

EDUCACIÓN, CREATIVIDAD E INTELIGENCIA ARTIFICIAL: NUEVOS HORIZONTES PARA EL APRENDIZAJE. ACTAS DEL VIII CONGRESO INTERNACIONAL SOBRE APRENDIZAJE, INNOVACIÓN Y COOPERACIÓN, CINAIC 2025

María Luisa Sein-Echaluce Lacleta, Ángel Fidalgo Blanco y Francisco José García Peñalvo (coords.)

1ª Edición. Zaragoza, 2025

Edita: Servicio de Publicaciones. Universidad de Zaragoza.



Servicio de
Publicaciones
Universidad Zaragoza

EBOOK ISBN 978-84-10169-60-9

DOI 10.26754/uz.978-84-10169-60-9



Esta obra se encuentra bajo una licencia Creative Commons Reconocimiento – NoComercial (ccBY-NC). Ver descripción de esta licencia en <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Referencia a esta obra:

Sein-Echaluce Lacleta, M.L., Fidalgo Blanco, A. & García-Peñalvo, F.J. (coords.) (2025). *Educación, Creatividad e Inteligencia Artificial: nuevos horizontes para el Aprendizaje. Actas del VIII Congreso Internacional sobre Aprendizaje, Innovación y Cooperación. CINAIC 2025 (11-13 de Junio de 2025, Madrid, España)*. Zaragoza. Servicio de Publicaciones Universidad de Zaragoza. DOI 10.26754/uz.978-84-10169-60-9

Integración de SLM en la Educación Superior: un enfoque seguro y experimental en la Universidad de Alicante

SLM integration in Higher Education: a safe and experimental approach at the University of Alicante

Sergio Arjona-Giner, Faraón Llorens-Largo, Alberto Real-Fernández, Rafael Molina-Carmona
sergio.arjona@ua.es, Faraon.Llorens@ua.es, alberto.real@ua.es, rmolina@ua.es

Grupo de investigación Smart Learning
Departamento de Ciencia de la Computación e IA
Universidad de Alicante

Resumen- La llegada de los LLM ha supuesto un gran avance en el procesamiento del lenguaje natural. En la educación, dichos modelos, tienen el potencial de mejorar los procesos de enseñanza y aprendizaje a través de las *Smart Learning Applications*, las cuales pueden ofrecer experiencias personalizadas mediante asistentes inteligentes. No obstante, el uso de los LLM no está libre de riesgos, por ejemplo, la aceleración de la información debido a las alucinaciones. Del mismo modo, las aplicaciones basadas en IA deberán cumplir con el nuevo Reglamento Europeo de IA. Para afrontar estos desafíos se propone un enfoque de sistemas con múltiples agentes especializados que utilicen técnicas de recuperación de información, integrados de forma segura con las plataformas de gestión del aprendizaje. Actualmente, este enfoque se está probando de forma experimental en la Universidad de Alicante, donde se evalúa su eficacia en un entorno real de enseñanza.

Palabras clave: Sistema multiagente, LLM, SLM, RAG, Smart Learning.

Abstract- The advent of LLMs has been a breakthrough in natural language processing. In education, such models have the potential to improve teaching and learning processes through Smart Learning Applications, which can deliver personalised experiences through intelligent assistants. However, the use of LLMs is not without risks, for example, the acceleration of information due to hallucinations. Similarly, AI-based applications will need to comply with the new European AI Regulation. To address these challenges, a multiagent systems approach with specialised agents using information retrieval techniques, securely integrated with learning management platforms, is proposed. This approach is currently being pilot-tested at the University of Alicante, where its effectiveness is being evaluated in a real teaching environment.

Keywords: Multiagent System, LLM, SLM, RAG, Smart Learning.

1. INTRODUCCIÓN

El surgimiento de los *Large Language Models* (LLM) y sus derivados como los *Large Concept Models* (LCM) o *Large Reasoning Models* (LRM) constituye un punto de inflexión en el campo de la Inteligencia Artificial Generativa (genAI, por sus siglas en inglés). Estos modelos no solo destacan en la generación de textos o imágenes, sino también en su capacidad de razonar o tomar decisiones. Gracias a la comprensión del lenguaje natural de estos modelos, se ha impulsado el desarrollo de aplicaciones basadas en LLM, especialmente agentes

inteligentes (LLM-Agents) y su integración cooperativa en Sistemas Multiagente (MAS, por sus siglas en inglés).

