UNIVERSIDAD NACIONAL SAN AGUSTÍN DE AREQUIPA

FACULTAD DE INGENIERÍA DE PRODUCCIÓN Y SERVICIOS ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



TITULO

Tesis presentada por:

Bach. Huanca Sancho, Yeimy

Estephany

Para optar el Título Profesional de:

Ingeniero de Sistemas

Asesora:

Dra. Castro Gutierrez, Eveling Gloria

Arequipa - Perú 2024

RESUMEN

orem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged. It was popularised in the 1960s with the release of Letraset sheets containing Lorem Ipsum passages, and more recently with desktop publishing software like Aldus PageMaker including versions of Lorem Ipsum.

Palabras clave: Keyword1, Keyword2, Keyword3, Keyword4.

ABSTRACT

orem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged. It was popularised in the 1960s with the release of Letraset sheets containing Lorem Ipsum passages, and more recently with desktop publishing software like Aldus PageMaker including versions of Lorem Ipsum.

Palabras clave: Keyword1, Keyword2, Keyword3, Keyword4.

TABLA DE CONTENIDOS

			Pág	gina
Lı	STA D	DE TABLAS		vii
Lı	STA D	DE FIGURAS		viii
1	CAP	PITULO I. PLANTEAMIENTO DEL PROBLEMA		1
	1.1	Descripción de la realidad problemática		1
	1.2	Problema principal		3
	1.3	Objetivos		4
		1.3.1 Objetivo General		4
		1.3.2 Objetivos específicos		4
	1.4	Hipótesis de la investigación		4
	1.5	Variables e indicadores		4
		1.5.1 Variable independiente		4
		1.5.2 Variable dependiente		5
	1.6	Viabilidad de la investigación		6
		1.6.1 Viabilidad técnica		6
		1.6.2 Viabilidad operativa		6
		1.6.3 Viabilidad económica		6
	1.7	Justificación e importancia de la investigación		7
		1.7.1 Justificación		7
	1.8	Alcance		7
	1.9	Tipo y nivel de la investigación		8
		1.9.1 Tipo de investigación		8
		1.9.2 Nivel de la investigación		8
	1.10	Líneas de investigación		8
	1.11	Método y diseño de la investigación		9
		1.11.1 Diseño de la Investigación		9
	1.12	Cronograma		9
	1.13	Técnicas e instrumentos de recolección de información		9
		1 13 1 Técnicos		Q

		1.13.2 Instrumentos	10
2	CAI	PITULO II. MARCO TEÓRICO	11
	2.1	Antecedentes de la Investigación	12
	2.2	Estado del Arte	13
		2.2.1 Evolución de Técnicas	13
		2.2.2 Configuración del Nivel de Conocimiento	14
		2.2.3 Criterios de Validacion	15
		2.2.4 Situaciones en el Comportamiento de Agentes	16
		2.2.5 Objetivos para la construccion	16
		2.2.6 Mejoras a futuro	17
	2.3	Marco Conceptual	18
		2.3.1 Agentes Conversacionables	18
		2.3.2 Agentes Enseñables	19
		2.3.3 Efecto Protégé	19
3	CAI	PITULO III. DESARROLLO DE LA PROPUESTA	21
	3.1	Generalidades del diseño	22
	3.2	Esquema de la propuesta	25
	3.3	Componentes de la propuesta	27
4	CAI	PÍTULO IV. DESARROLLO DE PRUEBAS, ANÁLISIS E INTERPRETACIÓN	
	DE	RESULTADOS	29
	4.1	Presentación del Caso de Validación o Diseño del Experimento	
		4.1.1 Objetivo del Caso de Validación	29
		4.1.2 Módulos Implementados en la Arquitectura	29
		4.1.3 Diseño del Experimento	32
		4.1.4 Diseño del Experimento	33
		4.1.5 Metodología de Validación	35
		4.1.6 Aplicación del Caso de Validación	35
		4.1.7 Resultados Esperados	35
	4.2	Desarrollo de las Pruebas	35
		4.2.1 Aplicación de pruebas	38
	4.3	Interpretación de los Resultados	39
	4.4	Discusión de Resultados	39
C	ONCI	LUSIONES	41
RI	ECOI	MENDACIONES Y TRABAJOS FUTUROS	43

TABLA DE CONTENIDOS

BIBLIOGRAFÍA	45
ANEXO A	49
TITULO	49

LISTA DE TABLAS

TAB	LA]	Pá	giı	na
1.1	Cronograma de Actividades		 	 												9

LISTA DE FIGURAS

Fig	PURA Pá	gina
3.1	Nivel 0 Arquitectura para agentes enseñables	22
3.2	Nivel 0 Arquitectura para agentes enseñables	23
3.3	Nivel 1 Arquitectura para agentes enseñables a base de LLMs y técnicas para la	
	configuración dinámica de criterios de evaluación	24
4.1	Asignación de Rol Agente Evaluador de Taxonomía (Módulo D)	30
4.2	Asignación de Rol Agente Generador de Respuesta Base (Módulo E)	31
4.3	Asignación de Rol Agente de Fluidez Dialógica (Módulo G)	32
4.4	Vista de chat con el agente.	33
4.5	Seccion de personalización de las técnicas y los temas requerimiento en la sesión	34
4.6	Seccion meu detalle de sesiones y perfil	34
4.7	Diagramas pastel de los resultados.	39

CAPITULO I. PLANTEAMIENTO DEL PROBLEMA

1.1 Descripción de la realidad problemática

Según los ODS planteados por la ONU, el 4º ODS (ONU, 2024) nos habla acerca de la calidad de la educación. Algo que se debe mejorar cada vez más para llegar al objetivo del 2030. Hoy en día muchos estados tratan de implementar tecnologías de información(TI) (Sarmiento Loaiza et al., 2023; Pascagaza & Estrada, 2020), pero el mundo tiene nuevas herramientas cada día y la educación en muchos países no sabe cómo abordar todas esas actualizaciones, específicamente sobre la IA, que podrían mejorar la calidad y que si no se usan de manera correcta incluso podrían mermar su desempeño (UNESCO, Miao, & Holmes, 2024).

La UNESCO comenta que en la actualidad muchos modelos de texto generativos han sido altamente utilizados (UNESCO et al., 2024), con un incremento notable de usuarios. Uno de los ejemplares es Chat-GPT (Diego Olite, Morales Suárez, & Vidal Ledo, 2023); sin embargo, hay una lista de variaciones que también son usadas por distintas características, como , modelos open source, con requerimientos de exigencia computacional menor (García-Peñalvo, 2024), lo que permitiría un acceso más fácil a quienes tienen bajos recursos, u otros con acceso a internet en tiempo real, enfocados a la generación de imágenes, entre otros.

Hoy en día, los beneficios y dificultades traídos por la inteligencia artificial, principalmente las herramientas de texto generativas, son un tema de discusión total con respecto a la educación. Actualmente, la UNESCO se preocupa por su uso ético en la investigación y en la educación (UNESCO et al., 2024), donde se comenta que esta herramienta no vino a reemplazar a los docentes (UNESCO, 2019; Ronghuai Huang y Hui Zhang Fengchun Miao, 2021), ni a establecer malas prácticas en los estudiantes, sino a promover la investigación y el desarrollo de habilidades para los estudiantes y como una herramienta de apoyo para los docentes.

El uso ético de la herramienta ha traído conflictos sobre cómo usarla, y cómo tener buenas

prácticas al hacerlo (Ronghuai Huang y Hui Zhang Fengchun Miao, 2021). Se ve totalmente el cambio de cómo se busca la información y cómo se genera. El uso ético está relacionado con la forma en que usan la herramienta los estudiantes, que no solo es buscar y confiar ciegamente en lo que dice, sino que es apoyarse en esta y desarrollar así más habilidades de investigación, compartir ideas y reforzar el conocimiento (UNESCO et al., 2024). Uno de los pioneros en contemplar esta arma de doble filo que es la IA generativa en el contexto de la educación es Beijing (UNESCO, 2019), que en el año 2019 ya publicó el Consenso de Beijing sobre la inteligencia artificial y la educación, donde da recomendaciones sobre cómo se debería utilizar la inteligencia artificial para mejorar la educación y, sobre todo, mejorar la calidad de la educación acorde al ODS 4. Entre las recomendaciones se plantea que la IA debe ser una herramienta de ayuda tanto para el estudiante como para el docente, y no debería reemplazar las actividades que realiza el docente. Debería, por ejemplo, utilizarse para la gestión de la impartición de la educación, para el aprendizaje y la evaluación, o con agentes sensibles.

El otro lado de la moneda es cómo la IA generativa afecta el aprendizaje de los estudiantes. En la publicación de la UNESCO se describen casos reales del efecto de la IA generativa en el contexto de la educación donde el propósito de aprendizaje hacia los estudiantes no se cumplieron por el uso indebido de la herramienta (Ronghuai Huang y Hui Zhang Fengchun Miao, 2021). La UNESCO también expresa la preocupación por el uso de la herramienta de manera consciente en menores de 13 años y la necesidad de supervisión, pero acepta su uso por mayores a partir de los 13 años (UNESCO et al., 2024).

Entonces, el problema radica en encontrar su aplicación de manera eficiente, como lo indica (UNESCO et al., 2024), hallar formas creativas por parte de los docentes, asignaciones donde se necesite de creatividad, algo que no podría replicar la herramienta. El propósito es brindar a los estudiantes oportunidades para desarrollar capacidades cognitivas y razonamiento lógico independiente, haciendo que el aprendizaje o la investigación sea más efectivos. Incluso brindar entrenamiento para la adquisición a ritmo propio de habilidades fundamentales. Dentro de lo propuesto por UNESCO (UNESCO et al., 2024), algunos ejemplos de co-diseño, entre ellos uno resaltable, es el uso de la guía generativa como entrenador para la adquisición autodidacta de habilidades básicas en idiomas. El presente estudio propone una dinámica que desarrolle habilidades en los estudiantes por medio del "efecto Protégé". (Chase, Chin, Oppezzo, & Schwartz, 2009)

La inteligencia artificial ha superado muchas barreras, entre ellas el poder analizar lenguaje natural (Moreira et al., 2021). Sin embargo, eso se podía lograr en cierta parte con NLPs. Actualmente, las nuevas herramientas reciben instrucciones a partir de este lenguaje ambiguo (Fernández Genaro, 2023). Sus ventajas son innegables, lo que permite una variedad de temas y múltiples tipos de conversación con el uso de roles o mejorando el prompt, haciendo que se mejore el desempeño del GPT con su interacción con el usuario.

El campo de los agentes enseñables es un ámbito en la educación que busca, con difer-

entes técnicas, que el estudiante tenga una dinámica más atractiva que mejore su aprendizaje (H. J. S. Lee et al., 2024; K. J. Lee, Chauhan, Goh, Nilsen, & Law, 2021).

