial.	Artículo de Divulgación Científica.	N/A

14	The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies (Markus, Kors, & Rijnbeek, 2021)	Señala que, aunque la inteligencia artificial (IA) tiene un gran potencial para mejorar la salud y el bienestar de las personas, su adopción en la práctica clínica sigue siendo limitada debido a la falta de transparencia.	Sruvey (Encuesta).	N/A
15	A historical perspective of explainable Artificial Intelligence (Confalonieri, Coba, Wagner, & Besold, 2020)	Presenta una perspectiva histórica sobre el desarrollo de la XAI, discute su entendimiento actual y propone criterios para las explicaciones que jugarán un papel crucial en el desarrollo de sistemas de XAI comprensibles para los seres humanos en el futuro.	Artículo de Divulgación científica.	N/A
16	Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research (Haque, Islam, & Patrick, 2023)	Presenta una revisión sistemática de la XAI desde la perspectiva del usuario final. El estudio tiene como objetivo identificar las dimensiones de las necesidades de explicación de los usuarios finales, investigar el efecto de la explicación en las percepciones de los usuarios finales y proponer agendas de investigación futuras para XAI.	Revisión Sistemática de la Literatura.	Textual, Visual.
17	Aplicaciones de la Teoría de la Atribución Social en XAI (Zhang, 2022)	Aborda el problema central de la ética de la inteligencia artificial, que consiste en cómo otorgar a la IA el estatus de agente y ganar la confianza de los seres humanos.	Artículo de Divulgación Científica.	N/A
18	Niveles de inteligencia artificial explicable para explicaciones conversacionales alineadas con humanos (Dazeley, y otros, 2021)	Describe el creciente interés en la Inteligencia Artificial Explicable (XAI) y el aprendizaje automático interpretable (IML), y cómo la falta de explicaciones comprensibles para las decisiones de la IA puede generar preocupación en el público.	Artículo de Divulgación Científica.	N/A
19	Una encuesta sobre inteligencia artificial explicable (XAI): Hacia la XAI médica (Tjoa & Guan, 2021)	Menciona cómo la inteligencia artificial y el aprendizaje automático han demostrado desempeños notables en diversas tareas, pero la falta de transparencia en los algoritmos puede generar preocupación, especialmente en sectores como la medicina que requieren mayor rendición de cuentas.	Survey (Encuesta).	N/A
20	Contrafactuales en Inteligencia Artificial Explicable (XAI): Evidencia del razonamiento humano (Byrne, 2019)	Discute el uso creciente de los contrafactuales en aplicaciones de Inteligencia Artificial (IA), especialmente en IA Explicable (XAI). Estos pueden ayudar a crear modelos interpretables y hacer que las decisiones de sistemas opacos sean comprensibles para desarrolladores y usuarios.	Artículo de Divulgación Científica.	N/A
21	Artificial Cognition: An emerging discipline to explain the decision making of artificial neural networks (Iglesias, 2020)	Explica como la cognición artificial busca explicar el proceso de toma de decisiones que realizan las redes neuronales, basándose en métodos de ciencia cognitiva.	Revista de Congreso.	N/A
22	Una encuesta de y contrastes	Aborda la falta de interpretabilidad de algunos algoritmos de inteligencia	Survey (Encuesta).	N/A