Los MAS representan una estructura de IA caracterizada por la interacción cooperativa o competitiva entre múltiples agentes autónomos. Cada agente dentro de este sistema tiene la capacidad de ejecutar tareas específicas y acceder a herramientas determinadas, al tiempo que colabora con otros agentes del sistema para lograr el objetivo común. Estos sistemas permiten resolver problemas complejos que superen las capacidades de sistemas individuales. En este sentido, la incorporación de los LLM en los agentes los dota de una profunda comprensión lingüística y conocimientos generalizados, potenciando su habilidad de toma de decisiones.

La aplicación de los MAS se ha extendido al ámbito educativo, ofreciendo oportunidades transformadoras a través de las *Smart Learning Applications* (SLApp). Estas herramientas educativas son capaces de enriquecer la experiencia educativa del estudiantado, ya sea proporcionando recursos personalizados o mediante asistentes conversacionales; y del profesorado asistiendo a la creación de contenidos o corrección de actividades

2. LLM Y AGENTES INTELIGENTES EN EDUCACIÓN

A. Agentes potenciados por LLM

En el ámbito de la informática, un agente es una entidad que actúa de forma inteligente y autónoma a partir de unos inputs que recibe del entorno (físico o virtual).

Además de percibir el entorno, los agentes se caracterizan por poseer memoria, capacidad de planificación y ejecutar acciones. En el caso de los LLM-Agents, su memoria viene determinada por dos factores: por un lado, su memoria a corto plazo se origina en la conversación de un chat con cada usuario; y, por otro lado, su memoria a largo plazo viene determinada por comportamientos reflejados en la propia red neuronal del modelo o en una base de datos externa. En cuanto a la capacidad de planificación, es posible dotarla mediante técnicas de ingeniería de *prompts* fundamentadas en el Aprendizaje por Contexto (ICL, por sus siglas en inglés), las cuales permiten modificar el comportamiento del LLM. Entre los *prompts* más utilizados están *Chain-of-Thought* (CoT), *Reflection*, *ReAct* o *Program of Thought* (PoT), de hecho, los LRM como

Deepseek-r1 o o3-mini-high han sido entrenados para responder siempre en base al CoT (Huang et al., 2024). Por último, la capacidad de ejecutar acciones viene determinada por la posibilidad de añadir herramientas a los agentes (*tools*). Estas herramientas amplían las capacidades de los agentes permitiéndoles la ejecución de cualquier tipo de programa, por ejemplo, la integración de API, acceso a bases de datos o buscadores web, entre otros. La Figura 1 ilustra los elementos comentados sobre los LLM-Agent.

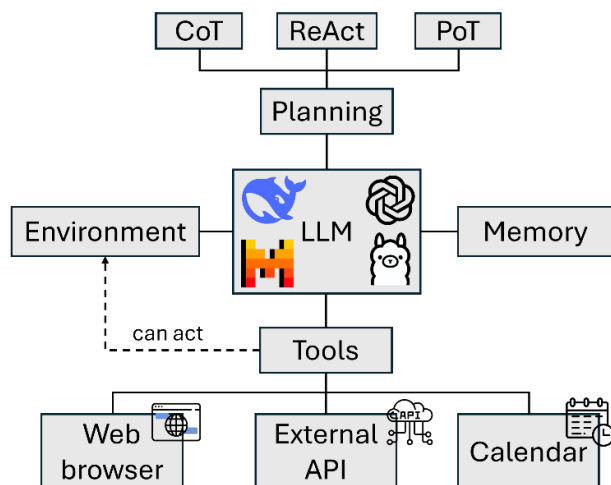


Figura 1. Elementos de un LLM-Agent

Debido a su enfoque especializado, los sistemas *monoagentes* son eficaces para un conjunto limitado de tareas. No obstante, para escenarios más complejos como el educativo, donde los riesgos inherentes a los LLM tienen mayor impacto, se hace necesario adoptar un enfoque *multiagente* que permita no solo delegar responsabilidades entre diferentes agentes especializados, sino también implementar mecanismos de supervisión y control para mitigar riesgos como las alucinaciones (Béchar & Ayala, 2024).