En una revisión literaria del campo de la inteligencia artificial generativa, exactamente con agentes enseñables, no se encontró una guía clara de cómo implementarlos en el ámbito de la educación. Si bien existen muchos tipos (Morales-Rodríguez & Domínguez-Martínez, 2011) dependiendo de hasta qué características abarca, sigue un patrón de muestra, entrada y procesamiento con herramientas que pueden variar

1.2 Problema principal

Los agentes enseñables han adquirido un valor significativo debido al notable impacto de la inteligencia artificial (H. J. S. Lee et al., 2024; UNESCO et al., 2024; UNESCO, 2019). Se han documentado casos donde estos agentes se emplean para que estudiantes universitarios refuercen sus conocimientos (UNESCO et al., 2024), uno de los enfoques que se le da dentr de la oducacion a estos agentes es el metodo LBT(Learn by teaching) que se basa en la enseñanza para mejorar el aprendizaje para potenciar el efecto "efecto Protégé" (Chase et al., 2009; Love, Law, Cohen, & Kuli ´c, 2022). Además, existe una investigación denominada Curiosity Notebook (Law et al., 2020; K. J. Lee et al., 2021) que ha dado lugar a una plataforma web donde los estudiantes pueden cargar información para enseñar a la plataforma, lo que contribuye a mejorar su comprensión.

Sin embargo, estas modalidades de enseñanza se limitan principalmente a la entrada de información mediante resaltado de documentos o mediante la introducción de texto tipo chat por parte del estudiante. Aunque válidas, la UNESCO plantea la posibilidad de mejorar esta interacción. Por ejemplo, mediante el uso de voz, permitiendo al estudiante explicar verbalmente al agente lo que desea enseñar, o mediante el uso de dibujos que el estudiante puede utilizar para comunicarse con el agente. (UNESCO, 2019; Ronghuai Huang y Hui Zhang Fengchun Miao, 2021; UNESCO et al., 2024)

Asimismo, se plantea la necesidad de establecer mecanismos de retroalimentación más efectivos para los estudiantes. En los artículos referenciados (K. J. Lee et al., 2021; Datta et al., 2021a), la retroalimentación es limitada, a menudo se reduce a mensajes predefinidos, o en algunos casos, no se proporciona retroalimentación alguna.

Para abordar esta problemática, es esencial construir un<mark>a arquitectura robusta y flexible</mark> que sirva como marco de referencia para la creación de agentes enseñables.

Además, la arquitectura propuesta debe incorporar principios de diseño centrados en el usuario, asegurando que los agentes enseñables sean accesibles, intuitivos y efectivos para estudiantes y educadores como lo menciona (K. J. Lee et al., 2021). Esto implica considerar aspectos como la facilidad de configuración y personalización, la transparencia en el proceso de toma de decisiones del agente y la capacidad de proporcionar retroalimentación relevante y contextualizada en tiempo real.

1.3 Objetivos

1.3.1 Objetivo General

Diseñar e implementar una instancia de arquitectura modular para agentes enseñables basada en un enfoque low code, que permita la personalización de criterios de evaluación y conocimiento en módulos específicos mediante el uso de LLMs y técnicas de Recuperación Asistida por Generación (RAG).

1.3.2 Objetivos específicos

- Investigar trabajos relacionados con agentes enseñables, agentes LBT o simuladores de estudiantes enfocados en el aprendizaje por enseñanza en estudiantes.
- Evaluar las arquitecturas existentes para la creación de agentes enseñables con enfoque en ("efecto Protégé" o LBT).
- Argumentar la modificación de la arquitectura propuesta, fundamentando la implementación de un entorno low code para la personalización de criterios de evaluacion mediante LLMs y técnicas RAG.
- Validar la arquitectura propuesta mediante una instancia que demuestre su efectividad en la personalización de parámetros, implementada en un entorno de configuración low code, y evaluar la experiencia del usuario a través de un cuestionario de usabilidad.

1.4 Hipótesis de la investigación

La implementación de una arquitectura flexible para agentes enseñables, basada en un enfoque low code con LLMs y técnicas RAG, aumentará la adaptabilidad del agente a diferentes criterios de evaluación, facilitando configuraciones específicas sin necesidad de programación avanzada.

1.5 Variables e indicadores

1.5.1 Variable independiente

Arquitectura flexible con enfoque low code usando LLMs y RAG

Indicadores	Índices
 Nivel de personalización de criterios de evaluación. Modularidad: Nivel de separación de componentes, facilitando actualizaciones y mantenimiento. Tasa de uso de módulos de personalización. 	 Índice 1: Lograr un ajuste de criterios sin necesidad de codificación. Índice 2: Cantidad de módulos separados y su facilidad de actualización. Índice 3: Utilización de módulos en al menos el 80% de los casos.

1.5.2 Variable dependiente

Adaptabilidad del agente a criterios de evaluación personalizados

Indicadores	Índices
 Satisfacción del usuario en cuestionarios de usabilidad. Flexibilidad para adaptar a distintos criterios configurables. 	 Índice 1: Al menos 85% de satisfacción en adaptabilidad y facilidad de uso. Índice 2: Evaluación de al menos 4/5 en pruebas de distintos temas y métodos.

1.6 Viabilidad de la investigación

1.6.1 Viabilidad técnica

La viabilidad técnica de la arquitectura propuesta se centra en facilitar la creación de Agentes Enseñables mediante el uso de inteligencia artificial, específicamente modelos de lenguaje grande (LLM). Estos modelos serán fundamentales como intermediarios para la conexión e intercomunicación entre diversas tecnologías, optimizando así resultados y procesos en conjunto.

1.6.1.1 Hardware y Software Necesarios:

- Computadoras con capacidades de procesamiento adecuadas para ejecutar modelos de LLM y otras tecnologías complementarias.
- Herramientas de desarrollo de software que incluyan entornos de programación para integración de LLM, como Python con bibliotecas como Transformers o Hugging Face.
- Plataformas de desarrollo de aplicaciones que permitan la integración fluida de componentes conversacionales y de interacción, como API de inteligencia artificial y servicios en la nube.
- Herramientas de desarrollo de interfaces de usuario que faciliten la interacción mediante texto y voz, mejorando así la accesibilidad y la experiencia del usuario.
- Modelos LLM entrenados para tareas específicas de enseñanza y aprendizaje, configurados para optimizar la interacción conversacional con los estudiantes.
- Software de gestión de bases de datos para almacenar y analizar datos recopilados durante las interacciones con los agentes enseñables.

1.6.2 Viabilidad operativa

El proyecto cuenta con el firme compromiso de los investigadores, quienes se dedicarán a la ejecución de todas las etapas del estudio. Además, el asesor de tesis ha demostrado un compromiso significativo para guiar y supervisar el desarrollo de la investigación, asegurando el cumplimiento de los objetivos y la calidad del trabajo. Los investigadores también han asegurado la disponibilidad de los recursos necesarios,

1.6.3 Viabilidad económica

En el caso de la viabilidad económica, está investigación es financiada por los tesistas Para llevar a cabo este proyecto, los investigadores asumen la responsabilidad de cubrir todos los costos necesarios para completar las diversas etapas propuestas y se contará con la infraestructura y equipamiento del CiTeSoft. Además, está abierto a participar en concursos y solicitudes de financiamiento externo para asegurar los recursos financieros necesarios

1.7 Justificación e importancia de la investigación

1.7.1 Justificación

El efecto protege ha sido objeto de estudio durante varios años, destacando desde el trabajo pionero de (Chase et al., 2009), que mostró el gran impacto de los agentes enseñables en el fomento de este efecto. Basándose en este estudio, se desarrollaron investigaciones adicionales, entre las más relevantes se encuentra el estudio Curiosity Notebook tanto del 2020 y 2021 (Law et al., 2020; K. J. Lee et al., 2021) que creó una infraestructura web para configurar y utilizar agentes enseñables, destacando funciones críticas que estos agentes deben cumplir. Sin embargo, hasta ese punto, no se contemplaban las nuevas tendencias de la inteligencia artificial, como los modelos de lenguaje grande (LLM). A partir de los trabajos de Curiosity Notebook (K. J. Lee et al., 2021), se han desarrollado nuevos enfoques en los cuales se utilizan agentes enseñables para diversas tareas, como aprender programación, integrando LLM (H. J. S. Lee et al., 2024) en el desarrollo de estos agentes. Estos estudios presentan nuevos objetivos de diseño y destacan la importancia de ciertas propiedades para que los agentes enseñables funcionen correctamente. No obstante, no se presentan muchas arquitecturas en general que puedan tomarse como referencia , hasta la propuesta en (Paul, 2015). También se han encontrado un estudio que simulan ser estudiantes para proporcionar retroalimentación a profesores (Datta et al., 2021a) en ese estudio se propone una arquitectura para agentes conversacionales, que presenta una arquitectura para estos agentes, pero no ofrece un marco de trabajo integral para la creación de agentes enseñables en distintos contextos. Es necesario contar con una arquitectura de referencia que cumpla con los objetivos de diseño y las propiedades descritas en los trabajos relacionados, si bien los estudios previos han demostrado el valor de los agentes enseñables en el fomento del efecto protege, la integración de modelos avanzados de inteligencia artificial, como los modelos de lenguaje grande (LLMs), aún no ha sido explorada en profundidad en arquitecturas accesibles y configurables para agentes enseñables. Las metodologías low code presentan una oportunidad única para reducir las barreras de implementación y aumentar la adopción en escenarios educativos variados. Al proporcionar una arquitectura de referencia modular y flexible que integre LLMs y RAG para la personalización dinámica, este trabajo no solo aborda las limitaciones actuales en cuanto a la adaptabilidad y escalabilidad, sino que también propone una solución práctica alineada con las necesidades tecnológicas y pedagógicas actuales. Esto permite la creación de agentes enseñables que se adaptan a múltiples contextos, ofreciendo una base para la investigación y aplicación de inteligencia artificial en la educación."

1.8 Alcance

• Construcción de una arquitectura para Agentes Enseñables (AEs): Desarrollo de una plataforma educativa que implemente AEs.

- Evaluación del impacto de los AEs en el contexto educativo: Enfocado específicamente en estudiantes de educación secundaria.
- Análisis de la interacción con los AEs: Estudio de cómo los AEs influyen en la motivación, el esfuerzo, y el rendimiento académico de los estudiantes.

• Examen de variables específicas:

- Motivación de los estudiantes.
- Tiempo dedicado a las actividades de aprendizaje.
- Rendimiento académico en materias específicas.

• Uso de métodos mixtos para recolección de datos:

- Cuestionarios para obtener datos cuantitativos.
- Entrevistas para recolectar información cualitativa.
- Análisis de desempeño académico para evaluar el impacto en el aprendizaje.
- Observación de las interacciones en la plataforma educativa: Estudio del impacto de la interacción con los AEs en los estudiantes.