	contrafactuales de generación de explicaciones para explicables Inteligencia artificial (Stepin, Jose M., Catala, & Pereira-Fariña, 2021)	artificial y la importancia de proporcionar explicaciones de sus decisiones.		
23	A historical perspective of explainable Artificial Intelligence (Confalonieri, Coba, Wagner, & Besold, 2020)	Aborda la importancia de la explicabilidad en la inteligencia artificial, especialmente en aplicaciones como la conducción autónoma, el diagnóstico médico o la banca y las finanzas, donde es crucial transmitir seguridad y confianza a los usuarios en las decisiones automatizadas.	Artículo de Divulgación Científica.	N/A
24	inteligencia artificial y etica (Sebio Martin, 2020)	Explora las implicaciones éticas del uso de la Inteligencia Artificial en la toma de decisiones. Se analiza el concepto de Inteligencia Artificial, los posibles problemas éticos y las soluciones actuales y futuras. Se concluye que es importante que los seres humanos mantengan el control sobre la tecnología y su uso.	Tesis.	N/A
25	Explainable Artificial Intelligence: a Systematic Review (Vilone & Longo, Explainable Artificial Intelligence: a Systematic Review, 2020)	Destaca el crecimiento significativo de la Inteligencia Artificial Explicable (XAI) en los últimos años debido a la amplia aplicación del aprendizaje automático, especialmente el aprendizaje profundo, que ha llevado al desarrollo de modelos altamente precisos, pero sin explicabilidad e interpretabilidad.	Revisión Sistemática de la Literatura.	Numérico, Textual, Mixto, Reglas, Visual, Otros.
26	Explainable AI: A Review of Machine Learning Interpretability Methods (Linardatos, Papastefanopoulos, & Kotsiantis, 2020)	Aborda el aumento en la adopción industrial de la inteligencia artificial, pero también sobre la complejidad de los modelos de aprendizaje automático, que a menudo se vuelven "cajas negras" que dificultan la comprensión de su funcionamiento y toma de decisiones.	Artículo de Divulgación Científica.	Otros.
27	Explainable AI: Foundations, Applications, Opportunities for Data Management Research (Pradhan, Lahiri, Galhotra, & Salimi, 2022)	Describe cómo los sistemas de toma de decisiones algorítmicas están siendo adoptados en diversas áreas y tareas, pero también cómo la confianza en ellos es cada vez más importante.	Artículo de Divulgación Científica.	Reglas, textual, mixto, otros.
28	Towards Quantification of Explainability in Explainable Artificial Intelligence Methods (Rabiul Islam, Eberle, & Ghafoor)	Menciona que la Inteligencia Artificial (IA) se ha convertido en una parte integral de diferentes dominios, pero explicar las decisiones de los sistemas de IA es un desafío debido a su alta complejidad y a las posibles implicaciones en los intereses y derechos humanos.	Artículo de Divulgación Científica.	Otros.
29	On the importance of experimental psychology for explainable artificial intelligence (Leuven, 2022)	Explica que la Inteligencia Artificial Explicable (XAI) ha ganado popularidad en la investigación reciente debido a que los usuarios humanos desean comprender por qué la IA llega a una decisión en particular y se comporta como lo hace.	Artículo de Divulgación Científica.	Otros.

30	Explainability of AI-predictions based on psychological profiling (Neugebauer, Rippitsch, Sobieczky, & Geib, 2021)	Propone un nuevo método de predicción para el desempeño de empresas emergentes basado en perfiles psicológicos, utilizando un enfoque de inteligencia artificial explicable llamado "Before and After prediction Parameter Comparison (BAPC)".	Artículo de Divulgación Científica.	Otros
31	Classification of Explainable Artificial Intelligence Methods through Their Output Formats (Vilone & Longo, Classification of Explainable Artificial Intelligence Methods through Their Output Formats, 2021)	Presenta una revisión sistemática de los métodos utilizados para explicar el funcionamiento de modelos de aprendizaje automático y profundo, los cuales a menudo son difíciles de interpretar debido a sus estructuras complejas.	Revisión Sistemática de la Literatura.	Número, Textual, Mixto, Reglas, Visual, Otros.
32	Explainable AI as evidence of fair Decisions (Leben, 2023)	Propone que las explicaciones son valiosas para aquellos que son impactados por las decisiones de un modelo (pacientes modelo) en la medida en que proporcionan evidencia de que una decisión adversa previa fue injusta.	Artículo de Divulgación Científica.	Otros.
33	Explainable Artificial Intelligence for Neuroscience: Behavioral Neurostimulation (Fellous, Sapiro, Rossi, Mayberg, & Ferrante, 2019)	Discute el valor potencial de XAI para el campo de la neuroestimulación tanto para la investigación científica básica como para fines terapéuticos, así como las preguntas pendientes y obstáculos para el éxito del enfoque XAI.	Artículo de Divulgación Científica.	N/A
34	Examining the effects of power status of an explainable artificial intelligence system on users' perceptions (Ha, June Sah, & Sangwon)	Investiga cómo las personas perciben la capacidad de los sistemas de inteligencia artificial explicables (XAI) en función de sus cualidades atributivas percibidas y cómo el estado de poder y la antropomorfización afectan el proceso de atribución.	Artículo de Divulgación Científica.	N/A
35	Emotion-Aware Explainable Artificial Intelligence for Personality, Emotion, and Mood Simulation (Valle-Cruz, Lopez Chau, Ramos Corchado, & Muños Jimenez, 2022)	Persigue dos objetivos principales. En primer lugar, se realiza un análisis del estado actual de la Inteligencia Artificial Explicable (XAI) aplicada al cómputo afectivo, el análisis de emociones y el análisis de sentimientos. En segundo lugar, se prueba un marco de trabajo de XAI consciente de las emociones para determinar el estado afectivo de entidades autónomas predefinidas en contextos de incertidumbre.	Revisión Sistemática de la Literatura.	N/A
36	Explainable AI: A Hybrid Approach to Generate Human-Interpretable Explanation for Deep Learning Prediction (De, Giri, Mevawala, Nemani, & Deo, 2020)	Aborda la importancia de la explicabilidad en la toma de decisiones basada en modelos de inteligencia artificial (IA), especialmente en los modelos de aprendizaje profundo (deep learning), que son cada vez más complejos.	Artículo de Divulgación Científica.	N/A
37	Artificial cognition: How experimental psychology can help generate explainable artificial intelligence (T.Taylor & Taylor, 2020)	Describe cómo los psicólogos cognitivos pueden contribuir a la XAI, ya que la mente humana también es una caja negra y los psicólogos cognitivos tienen más de 150 años de experiencia modelando la mente a través de la experimentación. Se proporciona una revisión de la XAI para los psicólogos, argumentando que los	Artículo de Divulgación Científica.	N/A