B. Retrieval-Augmented Generation

El *Retrieval-Augmented Generation* (RAG) es un *framework* ICL propuesto por Lewis et al. (2020), que consiste en incorporar conocimiento externo a los LLM: documentos, vídeos, audio, etc. Está compuesto de dos fases.

La primera fase es la de *ingestion*, y consiste en la creación de la base de conocimiento externa mediante la creación de una base de datos. Esta se puede realizar troceando los documentos en fragmentos (*chunks*) y convirtiéndolos en vectores (mediante un modelo de *embeddings*) almacenados en una base de datos vectorial, o mediante un enfoque basado en grafos que modele las relaciones entre conceptos para capturar conexiones más complejas (Edge et al., 2025).

La segunda fase es donde se produce el ICL y se realiza en tres pasos:

- *Retrieval*. Pretende extraer los *chunks* de información más relevantes de la base de conocimiento en base al *prompt* del usuario. En el caso de haber usado una base de datos vectorial esta búsqueda se basa en similitud semántica y, con el enfoque basado en grafos mediante la exploración de relaciones.
- *Augmentation*. Una vez recuperada la información relevante, se procede a enriquecer el *prompt* original.

- *Generation*. Finalmente, el LLM utiliza el *prompt* enriquecido para generar la respuesta final.

La integración del RAG ha demostrado ser efectiva para reducir las alucinaciones en los LLM, ya que fundamenta la generación en datos verificables (Béchar & Ayala, 2024). Además, es posible usar los LLM-Agents con el objeto de mejorar el proceso de RAG. Por ejemplo, el Corrective RAG (CRAG) incorpora un agente evaluador que se encarga de verificar los *chunks* recuperados (Jeong, 2024).

C. Smart Learning Applications

Las SLApp son aplicaciones surgidas del constructo *Smart Learning*, un modelo educativo que usa la IA para personalizar y mejorar el aprendizaje. A diferencia del modelo educativo tradicional, que suele apoyarse en una metodología lineal y uniforme, *Smart Learning* está orientado a satisfacer las necesidades específicas de cada estudiante.

Gracias a los avances en IA, las SLApps están ganando terreno en el ámbito educativo. Estas aplicaciones se integran de forma fluida con los Sistemas de Gestión del Aprendizaje (LMS, por sus siglas en inglés) y permiten el acceso seguro del alumnado y profesorado.

3. HIPÓTESIS Y OBJETIVOS

El sistema educativo actual, inmerso en la era digital, enfrenta retos como la necesidad de ofrecer un aprendizaje personalizado o proporcionar a los estudiantes retroalimentación rápida. Además, los docentes requieren de herramientas que les faciliten la organización de sus tareas estandarizadas, permitiéndoles dedicar más tiempo y atención al desarrollo del alumnado.

En este contexto, se hace imprescindible una transformación digital apoyada en la genAI, ya que las herramientas actuales de IA han demostrado capacidades que permiten desarrollar asistentes inteligentes en el ámbito educativo. Por ello, se plantea la siguiente hipótesis:

Es posible mejorar el proceso educativo mediante asistentes inteligentes sobre un MAS que use las capacidades de la genAI.

Para llevar a la práctica esta hipótesis es necesario que se aborden aspectos fundamentales como el diseño, la implantación y los desafíos académicos. La intención es crear un modelo de asistentes que cumpla con los siguientes objetivos:

- *Objetivo legal*. Cumplir con el Reglamento Europeo de IA (2024), de ahora en adelante RIA, cuya implantación comienza en 2025.
- *Objetivo académico*. Asegurar que el modelo de agentes esté alineado con los objetivos académicos y proteja la privacidad del usuario.
- *Objetivo de integración*. Diseñar una arquitectura que se integre de forma fluida con los LMS, permitiendo distinguir entre roles de profesorado y alumnado.

4. ARQUITECTURA PARA LA INTEGRACIÓN DE ASISTENTES INTELIGENTES EN EDUCACIÓN

A lo largo de esta sección se muestra, por un lado, la arquitectura obtenida en base a los objetivos propuestos y, por otro lado, la aplicación resultante.

A. Cumplimiento de los Objetivos

1) Objetivo legal

Para garantizar el objetivo legal, se debe clasificar el sistema de IA según el RIA, el cual establece tres categorías en función del riesgo del sistema: prácticas prohibidas, sistemas de alto riesgo y sistemas de riesgo limitado.