1.9 Tipo y nivel de la investigación

1.9.1 Tipo de investigación

La investigación se clasifica como una Investigación Aplicada. Considerando que la investigación busca aplicar los conocimientos adquiridos para poder desarrollar una arquitectura para agentes enseñables, esto con un enfoque cuantitativo.

1.9.2 Nivel de la investigación

La investigación se enmarca en un enfoque cuantitativo de nivel aplicativo, dado que se busca desarrollar una arquitectura para agentes enseñables que sea altamente configurable, para poder ajustarse a las diversas necesidades.

1.10 Líneas de investigación

La presente investigación se enmarca en la línea de investigación "Arquitectura, Redes y Sistemas Distribuidos", ya que busca desarrollar una arquitectura para la creación de agentes enseñables

1.11 Método y diseño de la investigación

1.11.1 Diseño de la Investigación

El diseño de la investigación es experimental, ya que se probará la arquitectura propuesta en entornos reales y controlados para evaluar el impacto en el aprendizaje. Este diseño experimental permitirá establecer relaciones causales entre la utilización de los agentes enseñables y los resultados de aprendizaje de los estudiantes.

1.12 Cronograma

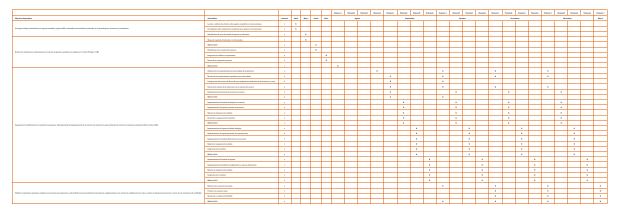


Table 1.1: Cronograma de Actividades

1.13 Técnicas e instrumentos de recolección de información

1.13.1 Técnicas

- Implementación con Modelos de Lenguaje Grande (LLMs) Se utilizarán LLMs como parte integral de la instancia para proporcionar respuestas contextuales y adaptativas. Estos modelos se configurarán para interactuar con los estudiantes, interpretar las reglas ingresadas por los encargados por la configuración y se encargaran del manejo del conocimiento del agente entre otras tareas.
- Recuperación Asistida por Generación (RAG) Se aplicará la técnica de Recuperación Asistida por Generación (RAG) para mejorar la capacidad del agente enseñables de acceder a los criterios de evaluación para cada modulo y para la generación y actualización del conocimiento del agente.
- Evaluación de Usabilidad: Para validar la efectividad de la arquitectura propuesta, se implementará un cuestionario de usabilidad que medirá la experiencia del usuario. Este cuestionario evaluará aspectos como la satisfacción del usuario, la facilidad de uso y la

efectividad del agente en la personalización del aprendizaje. La retroalimentación obtenida será crucial para realizar ajustes y mejorar la arquitectura.

1.13.2 Instrumentos

Para implementar las técnicas descritas en la sección anterior, se utilizarán los siguientes instrumentos:

• Implementación con Modelos de Lenguaje Grande (LLMs)

- Herramientas de desarrollo: Se emplearán Jupyter Notebook y Visual Studio
 Code para el desarrollo y ajuste de los modelos.
- Bibliotecas y frameworks: Se utilizará Hugging Face Transformers para implementar y trabajar con LLMs, junto con la biblioteca Gemini para la interacción del agente.
- Plataforma de almacenamiento: ChromaDB se empleará para almacenar y gestionar el conocimiento del agente.

• Recuperación Asistida por Generación (RAG)

 Contexto de decisión: Los criterios de evaluación se subirán mediante archivos que proporcionarán contexto para las decisiones del agente, evitando la necesidad de algoritmos de búsqueda por el momento.

• Evaluación de Usabilidad

- Cuestionario de Usabilidad: Se diseñará un cuestionario utilizando Google Forms para medir la satisfacción del usuario y la efectividad del agente.
- Herramientas de análisis: Se empleará Python, junto con Jupyter Notebook o Visual Studio Code, para analizar los resultados del cuestionario y extraer conclusiones sobre la experiencia del usuario.

CAPITULO

CAPITULO II. MARCO TEÓRICO

2.1 Antecedentes de la Investigación

En el artículo de (Chase et al., 2009) se presenta uno de los primeros estudios en proponer agentes enseñables (teachable agents). En este sistema, los estudiantes interactúan con un Agente Enseñable (AE) que puede simular razonamiento según las instrucciones recibidas. Dos estudios destacan el efecto protegido: los estudiantes muestran mayor esfuerzo al aprender para sus AEs que para sí mismos. En un estudio con estudiantes de octavo grado, aquellos que creían que enseñaban a sus AEs dedicaron más tiempo a actividades de aprendizaje y mostraron un mayor rendimiento, especialmente entre los estudiantes con menor desempeño. Un segundo estudio con estudiantes de quinto grado reveló que los niños que interactuaron con AEs también dedicaron más tiempo al aprendizaje activo. Estos estudiantes, al atribuir estados mentales y responsabilidad a sus AEs, demostraron una mayor disposición para reconocer errores y reflexionar sobre las causas de los fallos. Esto sugiere que los AEs no solo motivan el aprendizaje al tratarlos como seres sociales, sino que también ayudan a los estudiantes a mejorar su conocimiento sin dañar su autoestima cuando cometen errores.

En el trabajo de (Valdivia-Cisneros, Vidal, & Castro-Gutierrez, 2023) se presenta una arquitectura multimodal y sin contacto para sistemas de realidad virtual enfocados en la rehabilitación de miembros superiores. El artículo muestra una metodología para desarrollar una propuesta de arquitectura que consta de seis pasos: 1) Identificación de los requisitos de calidad del sistema, 2) Identificación de los requisitos arquitectónicamente significativos, 3) Diseño de los componentes de la arquitectura, 4) Clasificación de los componentes, 5) Validación de las decisiones de diseño, y 6) Análisis y evaluación de la arquitectura de software. Esta metodología adapta el trabajo de (Parisaca et al., 2023) y se centra en sistemas de rehabilitación. Si bien no es un tema relacionado con agentes enseñables, la metodología clara es un aporte significativo.

El artículo de (Parisaca et al., 2023) presenta un nuevo enfoque en la arquitectura de software para sistemas de rehabilitación virtual, enfocándose en requisitos de calidad del sistema y atributos arquitectónicamente significativos. Utiliza una metodología detallada que incluye la identificación de requisitos de calidad, componentes arquitectónicos estáticos y dinámicos, diseño de componentes, documentación de arquitectura, y validación mediante técnicas cuantitativas como simulaciones y análisis de usuario. La evaluación se realizó utilizando el Cuestionario de Experiencia del Usuario (UEQ) (Schrepp, 2015), destacando la eficiencia y fiabilidad del sistema con dispositivos como Leap Motion y guante háptico. Este enfoque proporciona una contribución significativa al construir arquitecturas dinámicas que permiten la integración flexible de tecnologías de rehabilitación virtual, mejorando la experiencia del usuario sin comprometer la funcionalidad y la eficacia terapéutica.

Como partes más destacables del libro (Gorton, 2006) se puede observar que en los Capítulos 4 a 7, que constituyen el núcleo técnico de la obra. El Capítulo 4 introduce tecnologías de middleware comúnmente empleadas en soluciones de aplicaciones, mientras que el Capítulo 5 detalla un proceso iterativo de arquitectura de software en tres etapas, delineando las tareas esenciales del

arquitecto. Por otro lado, el Capítulo 6 se enfoca en la documentación arquitectónica, destacando las nuevas notaciones en UML 2.0 que son cruciales para este propósito. Finalmente, el Capítulo 7 aplica estas tecnologías de middleware para abordar los requisitos de atributos de calidad en un estudio de caso específico, utilizando la plantilla de documentación presentada previamente. Esta discusión detallada y el uso de plantillas en UML 2.0 proporcionan una base sólida para la validación continua de la arquitectura, reforzada por el enfoque iterativo descrito en el Capítulo 5.

El artículo de (UNESCO et al., 2024) sobre inteligencia artificial generativa (IAGen) en educación destaca preocupaciones actuales y propone ideas no probadas pero potencialmente beneficiosas. Se discute cómo la IAGen podría automatizar tareas cognitivas y generar contenido educativo, aunque también plantea desafíos como seguridad, privacidad de datos y riesgo de sustituir la intervención humana en actividades creativas. Las políticas actuales son limitadas, con pocos países regulando específicamente el uso de IAGen en educación. Se recomienda un enfoque ético y colaborativo para su implementación, involucrando a docentes, estudiantes e investigadores en su construcción. Una idea promovida es utilizar IAGen como entrenador personalizado para la adquisición autodidacta de habilidades en idiomas y artes, mejorando habilidades lingüísticas mediante interacciones personalizadas. Sin embargo, se advierte sobre riesgos como la dependencia excesiva de la tecnología y la reducción del pensamiento crítico si no se implementan estrategias pedagógicas adecuadas.

2.2 Estado del Arte

2.2.1 Evolución de Técnicas

Para simular el comportamiento de un estudiante sin conocimientos previos a través de un proceso de evolución, se pueden utilizar Modelos de Lenguaje de Gran Escala (LLM) para crear conversaciones más realistas. Sin embargo, antes de la relevancia de los LLM, era necesario entrenar modelos para clasificar las entradas del usuario y responder adecuadamente.

Un ejemplo de esto se encuentra en el trabajo de (Datta et al., 2021b). En este estudio, se utilizaron grabaciones de conversaciones entre alumnos y docentes, enfocadas en la comprensión de la geometría como factor de escala. Estas grabaciones, provenientes de un proyecto de investigación previo, fueron clasificadas manualmente en una de las cuatro categorías de la IQA, una currícula adaptada para docentes. Este proceso permitió entrenar un modelo BERT que luego clasificaría las entradas de los usuarios, en este caso, docentes próximos a egresar y ejercer.

El mismo año,2021, se presentó (K. J. Lee et al., 2021), un trabajo que propone una plataforma web para simular un agente enseñable y configurable. En esta plataforma, se podía cargar información y simular conversaciones donde el usuario enseñaba al agente a clasificar elementos como animales o tipos de pinturas. Sin embargo, el artículo no especificaba las técnicas, modelos o tecnologías utilizadas para generar estas conversaciones.

En contraste, para el año 2024, se han publicado artículos que trabajan con LLM de manera más avanzada. Aunque estos modelos no requieren obligatoriamente un nuevo entrenamiento, en los estudios (H. J. S. Lee et al., 2024) y (Lu & Wang, 2024) se aplicaron métodos específicos para optimizar su desempeño. En el primer caso, se utilizó el ajuste fino (fine-tuning) del modelo, mientras que en el segundo se estableció un conjunto de reglas iniciales que determinaban los puntos de conocimiento y las áreas de incomprensión. Esto permitió generar distintos tipos de estudiantes mediante la variación de estos parámetros.