		métodos actuales tienen un punto ciego que puede complementarse con la tradición experimental cognitiva.		
38	Una encuesta de análisis visual para métodos de inteligencia artificial Explicable (Alicioglu & Sun, 2022)	Explora las tendencias y desafíos actuales del análisis visual en la interpretación de modelos de aprendizaje profundo (DL) mediante la adopción de métodos de Inteligencia Artificial Explicable (XAI), con el objetivo de presentar futuras direcciones de investigación en esta área.	Survey (Encuesta).	N/A
39	Descripción general de la inteligencia artificial explicable para Pronóstico y Gestión de la Salud de Activos Industriales Basado en elementos de informes preferidos para revisiones sistemáticas y metanálisis (Mohd Nor, Rao Pedapat, Muhammad, & Leiva, 2021)	Presenta una revisión sistemática y metaanálisis de los trabajos que se han realizado en relación con la aplicación de la inteligencia artificial explicable (XAI) en la disciplina de Pronósticos y Gestión de la Salud (PHM) de activos industriales.	Revisión Sistemática de la Literatura.	N/A
40	Inteligencia artificial y valoración de la prueba: las garantías jurídico-constitucionales del órgano de control (Castellano, 2021)	Analiza cómo la inteligencia artificial está impactando en la administración de justicia y se explora los fundamentos y campos de aplicación de los algoritmos basados en el aprendizaje autónomo.	Artículo de Divulgación Científica.	N/A
41	Explainable artificial intelligence based analysis for interpreting infant fNIRS data in developmental cognitive neuroscience. (Andreu-Perez, y otros, 2021)	Presenta una nueva técnica de análisis de datos de neuroimagen funcional llamada xMVPA, que utiliza la Inteligencia Artificial Explicable (XAI) para explorar el desarrollo cognitivo del cerebro humano.	Artículo de Divulgación Científica.	Visual, Numérico, otros.
42	New Explainability Method based on the Classification of Useful Regions in an Image. (Guerrero Velázquez & Sossa Azuela)	Plantea la necesidad de evaluar la fiabilidad de los modelos de aprendizaje automático utilizados en la industria, y cómo las métricas actuales no ofrecen una visión general de la consistencia o estabilidad de las predicciones.	Artículo de Divulgación Científica.	Visual, otros.
43	Survey of XAI in digital pathology. (Milda Pocevičiute, Eliertsen, & Lundstrom, 2020)	Este artículo presenta un estudio sobre la aplicación de la inteligencia artificial (IA) en el diagnóstico médico a través de la imagenología. Aunque la IA tiene un gran potencial, su implementación en la práctica médica plantea desafíos. Es importante que los algoritmos sean precisos, pero también comprensibles y confiables.	Survey (Encuesta).	Numérico, Textual, Mixto, Reglas, Visual, Otros.
44	DARPA's Explainable Artificial Intelligence Program. (Gunning & Aha, 2019)	Aborda como DARPA lanzó el programa XAI en 2017 con el objetivo de crear sistemas de IA que puedan proporcionar explicaciones efectivas sobre su comportamiento, fortalezas y debilidades a los usuarios humanos.	Artículo de Divulgación Científica.	Visual, otros.
45	Explainable Artificial Intelligence. (XAI) (Ridley, 2022)	Propone que las bibliotecas académicas adopten la inteligencia artificial explicable (XAI) como conjunto de herramientas para verificar y validar	Artículo de Divulgación Científica.	N/A

los recursos creados y entregados por
sistemas de aprendizaje automático.