Para evitar la clasificación de *prohibido*, el sistema no realizará análisis o inferencia de emociones en el alumnado. En cuanto a la clasificación de *alto riesgo*, el sistema no calificará resultados de aprendizaje del alumnado; únicamente podrá asistir estos procesos. No obstante, dado que el sistema interactúa de forma directa tanto con el profesorado como con el alumnado, se clasifica como sistema de *riesgo limitado*, siendo la única obligación la de informar que se utiliza genAI.

Además, el RIA contempla una clasificación adicional en cuanto al modelo utilizado, distinguiéndolo en función de su riesgo sistémico. El RIA considera a los modelos de IA con *riesgo sistémico* cuando presentan "capacidades de gran impacto en base a la evaluación a partir de herramientas y metodologías", cuando lo determine la Comisión Europea de oficio o mediante un grupo de expertos, o cuando cumplan ciertos criterios tales como el número de parámetros, el tamaño del conjunto de datos de entrenamiento o el número de usuarios finales registrados, entre otros. Cabe destacar que, hasta la fecha, la normativa únicamente ha establecido un umbral para la cantidad acumulada de cálculo utilizado durante el entrenamiento, fijado en 10^{25} operaciones en coma flotante (FLOP). Por ello, es imprescindible evaluar si el modelo empleado entra en la categoría de riesgo sistémico o no, para adoptar las medidas correspondientes. Según los datos de entrenamiento del modelo Llama3 (Meta, 2024), este se consideraría de alto riesgo aproximadamente a partir de los 110 billones (en métrica americana) de parámetros, aunque no es generalizable a todos los modelos.

Por lo tanto, para no utilizar modelos de *riesgo sistémico*, se optará por el uso de *Small Language Models* (SLM), los cuales no sobrepasan 100 billones de parámetros.

2) Objetivo de integración

Con el objetivo de potenciar la interoperabilidad de la SLApp en entornos educativos, se buscará que se integre de forma fluida con cualquier LMS compatible con el estándar *Learning Tools Interoperability* (LTI).

LTI es capaz de aportar una conexión segura y estandarizada entre la SLApp y los diferentes LMS. Además, proporciona roles, lo que permite distinguir y gestionar entre docentes, estudiantes y administradores. Esto garantiza que cada perfil acceda únicamente a las funcionalidades y recursos que le corresponden.

3) Objetivo académico

En cuanto al objetivo académico, se utilizará como base la arquitectura LAMB (*Learning, Assistant manager and Builder*) propuesta por Alier et al. (2025), aunque con algunas modificaciones para que permita soporte a MAS.

Una de las características clave de LAMB es el uso del estándar LTI para permitir la integración con el ecosistema educativo, permitiendo de este modo cumplir el objetivo de integración. Además, LAMB garantiza la privacidad del alumnado, pues controla el acceso a la información, asegurando que los datos personales no se transmitan a proveedores

externos de infraestructuras en la nube. En cuanto a la comunicación entre el LMS y LAMB, LTI respeta las políticas institucionales de privacidad, permitiendo que el LMS gestione la información del alumnado y del profesorado.

La arquitectura LAMB se estructura en varios componentes clave (Figura 2) que permiten diseñar, gestionar y desplegar asistentes de aprendizaje.

Uno de sus componentes clave es el *LAMB Knowledge Base*, ya que permite al profesorado crear su propia base de documentación utilizando RAG (mediante el módulo *Content Library Manager*). Con este componente, los docentes pueden cargar sus materiales didácticos (apuntes, vídeos, etc.). Este enfoque asegura que el asistente proporcione respuestas contextualizadas en base a los materiales aportados, asegurando, de este modo, el objetivo académico.