En otro trabajo, (Ma, Shen, Koedinger, & Wu, 2024) propone un sistema de aprendizaje interactivo llamado HypoCompass, que utiliza modelos de lenguaje (LLM) para ayudar a los estudiantes a aprender programación, específicamente en la depuración de código. Se aprovechan los LLM para generar código con errores y proporcionar explicaciones detalladas de dichos errores.

La evolución de la IA y la generación de texto ha avanzado significativamente en pocos años. Mientras que en 2021 se dependía de la clasificación manual y el entrenamiento de modelos específicos como BERT, en 2024 se aprovechan las capacidades preentrenadas de los LLM, complementándolas con técnicas como el ajuste fino y la definición de reglas iniciales para optimizar su funcionalidad en contextos educativos.

2.2.2 Configuración del Nivel de Conocimiento

Un punto importante que comparten todos los artículos es la configuración del nivel de conocimiento que debe tener el agente, idea que abarcan de distintas formas. Este nivel debería ser configurable para poder ajustar la implementación a diversos entornos.

Por ejemplo, en el trabajo de Notebook Curiosity (K. J. Lee et al., 2021), se implementa en dos ambientes: uno con estudiantes de primaria y otro con universitarios. El nivel de conocimiento tiene que ser configurable en ambos casos. De manera similar, en el artículo (Datta et al., 2021a), se menciona que el nivel de conocimiento debe ser ajustable. En este artículo, se utiliza una comparación semántica entre el input del usuario y la base de conocimientos para manipular el nivel de conocimiento del agente. La comparación semántica es una técnica efectiva, ya que permite determinar con un umbral qué tan "exacto" es el conocimiento del agente.

En el caso de (Lu & Wang, 2024), se usaron reglas predefinidas para establecer la distribución de conocimientos. Se creó un conjunto de 45 estudiantes generativos, distribuidos en diferentes niveles de conocimiento de la siguiente manera:a)10 estudiantes con dificultades (menos reglas dominadas y más confusiones).b)30 estudiantes promedio. c)5 estudiantes avanzados (más reglas dominadas y menos confusiones).Las reglas se utilizan como parámetros para establecer el conocimiento: a)Reglas Dominadas: Los estudiantes tienen mayor probabilidad de responder correctamente si la respuesta correcta es una regla que dominan. b)Reglas Confusas: Tienen alta probabilidad de equivocarse si la respuesta correcta es una regla sobre la que están confundidos. c)Reglas Desconocidas: Pueden acertar a veces, especialmente si no hay reglas confusas presentes en las opciones.

En el artículo (H. J. S. Lee et al., 2024), se utilizó el ajuste fino (fine-tuning), la configuración del rendimiento de la red neuronal de memoria dinámica y la modificación directa de la base de datos del conocimiento inicial, que se corregía a medida que el modelo aprendía.

Por otro lado, (Ma et al., 2024) utiliza modelos de lenguaje (LLM) para simular el conocimiento de estudiantes novatos en informática. Se emplean técnicas de prompt para generar código con errores, lo que permite a los usuarios corregir y depurar dicho código.

La configuración del nivel de conocimiento del agente es un aspecto crucial en todos los trabajos revisados. Aunque los métodos varían, todos apuntan a crear agentes más adaptables y precisos. Desde la comparación semántica y la configuración manual hasta el uso de reglas predefinidas y el ajuste fino de modelos, cada técnica tiene sus ventajas y se adapta a diferentes contextos educativos. Esta evolución muestra cómo los avances en IA y la capacidad de ajustar el nivel de conocimiento del agente pueden mejorar significativamente la efectividad de las herramientas educativas.

2.2.3 Criterios de Validacion

La validación de las implementaciones en los artículos revisados comparte formas similares de evaluación mediante experimentos comparativos entre grupos.

En "Notebook Curiosity" (K. J. Lee et al., 2021), se compararon las características y comportamientos de dos grupos (C1 y C2) en términos de sus tasas de clics en los botones y la frecuencia de revisión del cuaderno del agente. Se ajustaron modelos lineales gaussianos para explorar las relaciones entre las puntuaciones de motivación intrínseca y los resultados en cuestionarios posteriores al estudio.

Por otro lado, en el artículo (Datta et al., 2021a), la validación no incluyó una comparación con otra muestra. En cambio, se probó con un grupo de profesores que estaban por egresar, simulando estudiantes de 4° y 5° año de secundaria.

En (Lu & Wang, 2024), se compararon las respuestas de 45 estudiantes generativos con las de 100 estudiantes universitarios reales al mismo conjunto de preguntas. Se utilizó una medición de correlación para evaluar si las respuestas generadas por los estudiantes generativos eran similares a las de los estudiantes reales, validando que los perfiles generativos capturaban adecuadamente las diferencias en el conocimiento y las confusiones de los estudiantes.

Para el caso de (H. J. S. Lee et al., 2024) se realizaron pruebas individuales con un conjunto de 40 personas principiantes en algoritmos para evaluar al agente resultante, la validación se basó en la meta-cognición, midiendo la satisfacción del usuario y monitorizando la densidad y calidad de los mensajes en la conversación. Este enfoque permitió evaluar cómo los cambios en el modelo afectaban la interacción y el aprendizaje del usuario.

La forma más común de validación es la comparación entre dos grupos y el análisis pre y post sesión. Aunque algunos trabajos no implementaron una validación formal, otros utilizaron criterios específicos y detallados para asegurar la efectividad de sus modelos.

2.2.4 Situaciones en el Comportamiento de Agentes

Existen diversas situaciones en las que los agentes presentan particularidades en su funcionamiento y comportamiento. En el artículo (K. J. Lee et al., 2021), cada respuesta del agente se asocia con una emoción, utilizando etiquetas emocionales como "curioso" para controlar movimientos y gestos, esto en una implementación con un robot. Además, la implementación del robot incluye sensores en la cabeza y los pies, permitiendo que los niños den palmadas en la cabeza del robot cuando responde correctamente a una pregunta. Este enfoque mejora la interacción humano-robot y refuerza positivamente el aprendizaje del estudiante.

Así mismo cuando un estudiante enseña al agente, el nuevo conocimiento se almacena en un cuaderno que sirve como reporte detallado de lo enseñado. Esto contrasta con el enfoque de (Datta et al., 2021b), donde se clasifica el tipo de preguntas generadas por docentes en formación. Se busca que formulen preguntas de "indagación y exploración", "expositivas y de indicaciones", y "de procedimiento y fácticas", según las categorías de IQA (Análisis de la Calidad de las Preguntas).

Por otro lado en el estudio (Lu & Wang, 2024), se identifican preguntas fáciles y difíciles para comparar resultados entre estudiantes reales y generativos. Esto permite establecer reglas que guían el comportamiento de los agentes, ajustando su desempeño según estas evaluaciones.

La investigación (H. J. S. Lee et al., 2024) identifica una taxonomía de mensajes que un estudiante podría usar al enseñar. Las categorías incluyen Instrucción (fijación, dominante, ánimo), Incitación (encontrar desafíos, reticente, comprobación, sugerente, pidiendo ayuda), y Declaración (comprensión, elaboración, creación de sentido, aceptación, comentario).

En la comparación de estos enfoques, el enfoque basado en emociones proporciona una dimensión adicional al simple hecho de clasificar preguntas, permitiendo al agente influir en el tono de las respuestas y ejecutar acciones específicas para mejorar la experiencia del usuario. Por otro lado un tambien se estudiaron agentes donde la conversaciones se clasifica por taxonomias esto para poder tener una conversacion mas adecuada.

2.2.5 Objetivos para la construccion

Los objetivos de diseño son fundamentales, ya que definen el enfoque de cada implementación de los agentes. En el trabajo de (K. J. Lee et al., 2021), se establecen claramente los siguientes objetivos:

- Características del Agente: Permitir la modulación de características del agente para estudiar su impacto en el aprendizaje, ofreciendo retroalimentación y formulando preguntas auténticas.
- Cuantificación de Estrategias de Enseñanza: Facilitar la elección de actividades de enseñanza por parte de los estudiantes para analizar cómo estas estrategias mejoran el aprendizaje.

- Tarea y Material de Aprendizaje Escalable: Apoyar tareas de aprendizaje escalables para diferentes edades, comparando la efectividad de los agentes en distintos grupos etarios.
- Enseñanza Coordinada en Grupo: Permitir la enseñanza individual, en parejas o grupos, asegurando igual acceso a las oportunidades de enseñanza.
- Formas Flexibles de Agentes: Ofrecer agentes en texto, voz o físicos, para evaluar cómo diferentes formatos afectan el aprendizaje y la enseñanza.

En contraste, el artículo de (Datta et al., 2021b) no define textualmente los objetivos, pero se pueden inferir: configurabilidad, capacidad de recolección de información, y evaluación de la calidad de las preguntas (IQA). Por otro lado en el trabajo (H. J. S. Lee et al., 2024), los objetivos se centran en mejorar la eficacia del cambio de modo, aumentar la construcción de conocimiento, y validar la reconfigurabilidad, persistencia y adaptabilidad del sistema. Para el trabajo de (Lu & Wang, 2024), los objetivos incluyen simular el conocimiento estudiantil, evaluar preguntas efectivamente, y mejorar el diseño de preguntas para comprender cómo estudiantes generativos y reales perciben su dificultad y claridad.

El trabajo de (Ma et al., 2024) tiene como objetivo ayudar a los estudiantes a desarrollar habilidades en la construcción de hipótesis, tanto completas como precisas. Esto implica generar un conjunto exhaustivo de casos de prueba que cubran los posibles errores (LO1) y proporcionar una explicación precisa de por qué el programa es incorrecto, dada la información de los casos de prueba fallidos (LO2).

Se observa que la configurabilidad es un objetivo compartido entre los artículos analizados. En el primer artículo, se destaca la necesidad de contar con formas flexibles para los agentes, un aspecto que también se menciona en el segundo artículo dentro de las secciones dedicadas a los trabajos futuros al igual que en el tercer artículo. Tambien se destaca la configuracion de la dificultad y la persistencia de los datos.

2.2.6 Mejoras a futuro

En el artículo de (K. J. Lee et al., 2021), se recomienda mejorar la configuración del agente para que pueda operar con múltiples instancias y fomentar la colaboración entre grupos. Además, se sugiere usar modelos de aprendizaje automático para simular conversaciones, permitiendo al agente utilizar el conocimiento enseñado y mostrar sus errores para fomentar nuevas enseñanzas.

El artículo de (Datta et al., 2021b) propone expandir la base de conocimientos con nuevos datos clasificados, aunque esto es menos necesario gracias a los modelos de lenguaje (LLM) que pueden inferir e interpretar el lenguaje humano. También se aconseja utilizar modelos y técnicas de procesamiento de lenguaje natural (NLP) más actuales, ya que herramientas como BERT pueden ser insuficientes para ciertos proyectos.

En (H. J. S. Lee et al., 2024), se sugiere monitorear la dinámica entre la percepción de los estudiantes, la ganancia de aprendizaje y la metacognición en ambientes más grandes, debido

a las variaciones observadas entre los participantes por sus diferentes niveles de conocimiento previo y estilos de enseñanza.