Fuente: Elaboración propia.

La XAI se está convirtiendo en una herramienta fundamental para la psicología como consecuencia de lo anteriormente planteado. Esto se debe a que puede identificar enfermedades mentales y predecir el estado emocional y el comportamiento de los usuarios. Esta herramienta está mejorando significativamente la calidad de vida de muchas personas. En conclusión, esta revisión exhaustiva de la literatura ha demostrado el valor de la Inteligencia Artificial eXplicable en el campo de la psicología.

Bibliografía

- Adadi, A., & Berrada, M. (16 de Septiembre de 2018). Mirando dentro de la caja negra: una encuesta sobre Inteligencia Artificial Explicable (XAI). *IEEE Access*, 6, 52138-52160. doi: 10.1109/ACCESS.2018.2870052
- Alicioglu, G., & Sun, B. (2022). Una encuesta de análisis visual para métodos de inteligencia artificial explicable. *Computers & Graphics*, 102, 502-520. doi:

<https://doi.org/10.1016/j.cag.2021.09.002>

- Andreu-Perez, J., Emberson, L. L., Kiani, M., Filippetti, M. L., Hagra, H., & Rigato, S. (15 de Septiembre de 2021). Explainable artificial intelligence based analysis for interpreting infant fNIRS data in developmental cognitive neuroscience. *Communications Biology*, 4(1077). doi:10.1038/s42003-021-02534-y
- Bauer, K., von Zahn, M., & Hinz, O. (29 de Noviembre de 2021). Expl(AI)ned: The Impact of Explainable Artificial Intelligence on Users' Information Processing. *Safe Working Paper*, 315. doi: <http://dx.doi.org/10.2139/ssrn.3872711>
- Broniatowski, D. (abril de 2021). Psychological Foundations of Explainability and Interpretability in Artificial Intelligence. National Institute of Standards and Technology, 18-56. doi: <https://doi.org/10.6028/NIST.IR.8367>
- Byrne, R. (2019). Contrafactuales en Inteligencia Artificial Explicable (XAI): Evidencia del razonamiento humano., (págs. 6276-6282). doi:

<https://doi.org/10.24963/ijcai.2019/876>

Castellano, P. S. (22 de Noviembre de 2021). Inteligencia Artificial y valoración de la prueba: las garantías Jurídico-Constitucionales del órgano de control. *Themis Revista de derecho*(79), 283-297. doi: <https://doi.org/10.18800/themis.202101.015>

Confalonieri, R., Coba, L., Wagner, B., & Besold, T. R. (19 de Octubre de 2020). A historical perspective of explainable Artificial Intelligence. *Wires Data Mining and knowledge discovery*, 11(1). doi: <https://doi.org/10.1002/widm.1391>

Dazeley, R., Vamplew, P., Foale, C., Young, C., Aryal, S., & Cruz, F. (Octubre de 2021). Niveles de inteligencia artificial explicable para explicaciones conversacionales alineadas con humanos. *Artificial Intelligence*, 299(103525). doi: <https://doi.org/10.1016/j.artint.2021.103525>

De, T., Giri, P., Mevawala, A., Nemani, R., & Deo, A. (13 de Mayo de 2020). Explainable AI: A Hybrid Approach to Generate Human-Interpretable Explanation for Deep Learning Prediction. *Procedia Computer Science*, 168, 40-48. doi:

<https://doi.org/10.1016/j.procs.2020.02.255>

Explanations of artificial intelligence: Author proposes model that highlights evidence of fairness. (15 de Febrero de 2023). Techplore.

Fellous, J.-M., Sapiro, W., Rossi, A., Mayberg, H., & Ferrante, M. (13 de Diciembre de 2019). Explainable Artificial Intelligence for Neuroscience: Behavioral neurostimulation. *frontiers in neuroscience*. doi:10.3389/fnins.2019.01346

Ghorbani, A. (2017). Explorando la utilización de la Inteligencia Artificial Explicable para el diagnóstico de enfermedades mentales. *Journal of Neurology and Clinical Neuroscience*, 1(2), 1-2.