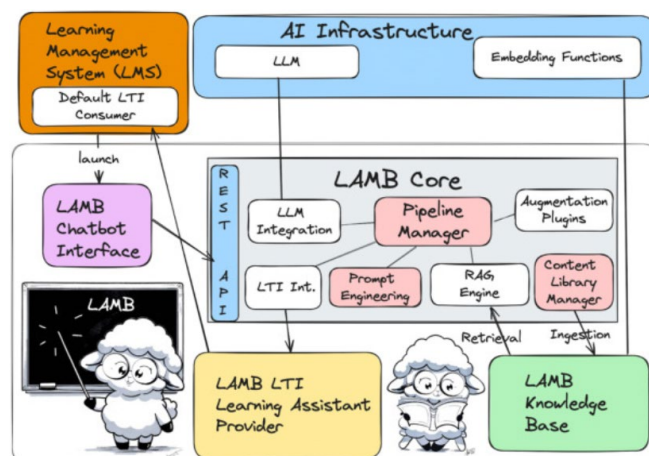


Figura 2. Arquitectura LAMB (Alier et al., 2025)

El componente *LAMB Core* constituye el núcleo central de la arquitectura, encargado de orquestar la comunicación entre todos los módulos del sistema. Sus principales componentes son:

- *RAG Engine*. Se encarga de recuperar los *chunks* del *LAMB Knowledge Base*.
- *Prompt Engineering*. Permite crear probar y gestionar los *prompts* que definen el comportamiento de los asistentes.
- *LLM Integration*. Gestiona la interacción con los modelos del lenguaje.
- *Augmentation Plugins*. Permiten personalizar y ajustar el comportamiento del asistente, incorporando funcionalidades específicas.
- *LTI Integration*. Gestiona la comunicación con el LMS.
- *Pipeline Manager (PM)*. Coordina el flujo de ejecución del asistente. Integra los datos recuperados del *LAMB Knowledge Base*, aplica los *Augmentation Plugins* y gestiona la interacción con el *LLM Integration Layer* para generar y devolver las respuestas generadas al usuario.

B. Aplicación resultado

En base a los objetivos desarrollados, se ha desarrollado una SLApp en formato de aplicación web que cumple con los objetivos mencionados anteriormente. La aplicación se ha puesto en marcha de forma experimental en la Universidad de Alicante mediante el LMS Khipulearn (García-Sigüenza et al., 2023).

En ella, se le permite al profesorado crear asistentes conversacionales y asociarlos a los cursos que imparta, donde cada asistente es una instancia en la que se le permite configurar el *Knowledge Base* (Figura 3), el SLM y sus parámetros (Figura 4), y el *prompt* de sistema mediante el módulo de *Prompt Engineering* (Figura 5). Además, en la Figura 6 se muestra una captura de la interfaz conversacional de la aplicación, la cual será accesible tanto por el profesorado como el alumnado.