Para (Lu & Wang, 2024), se identifica la viabilidad con instructores no especializados como un área de mejora, señalando la necesidad de una preconfiguración adecuada y un análisis del tiempo requerido para que los instructores utilicen el agente de manera práctica.

- Investigar estrategias de delegación de tareas más personalizadas que satisfagan las necesidades de aprendizaje individuales.
- Modularizar los componentes de HypoCompass para que sean más accesibles y adaptables a diferentes contextos educativos.
- Realizar estudios de implementación en el aula más amplios para evaluar la eficacia de HypoCompass en entornos del mundo real.

Estos puntos resaltan la necesidad de construir una arquitectura flexible y adaptable a diversos contextos educativos. Integrar nuevas tecnologías para facilitar conversaciones más naturales permitirá que los agentes se adapten mejor a las dinámicas de los grupos de estudiantes. Además, la arquitectura debe permitir que los agentes utilicen las enseñanzas adquiridas durante la sesión y se ajusten a las variaciones individuales de los estudiantes, mejorando así la personalización del aprendizaje y ofreciendo retroalimentación instantánea basada en el progreso del estudiante.

2.3 Marco Conceptual

2.3.1 Agentes Conversacionables

- El término "Agentes Conversables" según (Wu et al., 2023) se refiere a entidades dentro del marco AutoGen que son capaces de participar activamente en conversaciones con otros agentes, humanos, o una combinación de ambos. Estos agentes están diseñados para recibir, reaccionar y responder a mensajes de manera autónoma, lo que facilita la creación de aplicaciones complejas mediante un paradigma de programación centrado en la conversación.
- La definición según el artículo (Li, Hammoud, Itani, Khizbullin, & Ghanem, 2023) se refiere a programas informáticos diseñados para imitar la interacción humana mediante el uso del lenguaje natural. Estos agentes son capaces de comunicarse con humanos y otros agentes de inteligencia artificial, facilitando la resolución de tareas complejas a través de la conversación y la cooperación mutua.
- Según (Kowatsch et al., 2021) indica que son programas informáticos diseñados para imitar la interacción humana mediante el uso del lenguaje natural. Estos agentes permiten la

comunicación tanto entre agentes de inteligencia artificial como con humanos, facilitando la colaboración y la resolución de tareas complejas a través de conversaciones interactivas y cooperativas.

2.3.2 Agentes Enseñables

- Según (Chhibber & Law, 2021), son sistemas de inteligencia artificial que pueden aprender y adaptarse mediante la interacción con usuarios humanos no expertos, a través de una interfaz conversacional. Estos agentes se entrenan utilizando retroalimentación humana directa durante conversaciones, lo que les permite mejorar su rendimiento en tareas específicas, como la clasificación de textos. Este enfoque aprovecha las interacciones naturales para ajustar y optimizar modelos de aprendizaje automático, facilitando la personalización y la mejora continua basada en las aportaciones de los usuarios.
- En el artículo (Love, Law, Cohen, & Kulić, 2022), se describen agentes de inteligencia artificial diseñados para aprender mediante la enseñanza que les proporcionan los estudiantes. En este contexto, el agente adopta el rol de un aprendiz o novato, mientras que el estudiante actúa como su maestro. Esta dinámica es significativa porque puede activar el efecto "Protégé" en el estudiante-maestro, un fenómeno pedagógico que aumenta la participación en la tarea de enseñanza y mejora los resultados cognitivos del estudiante. En el artículo se muestra una interacción a través de diálogos en lenguaje natural, lo que fomenta la comprensión y el compromiso afectivo del estudiante.
- De la misma manera, (Silvervarg, Wolf, Blair, Haake, & Gulz, 2021) define estos entornos de software como lugares donde los estudiantes enseñan a agentes, en lugar de solo recibir instrucción. Estos agentes mejoran el aprendizaje de los estudiantes, fomentando habilidades de razonamiento, comprensión conceptual y procesamiento metacognitivo. El "efecto Protégé" motiva a los estudiantes a dedicar más esfuerzo y tiempo a las actividades de aprendizaje, ya que ven al agente como un aprendiz real, compartiendo con él la responsabilidad de los errores. Esto facilita que los estudiantes reconozcan y discutan sus errores, abriéndose más a la retroalimentación crítica y constructiva.es, abriéndose más a la retroalimentación crítica y constructiva.

2.3.3 Efecto Protégé

• Segun (Rubin, Noy, & Musen, 2007) se refiere a la relación mentor-protegido, que se define como una relación simbiótica y mutuamente beneficiosa entre dos individuos, en la que el mentor influye significativamente en la carrera del protegido. Esta relación es especialmente relevante en el contexto de los adultos y sus trayectorias profesionales, aunque hay poca información sobre aquellos que no siguen un camino ocupacional específico o no buscan activamente un mentor.

- El "efecto Protégé" segun (Okurame, 2008) se refiere a la influencia que la combinación de géneros entre el protégé y el mentor tiene sobre el compromiso organizacional. En un estudio realizado en Nigeria, se encontró que los hombres mostraban mayor compromiso organizacional cuando sus mentores también eran hombres, mientras que las mujeres mostraban mayor compromiso cuando sus mentoras eran mujeres. La combinación de protégé masculino con mentora femenina resultó en el menor nivel de compromiso organizacional. Este efecto destaca la importancia del género en las relaciones de mentoría y su impacto en la actitud de los empleados hacia la organización.
- El articulo de (Chase et al., 2009) denota que se refiere a la motivación adicional que los estudiantes experimentan al enseñar a un agente digital, como un Tamagotchi o un agente enseñable (TA). Este efecto surge cuando los estudiantes sienten una responsabilidad hacia el agente, lo cual les lleva a dedicar más esfuerzo y tiempo en su aprendizaje para asegurar que el agente "aprenda" y se desarrolle adecuadamente. Esta responsabilidad y el sentido de cuidado hacia el agente motivan a los estudiantes a mejorar su propio entendimiento del contenido para poder enseñarlo mejor.

SAPITULO

CAPITULO III. DESARROLLO DE LA PROPUESTA

3.1 Generalidades del diseño

La arquitectura propuesta para un agente enseñable, centrada en la dinámica "enseñar para aprender" o *Learning by Teaching* en inglés, describe un sistema en el que un estudiante humano (usuario) enseña a un agente virtual con el fin de mejorar y reforzar sus propios conocimientos a través de este proceso de enseñanza, fortaleciendo el efecto *Protege*. La interacción se estructura en varias capas y módulos que permiten la clasificación, evaluación y contextualización de la información proporcionada por el usuario. A continuación, se presenta una descripción formal del diagrama en la version inicial y su modificacion:

Arquitectura nivel 0 (Paul, 2015): El usuario interactúa con el sistema enseñando al agente, y la información pasa por un proceso de estandarización. Luego, la información es clasificada, evaluada y procesada en diferentes niveles, desde la taxonomía hasta la generación de una respuesta adecuada y contextualizada, para que el agente responda correctamente al usuario según el conocimiento preestablecido y aprendido. Finalmente, se genera un reporte que se entrega al encargado, quien puede supervisar la sesión y aportar datos adicionales para futuras interacciones.

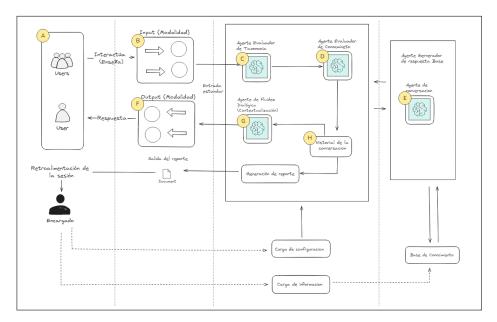


Figure 3.1: Nivel 0 Arquitectura para agentes enseñables . Fuente: (Paul, 2015)

Nota: La arquitectura toma en cuenta los diferentes módulos necesarios dependiendo a la revisión hecha .

Tras una revisión de la literatura y las técnicas actuales en el ámbito de los agentes enseñables, se identificó la necesidad de incluir un módulo adicional: el agente administrador de memoria-conocimiento (J). Aunque en nuestro artículo previo se propuso una arquitectura más simplificada, la inclusión de este módulo es una mejora que responde a varios factores críticos observados en estudios recientes. Este módulo está diseñado para optimizar el rendimiento adaptativo del agente mediante una estructura modular que facilita la configuración y personalización del nivel de conocimiento del agente de acuerdo con diferentes contextos educativos, como los descritos en la Sección 2.2.2.

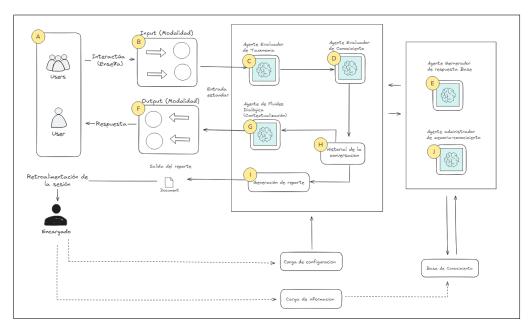


Figure 3.2: Nivel 0 Arquitectura para agentes enseñables. Fuente: (Paul, 2015)

Nota: La arquitectura ah sido modificada para integrar un modulo necesario para la administración del conocimiento del agente que representa las diferentes maneras de implementar la lógica del conocimiento del agente.

La necesidad de modularidad en la arquitectura se justifica por las conclusiones de estudios como los de Notebook Curiosity (K. J. Lee et al., 2021) y (Datta et al., 2021a), que resaltan la importancia de ajustar el conocimiento del agente en función del nivel de los estudiantes, ya sea de primaria o universitarios. En particular, el uso de configuraciones dinámicas permite que el agente ofrezca respuestas que se alineen con las expectativas y conocimientos previos de los usuarios en diversos contextos.

Además, en el caso de (Lu & Wang, 2024), se observa el uso de reglas predefinidas para simular distintos niveles de conocimiento, lo que permite una distribución flexible de los parámetros de dominio en el agente. El módulo de administrador de memoria-conocimiento que proponemos facilitará la integración de reglas y configuraciones personalizadas, como reglas dominadas y confundidas, directamente en la arquitectura, lo que resulta en un agente que puede ajustar su conocimiento y respuesta en tiempo real según el contexto de enseñanza y el perfil del usuario.

Asimismo, técnicas avanzadas como el ajuste fino (fine-tuning) y la manipulación directa de la base de datos de conocimiento, descritas en (H. J. S. Lee et al., 2024), nos han inspirado a

agregar capacidades de auto-optimización en este módulo. Así, el agente podrá ajustar su nivel de conocimiento y precisión de respuesta con cada interacción, lo que asegura una experiencia de usuario más precisa y adaptable.