Gill, A., Smith, K., & S. D. (2020). Predicción de la respuesta al tratamiento con algoritmos de Inteligencia Artificial Explicable: Un estudio con pacientes con trastornos de ansiedad. *Revista de Psicología Clínica*, 17(3), 302-309.

Giuste, F., Shi, W., Zhu, Y. Z., Naren, T., Isgut, M., Sha, Y., . . . Wang, M. D. (23 de Diciembre de 2021). Explainable Artificial Intelligence Methods in Combating Pandemics: A Systematic Review. *IEEE Journal*, 23. doi:

- <https://doi.org/10.48550/arXiv.2112.12705>
- Guerrero Velázquez, T. M., & Sossa Azuela, J. H. (s.f.). New Explainability Method based on the Classification. *Computación y Sistemas*, 25(4), 719-728. doi:10.13053/CyS-25-4-4049
- Gunning, D., & Aha, D. W. (2019). DARPA's Explainable. Deep Learning and Security.
- H. HSIAO, J., HT NGAI, H., QIU, L., YANG, Y., & Chen CAO, C. (2021). Roadmap of Designing Cognitive Metrics for Explainable Artificial Intelligence (XIA). 29.
- Ha, T., June Sah, Y., & Sangwon, L. (s.f.). Examining the effects of power status of an explainable artificial intelligence system on users perceptions. *ISSN Journal*, 41(5), 946-958. doi:10.1089/144929X.2020.1846789
- Haque, A. B., Islam, A. N., & Patrick, M. (Enero de 2023). Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. *Technological Forecasting and Social Change*, 186(122120). doi: <https://doi.org/10.1016/j.techfore.2022.122120>
- Holzinger, A., Goebel, R., Fong, R., Moon, T., Muller, K., & Samek, W. (18 de julio de 2020). xxAI - Beyond. *International Workshop Held in Conjunction with ICML 2020*, 4-397.
- Iglesias, A. (4 de Abril de 2020). Artificial cognition: an emergent discipline to explain decision making in artificial neural networks. *CognitiveScience*, 16(1), 10-13.
- Kindermans, P. J., & Al-Shehivat, M. (2017). Interpretability Beyond Feature Attribution: Quantitative Testing with Concept Activation Vectors (TCAV).
- Kumar, S., Subramanian, S., & Balasubramanian, P. (2018). Predicción del estado emocional, usando Inteligencia Artificial Explicable. *Journal of Computational Science*, 24(4), 814-819.
- Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., . . . Baum, K. (Julio de 2021). ¿Qué queremos de la Inteligencia Artificial Explicable (XAI)? - Una perspectiva de las partes interesadas sobre la XAI y un modelo conceptual que guía la investigación interdisciplinaria de la XAI. *ArtificialIntelligence*, 296(103473). doi:

- <https://doi.org/10.1016/j.artint.2021.103473>
- Leben, D. (14 de Febrero de 2023). Explainable AI as evidence of fair decisions. *frontiers in psychology*, 14. doi:10.3389/fpsyg.2023.1069426
- Lesit, A. K., Klee, M., Hyun Kim, J., Rehkopf, D. H., Bordas, S. P., & Muniz-T, G. (19 de Octubre de 2022). Mapping of machine learning approaches for description, prediction, and causal inference in the social and health sciences. *ScienceAdvances*, 8(42). doi:10.1126/sciadv.abk1942
- Leuven, K. (2022). On the importance of experimental psychology for explainable artificial intelligence .
- Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (25 de Diciembre de 2020). Explainable AI: Review of Machine Learning Interpretability Methods. *entropy*. doi:10.3390/e23010018
- Loor Zambrano, B., Tello Salvador, F., Alcívar Cevallos, R., & Vaca Cardenas, L. (2020). Aproximaciones de métodos predictivos y de agrupamiento utilizados en eventos de emergencia: una literatura sistemática.
- Luo, Y., Zhou, Y., Shang, J., & Xu, X. (2020). Uso de Inteligencia Artificial Explicable para el análisis de patrones de comportamiento en la práctica clínica de psicología. *Revista de Psicología Clínica*, 18(1), 15-23.
- Markus, A., Kors, J., & Rijnbeek, P. (6 de diciembre de 2021). The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices and evaluation strategies. *SciencesDirect*, 1-11. doi: <https://doi.org/10.1016/j.jbi.2020.103655>
- Merry, M., Pat, R., & Jim, W. (9 de Diciembre de 2021). A mental models approach for defining explainable artificial intelligence. *BMC Medical Informatics and Decision Making*, 21(344). doi: <https://doi.org/10.1186/s12911-021-01703-7>
- Milda Pocevicute, M., Eliertsen, G., & Lundstrom, C. (14 de Agosto de 2020). Survey of XAI in digital pathology.
- Miller, T. (Febrero de 2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1-38. doi:<https://doi.org/10.1016/j.artint.2018.07.007>