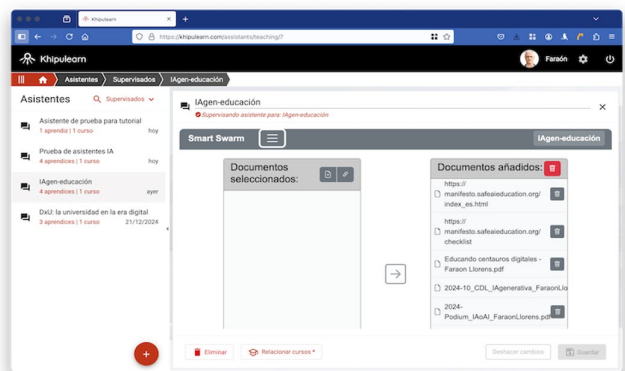


Figura 3. Gestión de conocimiento del asistente

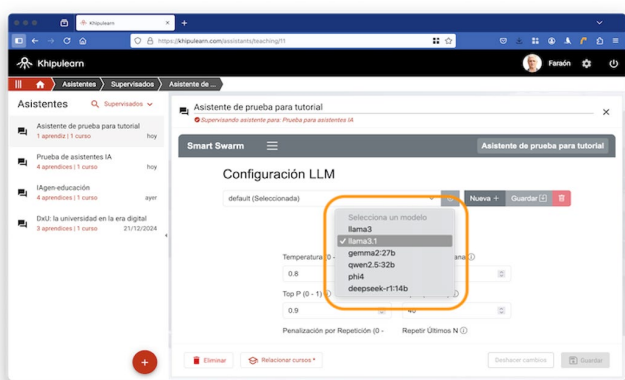


Figura 4. Configuración del SLM y de sus parámetros

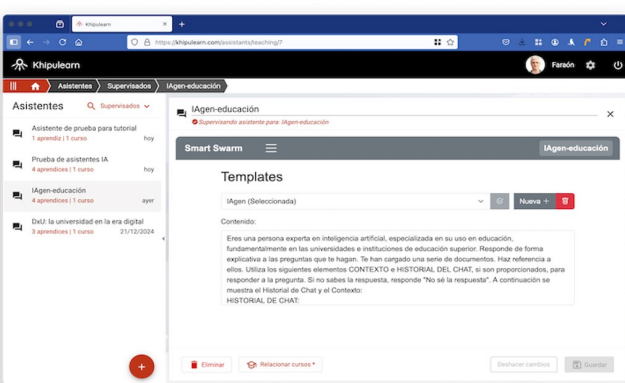


Figura 5. Configuración del *system prompt* del asistente

5. CONCLUSIONES

En este trabajo se ha explorado el potencial de los agentes inteligentes basados en LLM para el ámbito educativo y se ha propuesto un modelo de integración de asistentes en plataformas educativas. Además, la implantación en la Universidad de Alicante abre la puerta a futuros experimentos que permitan refinar un modelo de agentes que interactúen tanto con profesores como estudiantes.

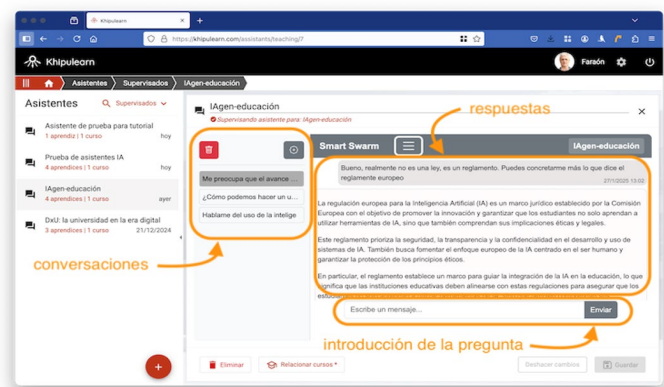


Figura 6. Ejemplo de la interfaz conversacional del asistente

REFERENCIAS

- Alier, M., Pereira, J., García-Peñalvo, F. J., Casañ, M. J., & Cabré, J. (2025). LAMB: An open-source software framework to create artificial intelligence assistants deployed and integrated into learning management systems. *Computer Standards & Interfaces*, 92, 103940. <https://doi.org/10.1016/j.csi.2024.103940>
- Béchar, P., & Ayala, O. M. (2024). *Reducing hallucination in structured outputs via Retrieval-Augmented Generation*. arXiv. <https://doi.org/10.48550/ARXIV.2404.08189>
- Edge, D., Trinh, H., Cheng, N., Bradley, J., Chao, A., Mody, A., Truitt, S., Metropolitansky, D., Ness, R. O., & Larson, J. (2025). *From Local to Global: A Graph RAG Approach to Query-Focused Summarization* (arXiv:2404.16130). arXiv. <https://doi.org/10.48550/arXiv.2404.16130>
- García-Sigüenza, J., Real-Fernández, A., Molina-Carmona, R., & Llorens-Largo, F. (2023). Two-Phases AI Model for a Smart Learning System. En P. Zaphiris & A. Ioannou (Eds.), *Learning and Collaboration Technologies* (Vol. 14040, pp. 42-53). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-34411-4_4
- Huang, X., Liu, W., Chen, X., Wang, X., Wang, H., Lian, D., Wang, Y., Tang, R., & Chen, E. (2024). *Understanding the planning of LLM agents: A survey* (arXiv:2402.02716). arXiv. <https://doi.org/10.48550/arXiv.2402.02716>
- Jeong, C. (2024). A Study on the Implementation Method of an Agent-Based Advanced RAG System Using Graph. *Knowledge Management Research*, 25(3), 99-119. <https://doi.org/10.15813/kmr.2024.25.3.005>
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. <https://doi.org/10.48550/ARXIV.2005.11401>
- Meta. (2024). *Introducing Meta Llama 3: The most capable openly available LLM to date*. <https://ai.meta.com/blog/meta-llama-3/>
- Reglamento Europeo de Inteligencia Artificial, UE 2024/1689 (2024). <https://eur-lex.europa.eu/eli/reg/2024/1689/oj?locale=es>