En resumen, la inclusión del agente administrador de memoria-conocimiento fortalece nuestra propuesta inicial sin contradecir el trabajo anterior, ya que responde a la necesidad de adaptar la configuración del conocimiento del agente a distintos escenarios educativos. Con esta ampliación, la arquitectura no solo es más robusta y modular, sino que también se alinea con las prácticas y técnicas modernas, permitiendo que el agente se configure y responda de manera eficaz en función de las necesidades específicas de cada contexto.

Arquitectura nivel 1: Los estudiantes reales enseñan contenido a este agente, que, a su vez, evalúa y responde según criterios de conocimiento y diálogo predefinidos. La arquitectura se compone de varios módulos interconectados que aseguran una respuesta adaptada y una fluidez dialógica en la interacción. Además, el sistema permite configuraciones avanzadas y personalización de criterios de evaluación por parte del educador o encargado, quien supervisa el proceso y recibe reportes finales sobre la interacción.

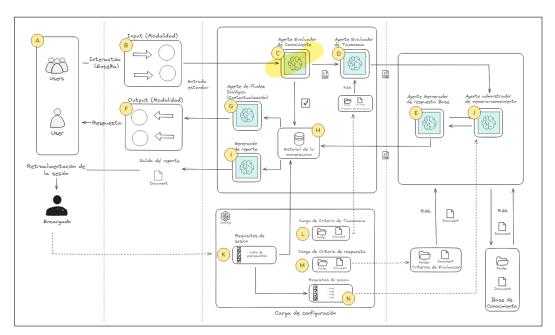


Figure 3.3: Nivel 1 Arquitectura para agentes enseñables a base de LLMs y técnicas para la configuración dinámica de criterios de evaluación.

Fuente: Elaboración propia

Nota: La arquitectura ah sido adaptada y especificada para el caso y uso de LLMs y técnicas RAG para la carga dinámica de criterios de evaluación personalizables y la administración de la memoria-conocimiento del agente.

3.2 Esquema de la propuesta

Usuarios y Roles

- Estudiantes reales: Representados por los usuarios que interactúan con el agente enseñable mediante texto.
- Encargado o Educador: Responsable de definir los criterios de evaluación y supervisar el proceso de enseñanza.

Módulos Principales de la Arquitectura

- Módulo C Agente Evaluador de Conocimiento: Evalúa el conocimiento transmitido en cada interacción para determinar su validez y cumplimiento de los objetivos.
- Módulo D Agente Evaluador de Taxonomía: Aplica criterios específicos de taxonomía, precargados o personalizados por el educador, para categorizar y evaluar la interacción.
- Módulo J Agente Administrador de Memoria y Conocimiento: Actualiza y organiza el conocimiento del agente, almacenándolo en una memoria estructurada para consultas futuras.
- Módulo E Agente Generador de Respuesta Base: Genera una respuesta inicial basándose en el contexto de la primera interacción.
- Módulo H Historial de Conversación: Registra cada interacción, configuraciones de la sesión, y los temas/prerrequisitos que deben abordarse.
- **Módulo G Agente de Fluidez Dialógica**: Contextualiza las respuestas, integrando el historial de la conversación para mantener la coherencia en el diálogo.
- Módulo F Convertidor de Modalidad de Respuesta: Convierte las respuestas generadas en formatos adecuados (texto, audio, etc.) según las preferencias del usuario o del educador.
- **Módulo K Carga de Requisitos de Sesión**: Permite definir los requisitos y prerrequisitos de la sesión para ser actualizados en el historial.
- Módulo L Carga de Criterios de Taxonomía: Contiene los criterios de taxonomía necesarios para la evaluación de las interacciones.
- **Módulo M Criterios de Respuesta**: Define los criterios que el agente generador de respuesta debe tener en cuenta al formular una respuesta base.

• **Módulo N - Configuración de Conocimiento Inicial**: Permite definir un conocimiento inicial para el agente, que el módulo J transformará en documentos estructurados para su base de conocimiento.

3.3 Componentes de la propuesta

Inicio de la Interacción

- Los estudiantes ingresan al sistema y realizan su interacción de enseñanza en formato de texto.
- El contenido proporcionado pasa al **Módulo C** (Agente Evaluador de Conocimiento), que analiza si el conocimiento entregado cumple con el objetivo de aprendizaje o necesita ajustes.

Evaluación y Clasificación del Conocimiento

- Si el contenido es adecuado, se envía en formato JSON al Módulo D (Agente Evaluador de Taxonomía), que categoriza la enseñanza según criterios de taxonomía previamente definidos o personalizados por el encargado.
- El Módulo D evalúa el contenido y envía los resultados en formato JSON, incluyendo los campos de clasificación requeridos, al Módulo J (Administrador de Memoria y Conocimiento).

Actualización de la Memoria de Conocimiento

- El **Módulo J** almacena y organiza el conocimiento del agente en una estructura de documentos específica, utilizando técnicas de *prompting* para asegurar la coherencia en respuestas futuras.
- Este contexto de memoria se activa cuando el agente necesita generar una respuesta o recordar interacciones pasadas, apoyando su capacidad de retener conocimiento acumulado.

Generación de la Respuesta Base

- Con el contexto de la memoria en uso, la interacción pasa al Módulo E (Generador de Respuesta Base), que produce una respuesta inicial que se envía en formato JSON al Módulo H (Historial de Conversación).
- Este historial registra todas las configuraciones, temas y prerrequisitos de la sesión, ayudando al sistema a recordar el progreso y el estado de los temas cubiertos.

Contextualización de la Respuesta

• El **Módulo G** (Fluidez Dialógica) recibe tanto el historial como la respuesta base para ajustar la respuesta en función de las interacciones previas, logrando una mayor coherencia en el diálogo.

• La respuesta contextualizada se envía al **Módulo F** (Convertidor de Modalidad de Respuesta), que la formatea de acuerdo con el medio solicitado por el estudiante o el educador (texto, audio, etc.).

Entrega de la Respuesta y Generación de Reportes

La respuesta final es presentada al usuario en el formato deseado, y la información acumulada durante la sesión se utiliza para generar reportes en el Módulo A (Generador de Reportes), destinados al educador.

Actualización Constante y Ajustes en Tiempo Real

- Los prerrequisitos de la sesión, almacenados en el Módulo K, son actualizados continuamente por el Módulo C para reflejar el progreso del estudiante.
- Los criterios de evaluación, taxonomía y respuesta, definidos en los módulos L, M y N,
 permiten que la arquitectura mantenga un conocimiento inicial adaptable, fortaleciendo
 la capacidad del agente para responder como un "estudiante" con niveles de conocimiento
 específicos.



CAPÍTULO IV. DESARROLLO DE PRUEBAS, ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

En este capítulo, se aborda

4.1 Presentación del Caso de Validación o Diseño del Experimento

El presente trabajo propone un plan de validación para la arquitectura flexible de agentes enseñables basada en un enfoque low code, con técnicas RAG y LLMs. Este caso de validación está orientado a evaluar la efectividad de los agentes en la enseñanza de locuciones locales arequipeñas, utilizando estudiantes de nivel secundario y profesores encargados de las sesiones como grupo de prueba. A continuación, se detalla el diseño del experimento y los módulos implementados para esta instancia.

4.1.1 Objetivo del Caso de Validación

El objetivo principal es demostrar que la arquitectura propuesta permite configurar criterios específicos de evaluación y enseñanza sin necesidad de programación avanzada, adaptándose a diferentes contextos educativos.

4.1.2 Módulos Implementados en la Arquitectura

Se han integrado los siguientes módulos para esta instancia del experimento:

4.1.2.1 Agente 1: Evaluador de Taxonomía (Módulo D)

El **Agente Evaluador de Taxonomía** tiene la tarea de evaluar las interacciones de los estudiantes basándose en criterios predefinidos o personalizados de taxonomía. Este agente categoriza las respuestas y evalúa la validez del conocimiento transmitido en cada sesión. Utiliza una taxonomía que puede ser ajustada según los requerimientos del educador.

Figure 4.1: Asignación de Rol Agente Evaluador de Taxonomía (Módulo D)

El agente analiza interacciones lingüísticas, identifica brechas de conocimiento y genera una lista priorizada de temas para profundizar el aprendizaje. Fuente: Elaboración propia

Tareas principales:

- Evaluar las interacciones basadas en criterios taxonómicos.
- Determinar la validez del conocimiento y la alineación con los objetivos de aprendizaje.
- Generar un informe sobre la calidad del contenido transmitido.

4.1.2.2 Agente 2: Generador de Respuesta Base (Módulo E)

El **Agente Generador de Respuesta Base** genera las respuestas iniciales basadas en el contexto de la conversación. Este agente utiliza el historial de interacción para construir respuestas que sean coherentes con el tema de la sesión, priorizando la claridad y la relevancia.

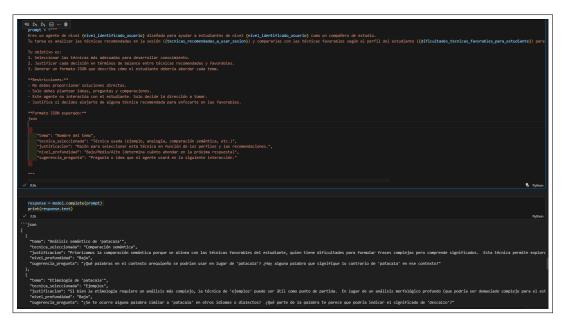


Figure 4.2: Asignación de Rol Agente Generador de Respuesta Base (Módulo E)

El agente analiza técnicas recomendadas y favorables según el perfil del estudiante, utiliza variables como las dificultades técnicas y las descripciones de los temas, y genera un plan estructurado en formato JSON para guiar el aprendizaje. Fuente: Elaboración propia

Tareas principales:

- Generar respuestas contextuales para iniciar la interacción.
- Mantener coherencia con los temas previos tratados en la conversación.
- Proporcionar respuestas claras y estructuradas para facilitar la comprensión.

4.1.2.3 Agente 3: Agente de Fluidez Dialógica (Módulo G)

El **Agente de Fluidez Dialógica** asegura que las respuestas generadas sean coherentes y fluidas, tomando en cuenta el historial de la conversación. Este agente tiene la función de mantener el flujo natural de la interacción, mejorando la experiencia del usuario mediante respuestas que siguen de manera lógica a las anteriores.

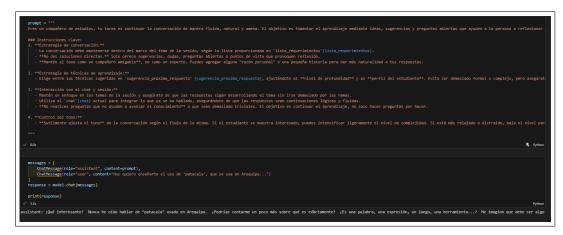


Figure 4.3: Asignación de Rol Agente de Fluidez Dialógica (Módulo G).