- Mohd Nor, A. K., Rao Pedapat, S., Muhammad, M., & Leiva, V. (1 de Diciembre de 2021). Descripción general de la inteligencia artificial explicable para pronóstico y gestión de la salud de activos industriales basado en elementos de informes preferidos para revisiones sistemáticas y metanálisis. *Sensors*, 21(23). doi:
<https://doi.org/10.3390/s21238020>
- Neugebauer, S., Rippitsch, L., Sobieczky, F., & Geib, M. (2021). Explainability of AI-predictions based on psychological profiling. Elsevier, 1-10.
- Phillips, P. J., Hahn, C. A., Fontana, P. C., Yates, A. N., Greene, K., Broniatowski, D. A., & Przybocki, M. A. (Septiembre de 2021). Four Principles of Explainable Artificial. 43. doi:<https://doi.org/10.6028/NIST.IR.8312>
- Pradhan, R., Lahiri, A., Galhotra, S., & Salimi, B. (12 de Junio de 2022). Explainable AI: Foundations, Applications, Opportunities for Data Management Research. *Sigmod*. doi:[10.1145/3514221.3522564](https://doi.org/10.1145/3514221.3522564)
- Rabiul Islam, S., Eberle, W., & Ghafoor, S. K. (s.f.). Towards Quantification of Explainability in Explainable Artificial Intelligence methods .
- Rahul, R., & Reddy, C. K. (2018). Explaining the Predictions of Any Classifier.
- Ridley, M. (2022). Explainable Artificial Intelligence (XAI). *Information technology and libraries*, 41(2). doi:<https://doi.org/10.6017/ital.v41i2.14683>
- Sebio Martin, M. (2020). Inteligencia artificial y ética.
- Stepin, I., Jose M., A., Catala, A., & Pereira-Fariña, M. (2021). Una encuesta de contrastes y contrafactuales Métodos de generación de explicables para explicables Inteligencia artificial. *IEEE Acces*, 9, 11974-12001. doi:[10.1109/ACCESS.2021.3051315](https://doi.org/10.1109/ACCESS.2021.3051315)
- T. Taylor, J. E., & Taylor, G. W. (6 de Noviembre de 2020). Artificial cognition: How experimental psychology can help generate explainable artificial intelligence. *Psychonomic Bulletin & Review*, 28(454-475). doi:[10.3758/s13423-020-01825-5](https://doi.org/10.3758/s13423-020-01825-5)
- Tjoa, E., & Guan, C. (27 de Octubre de 2021). Una encuesta sobre inteligencia artificial explicable (XAI): Hacia la XAI médica. *IEEE Transactions on Neural Networks and Learning Systems*, 32(11), 4793-4813.

doi:10.1109/TNNLS.2020.3027314

Valle-Cruz, D., Lopez Chau, A., Ramos Corchado, M. A., & Muños Jimenez, V. (1 de Julio de 2022). Emotion-Aware Explainable Artificial Intelligence for Personality, Emotion, and Mood Simulation. *Computación y Sistemas*, 26, 45-57. doi:<https://doi.org/10.13053/cys-26-1-4151>

Vaquero de Miguel, G. (Julio de 2020). Aproximaciones a la explicación de decisiones algorítmicas: Inteligencia Artificial Explicable.

Vilone, G., & Longo, L. (12 de Octubre de 2020). Explainable Artificial Intelligence: a Systematic Review. 78. doi:<https://doi.org/10.48550/arXiv.2006.00093>

Vilone, G., & Longo, L. (4 de Agosto de 2021). Classification of Explainable Artificial Intelligence Methods through Their Output Formats. *Machine learning & knowledge extraction*, 615-661. doi:10.3390/make3030032

Zhang, G. (2 de Abril de 2022). Aplicaciones de la Teoría de la Atribución Social en XAI. *MDPI*, 81(1), 101. doi:<https://doi.org/10.3390/proceedings2022081101>