El agente actúa como un compañero de estudio que mantiene la conversación fluida, natural y amena. Analiza los temas y las técnicas de aprendizaje según el perfil del estudiante, evitando soluciones directas y fomentando la reflexión mediante preguntas abiertas y sugerencias. Utiliza las conversaciones previas para garantizar coherencia y adaptarse al tono adecuado para cada momento. Fuente: Elaboración propia

Tareas principales:

- Mantener la coherencia en la conversación a través del contexto.
- Integrar los elementos previos de la interacción para asegurar la fluidez.
- Adaptar las respuestas para que se ajusten al ritmo y tono de la conversación.

4.1.3 Diseño del Experimento

El diseño del experimento está compuesto por las siguientes etapas:

- Preparación de la Sesión: Los profesores definen los temas y requisitorios para la enseñanza de locuciones locales arequipeñas mediante el módulo K - Carga de Requisitos de Sesión. El sistema inicializa los criterios de evaluación con el módulo L - Carga de Criterios de Taxonomía.
- 2. Interacción con los Estudiantes: Los agentes interactúan con los estudiantes simulando un proceso de enseñanza. Durante la sesión, el módulo H Historial de Conversación registra cada interacción, mientras que el módulo J Administrador de Memoria y Conocimiento organiza y actualiza el conocimiento adquirido.
- 3. Evaluación de Resultados: El módulo D Evaluador de Taxonomía analiza el cumplimiento de los objetivos pedagógicos definidos. La satisfacción de los usuarios y la flexibilidad del sistema son evaluadas mediante cuestionarios y observaciones cualitativas, considerando los siguientes índices:

- Nivel de satisfacción del usuario (índice mínimo: 70%).
- Flexibilidad para adaptar criterios configurables (calificación mínima: 4/5).

4.1.4 Diseño del Experimento

Para la interfaz se tiene un desendioso simple que permita al usuario personalizar su sesión, visualizar las sesiones , su cuenta y la función de chat con el compañero virtual



Figure 4.4: Vista de chat con el agente. Funcion de Chat con el agente directamente. Fuente: Elaboración propia

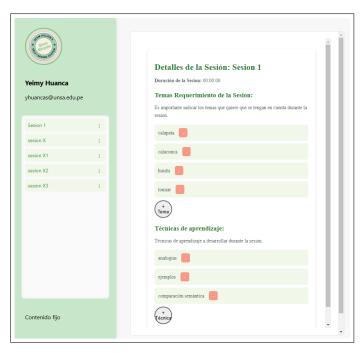


Figure 4.5: Seccion de personalización de las técnicas y los temas requerimiento en la sesión. Sección que personaliza los temas que se desarrollaran durante la sesión y las técnicas recomendadas que se quiere se desarrollen con el estudiante. Fuente: Elaboración propia



Figure 4.6: Seccion meu detalle de sesiones y perfil.

El sistema presenta una sección que permite el inicio de sesión y el manejo de sesiones. Fuente: Elaboración propia

4.1.5 Metodología de Validación

La validación se realizará utilizando métodos mixtos:

- Datos Cuantitativos: Los cuestionarios de satisfacción medirán la percepción de los usuarios sobre la facilidad de uso y adaptabilidad del sistema.
- Datos Cualitativos: Entrevistas con los profesores responsables identificarán áreas de mejora en la interacción y evaluación del agente.

4.1.6 Aplicación del Caso de Validación

En esta fase inicial, la arquitectura será evaluada exclusivamente en el contexto de locuciones locales arequipeñas. Se espera expandir este enfoque a otros dominios en el futuro.

4.1.7 Resultados Esperados

Se anticipa que los agentes diseñados:

- Logren un alto nivel de adaptabilidad a criterios específicos de evaluación sin codificación adicional.
- Fomenten un aprendizaje reflexivo y colaborativo en estudiantes y profesores.
- Alcancen una satisfacción del usuario superior al 70%.

Los resultados de esta validación guiarán futuras iteraciones y mejoras en la arquitectura propuesta.

4.2 Desarrollo de las Pruebas

La siguiente sección presenta la forma de evaluación llevada a cabo para el caso de validación , pera demostrar una alineación con el comportamiento esperado

Los sistemas de software representan su calidad en la delineación con sus objetivos de comportamiento, por lo tanto, es necesario realizar pruebas de usabilidad @@@@. A continuación se presenta una evaluación de la primera versión del Agente Enseñable para poder demostrar y determinar la usabilidad del Agente. Al aplicar las pruebas de usabilidad se busca medir la capacidad del agente para poder cumplir con los objetivos de interacción con el usuario.

El propósito de implementar una evaluación basada en la System Usability Scale (SUS) para nuestro sistema de agentes enseñables es garantizar que los usuarios (tanto estudiantes como profesores) encuentren la herramienta intuitiva, funcional y adecuada para su uso continuo. Dado que el sistema busca proporcionar una experiencia personalizada que facilite la conversación

y la generación de nuevos conocimientos, su éxito depende en gran medida de la aceptación y comodidad del usuario.

Aspectos clave de la evaluación:

- Interactividad y comodidad del usuario: El sistema permite al estudiante interactuar con un
 agente personalizado para explorar temas, resolver dudas y generar nuevos conocimientos.
 La percepción de comodidad durante estas sesiones es crucial para cumplir con el propósito
 del sistema, que es fomentar un aprendizaje fluido y sin barreras.
- Configuración flexible: Aunque el sistema ofrece configuraciones por defecto para agilizar
 el inicio, también permite personalizar las sesiones según las preferencias del usuario.
 Evaluar cómo los usuarios perciben esta flexibilidad ayudará a determinar si la herramienta
 satisface tanto a quienes prefieren una experiencia guiada como a quienes buscan mayor
 control.
- Facilidad de uso: Al ser una herramienta destinada a estudiantes y profesores, es esencial
 que las funcionalidades sean intuitivas y fáciles de aprender, incluso sin asistencia técnica
 especializada. La evaluación SUS nos permitirá identificar áreas donde la interfaz o la
 experiencia del usuario necesiten mejoras.

4.2.0.1 Preguntas incluidas en la evaluación

- 1. Conversaría con mi compañero constantemente para adquirir nuevos conocimientos.
- Encuentro el proceso de usar el sitio para conversar con mi compañero innecesariamente difícil o complicado.
- 3. Considero que el sitio es entendible y fácil de usar.
- Para poder usar el sitio para conversar con mi compañero, necesito que alguien técnico me ayude.
- 5. Pienso que varias funcionalidades están bien integradas en el sitio y son útiles.
- 6. Pensé que las funcionalidades no son amigables o fáciles de usar.
- 7. Creo que la mayoría de las personas podría aprender a usar esta herramienta con facilidad.
- 8. Encuentro el sistema demasiado difícil de usar.
- 9. Me siento cómodo usando el sitio para conversar con mi compañero.
- 10. Necesité aprender demasiadas cosas antes de poder usar este sistema.

El cuestionario fue diseñado para medir aspectos específicos relacionados con la usabilidad del sistema, organizados en cuatro dimensiones principales derivadas de la metodología SUS:

4.2.0.2 Facilidad de uso (Ease of Use)

Evalúa la percepción de los usuarios sobre la claridad, accesibilidad y facilidad de interacción con el sistema.

Preguntas relacionadas:

- P3: Considero que el sitio es entendible y fácil de usar.
- **P7:** Yo pienso que la mayoría de personas podría aprender a usar esta herramienta con facilidad.
- **P9:** Me siento cómodo usando el sitio para conversar con mi compañero.

4.2.0.3 Complejidad percibida (Complexity Perceived)

Indaga si el sistema presenta barreras, confusión o dificultades en su uso.

Preguntas relacionadas:

- P2: Encuentro el proceso de usar el sitio innecesariamente difícil o complicado.
- P6: Pensé que las funcionalidades no son amigables o fáciles de usar.
- P8: Encuentro el sistema demasiado difícil de usar.

4.2.0.4 Aprendibilidad (Learnability)

Determina si el sistema requiere conocimientos previos o una curva de aprendizaje significativa para su uso.

Preguntas relacionadas:

- **P4:** Para poder usar el sitio, necesito que alguien técnico me ayude.
- P10: Necesité aprender demasiadas cosas antes de poder usar este sistema.

4.2.0.5 Satisfacción y utilidad percibida (Satisfaction and Usefulness)

Mide la percepción del usuario sobre si el sistema cumple con sus objetivos y resulta satisfactorio en su uso.

Preguntas relacionadas:

- P1: Conversaría con mi compañero constantemente para adquirir nuevos conocimientos.
- P5: Pienso que varias funcionalidades están bien integradas en el sitio y son útiles.

Con lo anterior se esperan un rango de resultados para su analisis:

- Facilidad de uso: Puntajes altos (4-5), indicando que el sistema es comprensible y fácil de utilizar.
- Complejidad percibida: Puntajes bajos (1-2), ya que el sistema debería ser intuitivo. Sin embargo, podrían señalarse algunas barreras relacionadas con la fluidez en las conversaciones.
- **Aprendibilidad:** Puntajes bajos (1-2), reflejando que no se requiere aprendizaje previo para utilizar el sistema.
- **Satisfacción y utilidad:** Puntajes moderados (3-4), indicando que las funcionalidades son útiles, aunque pueden no ser completamente satisfactorias en términos de fluidez.

4.2.1 Aplicación de pruebas

Para la evaluación de la primera versión del sistema, se llevó a cabo un proceso de pruebas con un grupo representativo de 15 estudiantes de educación, seleccionados con el propósito de obtener perspectivas tanto desde su formación profesional como desde su experiencia como estudiantes. Este enfoque permitió garantizar un análisis integral que incluyera puntos de vista pedagógicos y de usabilidad.

4.2.1.1 Procedimiento

• Selección de participantes Se contactó a estudiantes del área de educación interesados en participar en la evaluación del sistema. Los criterios de selección se centraron en su disponibilidad y en su familiaridad con herramientas tecnológicas básicas, asegurando un nivel homogéneo entre los participantes.(a) Los participantes fueron seleccionados exclusivamente entre estudiantes matriculados en programas de formación en educación. Este criterio fue establecido debido a que su experiencia como futuros docentes les permite ofrecer una perspectiva dual: la de usuario final del sistema y la de profesional en formación capaz de implementar herramientas tecnológicas similares en entornos educativos. Esta doble visión enriquecerá la retroalimentación sobre la usabilidad del sistema desde un enfoque educativo y práctico. (b)Se consideró importante que los participantes contaran con habilidades mínimas en el uso de dispositivos tecnológicos y herramientas digitales, dado que este conocimiento previo garantiza que la evaluación se centre en la usabilidad y accesibilidad del sistema, evitando que las posibles barreras tecnológicas de los usuarios interfieran con los resultados de la evaluación. (c)Los participantes deberían haber utilizado sistemas educativos o plataformas tecnológicas en contextos académicos previos. Este criterio fue definido con el fin de que los estudiantes pudieran ofrecer una comparación más rica y detallada sobre la funcionalidad del sistema, basándose en experiencias previas con

herramientas tecnológicas educativas, lo que les permitirá identificar tanto las fortalezas como las áreas de mejora del sistema evaluado.

- Inducción al sistema Cada estudiante fue evaluado de manera individual. Antes de iniciar la prueba, se les brindó una breve explicación sobre los objetivos del sistema y una demostración básica sobre cómo utilizarlo. Esta introducción buscaba reducir posibles barreras iniciales de uso.
- **Prueba práctica** Los participantes interactuaron directamente con el sistema, realizando tareas comunes que representaran escenarios reales de uso. Esto incluyó navegar por la interfaz, utilizar funcionalidades principales y resolver ejercicios que reflejaran el propósito del sistema.
- Rellenado del cuestionario Tras la interacción con el sistema, cada estudiante completó el cuestionario basado en la metodología System Usability Scale (SUS). Este cuestionario evaluó dimensiones clave como facilidad de uso, complejidad percibida, aprendibilidad y satisfacción con el sistema.

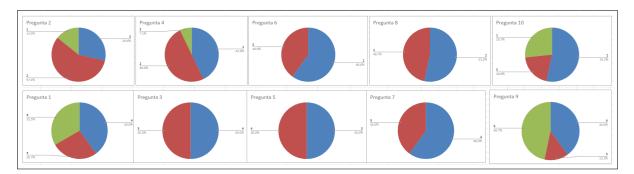


Figure 4.7: Diagramas pastel de los resultados.

Los gráficos indican la frecuencia de las respuestas por pregunta , las cuales tienen tanto pares como impares tienen un significado sobre el sistema. *Fuente: Elaboración propia*

4.3 Interpretación de los Resultados

4.4 Discusión de Resultados

CONCLUSIONES

PRIMERO:

It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editors now use Lorem Ipsum as their default model text, and a search for 'lorem ipsum' will uncover many web sites still in their infancy. Various versions have evolved over the years, sometimes by accident, sometimes on purpose (injected humour and the like).

SEGUNDO:

It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editors now use Lorem Ipsum as their default model text, and a search for 'lorem ipsum' will uncover many web sites still in their infancy. Various versions have evolved over the years, sometimes by accident, sometimes on purpose (injected humour and the like).

TERCERO:

It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editors now use Lorem Ipsum as their default model text, and a search for 'lorem ipsum' will uncover many web sites still in their infancy. Various versions have evolved over the years, sometimes by accident, sometimes on purpose (injected humour and the like).

CUARTO:

It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editors now use Lorem Ipsum as their default model text, and a search for 'lorem ipsum' will uncover many web sites still in their infancy. Various versions have evolved over the years, sometimes by accident, sometimes

on purpose (injected humour and the like).

QUINTO:

It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editors now use Lorem Ipsum as their default model text, and a search for 'lorem ipsum' will uncover many web sites still in their infancy. Various versions have evolved over the years, sometimes by accident, sometimes on purpose (injected humour and the like).

RECOMENDACIONES Y TRABAJOS FUTUROS

- 1. S
- 2. a
- 3. b

BIBLIOGRAFÍA

- Chase, C. C., Chin, D. B., Oppezzo, M. A., & Schwartz, D. L. (2009, 9). Teachable agents and the protégé effect: Increasing the effort towards learning. *Journal of Science Education and Technology*, 18, 334-352. Retrieved from https://link.springer.com/article/10.1007/s10956-009-9180-4 doi: 10.1007/S10956-009-9180-4/METRICS
- Chhibber, N., & Law, E. (2021). Towards teachable conversational agents. *arXiv* preprint *arXiv*:2102.10387.
- Datta, D., Phillips, M., Bywater, J. P., Chiu, J., Watson, G. S., Barnes, L., & Brown, D. (2021a). Virtual pre-service teacher assessment and feedback via conversational agents. Retrieved from https://aclanthology.org/2021.bea-1.20 (Define una arquitectura para "Agentes Conversables")
- Datta, D., Phillips, M., Bywater, J. P., Chiu, J., Watson, G. S., Barnes, L., & Brown, D. (2021b,
 4). Virtual pre-service teacher assessment and feedback via conversational agents. In
 E. L.-Q. R. L. C. M. N. P. I. Y. H. Z. T. B. Jill Horbach Andrea Kochmar (Ed.), (p. 185-198).
 Association for Computational Linguistics. Retrieved from https://aclanthology.org/2021.bea-1.20
- Diego Olite, F. M., Morales Suárez, I. d. R., & Vidal Ledo, M. J. (2023). Chat gpt: origen, evolución, retos e impactos en la educación. *Educación Médica Superior*, 37(2).
- Fernández Genaro, J. L. (2023). Modelos de machine learning para la ciencia de datos.
- García-Peñalvo, F. J. (2024, abr.). Inteligencia artificial generativa y educación: Un análisis desde múltiples perspectivas. *Education in the Knowledge Society (EKS)*, 25, e31942. Retrieved from https://revistas.usal.es/tres/index.php/eks/article/view/31942 doi: 10.14201/eks.31942
- Gorton, I. (2006). Essential software architecture. Springer Science & Business Media.
- Kowatsch, T., Schachner, T., Harperink, S., Barata, F., Dittler, U., Xiao, G., ... others (2021). Conversational agents as mediating social actors in chronic disease management involving health care professionals, patients, and family members: multisite single-arm feasibility study. *Journal of medical Internet research*, 23(2), e25060.
- Law, E., Ravari, P. B., Chhibber, N., Kulic, D., Lin, S., Pantasdo, K. D., ... Dillen, N. (2020, 4). Curiosity notebook: A platform for learning by teaching conversational agents. *Conference on Human Factors in Computing Systems Proceedings*. Retrieved from https://dl.acm.org/doi/10.1145/3334480.3382783 doi: 10.1145/3334480.3382783

- Lee, H. J. S., Kim, H. S. J., Jin, H., Lee, S., Shin, H., & Kim, J. (2024, 5). Teach ai how to code: Using large language models as teachable agents for programming education. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1-28. Retrieved from https://dl.acm.org/doi/10.1145/3613904.3642349 doi: 10.1145/3613904.3642349
- Lee, K. J., Chauhan, A., Goh, J., Nilsen, E., & Law, E. (2021, 10). Curiosity notebook: The design of a research platform for learning by teaching. *Proceedings of the ACM on Human-Computer Interaction*, 5. Retrieved from https://dl.acm.org/doi/10.1145/3479538 doi: 10.1145/3479538
- Li, G., Hammoud, H., Itani, H., Khizbullin, D., & Ghanem, B. (2023). Camel: Communicative agents for mind exploration of large language model society. Advances in Neural Information Processing Systems, 36, 51991–52008.
- Love, R., Law, E., Cohen, P. R., & Kulić, D. (2022). Natural language communication with a teachable agent. *arXiv preprint arXiv:2203.09016*.
- Love, R., Law, E., Cohen, P. R., & Kuli´c, D. K. (2022, 3). Natural language communication with a teachable agent. Retrieved from https://arxiv.org/abs/2203.09016v1
- Lu, X., & Wang, X. (2024). Generative students: Using llm-simulated student profiles to support question item evaluation. In *Proceedings of the eleventh acm conference on learning@ scale* (pp. 16–27).
- Ma, Q., Shen, H., Koedinger, K., & Wu, S. T. (2024). How to teach programming in the ai era? using llms as a teachable agent for debugging. In *International conference on artificial intelligence in education* (pp. 265–279).
- Morales-Rodríguez, M., & Domínguez-Martínez, J. (2011). Agentes conversacionales como un sistema de diálogo. *Memorias del V Encuentro de Investigadores del ITCM*.
- Moreira, D., Cruz, I., Gonzalez, K., Quirumbay, A., Magallan, C., Guarda, T., ... Castillo, C. (2021). Análisis del estado actual de procesamiento de lenguaje natural. *Revista Ibérica de Sistemas e Tecnologias de Informação*(E42), 126–136.
- Okurame, D. E. (2008). Effects of protégé-mentor gender mix on organisational commitment. Journal of International Women's Studies, 9(3), 1–12.
- ONU. (2024). Objetivos y metas de desarrollo sostenible desarrollo sostenible. https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/. (Consultado el 23 de mayo de 2024)
- Parisaca, E. E. S., Muñoz, S. J. M., Duarte, E. V., Gutierrez, E. G. C., Peraltilla, A. Y. C., & Pérez, S. A. (2023, 12). Dynamic software architecture design for virtual rehabilitation system for manual motor dexterity. *International Journal of Advanced Computer Science and Applications*, 14, 78-85. Retrieved from www.ijacsa.thesai.org doi: 10.14569/IJACSA.2023.0140210
- Pascagaza, E. F., & Estrada, L. C. C. (2020). Modernización de la educación virtual y su incidencia en el contexto de las tecnologías de la información y la comunicación (tic). *Revista Academia*

- y virtualidad, 13(2), 103-116.
- Paul, Y. (2015). Arquitectura agentes. All you need to know to apply the UEQ successfully in your project, 50–52.
- Ronghuai Huang y Hui Zhang Fengchun Miao, W. H. (2021). Inteligencia artificial y educación: guía para las personas a cargo de formular políticas., 57. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000379376
- Rubin, D. L., Noy, N. F., & Musen, M. A. (2007). Protege: a tool for managing and using terminology in radiology applications. *Journal of digital imaging*, 20, 34–46.
- Sarmiento Loaiza, R., et al. (2023). Políticas públicas como promotoras de la implementación de las tecnologías de la información y las comunicaciones-tic-en las instituciones educativas oficiales de medellín entre el 2016 y 2021 (Unpublished master's thesis). Facultad de Derecho y Ciencias Políticas.
- Schrepp, M. (2015). User experience questionnaire handbook. All you need to know to apply the UEQ successfully in your project, 50–52.
- Silvervarg, A., Wolf, R., Blair, K. P., Haake, M., & Gulz, A. (2021). How teachable agents influence students' responses to critical constructive feedback. *Journal of Research on Technology in Education*, 53(1), 67–88.
- UNESCO. (2019). Beijing consensus on artificial intelligence and education.. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000368303?posInSet=2&queryId=e945bb2d-7c23-41cc-8585-81de720408ed
- UNESCO, Miao, F., & Holmes, W. (2024). Guía para el uso de ia generativa en educación e investigación. Scientific and Cultural Organisation. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000389227
- Valdivia-Cisneros, E., Vidal, E., & Castro-Gutierrez, E. (2023). Multimodal contactless architecture for upper limb virtual rehabilitation. *International Journal of Advanced Computer Science and Applications*, 14(8).
- Wu, Q., Bansal, G., Zhang, J., Wu, Y., Zhang, S., Zhu, E., ... Wang, C. (2023). Autogen: Enabling next-gen llm applications via multi-agent conversation framework. *arXiv* preprint *arXiv*:2308.08155.

ANEXO A

TITULO