Prefijo DOI: 10.70417
ISSN: 2594-2298

Crossref

latindex

Compartiendo conocimientos

FEGLININ

Año 10, N°37, abril-junio 2026

REVISTA OFICIAL
DE LA

F.G.D.P
FEDERACIÓN GLOBAL DE PROFESIONES

Tópicos:

Tecnología; Innovación; Calidad; Productividad y Desarrollo Económico.



ISSN 2594-2298

© FEGLININ

| Año 10, No 37, abril - junio 2026 |
https://doi.org/10.70417/S30dE4cnB27

REVISIÓN SISTEMÁTICA DE LA LITERATURA: APLICACIÓN DE SISTEMAS DE RECUPERACIÓN DE INFORMACIÓN, NLP Y ARQUITECTURAS RAG EN EL DOMINIO JURÍDICO-ADMINISTRATIVO DE INSTITUCIONES EDUCATIVAS



SYSTEMATIC LITERATURE REVIEW: APPLICATION OF INFORMATION RETRIEVAL SYSTEMS, NLP, AND RAG ARCHITECTURES IN THE LEGAL-ADMINISTRATIVE DOMAIN OF EDUCATIONAL INSTITUTIONS

Ing. Edgar David Mariano Ruiz[1]
Dr. Arnulfo Catalán Villegas[2]
Dra. Mercedes Hernández de la Cruz[3]
Dr. Rene Edmundo Cuevas Valencia[4]

RESUMEN

Las Instituciones de Educación Superior se ven desafiadas al consultar administrativamente grandes volúmenes de normativas. Tradicionalmente, la búsqueda depende de motores léxicos, ineficientes ante consultas complejas. Esta revisión sistemática aborda cómo el Procesamiento de Lenguaje Natural en Sistemas de Recuperación de Información y la arquitectura de Generación Aumentada por Recuperación superan esta problemática. Siguiendo la metodología PRISMA, se analizaron artículos en bases de datos académicas para documentar la migración de repositorios estáticos a asistentes de Inteligencia Artificial impulsados por Grandes Modelos de Lenguaje y bases de datos vectoriales. Los hallazgos demuestran que las consultas semánticas mejoran significativamente la precisión y usabilidad. Sin embargo, persisten retos considerables como la pérdida de contexto legal por una segmentación inadecuada de documentos (chunking), además de una notable escasez de trabajos en el dominio administrativo hispanohablante.

ABSTRACT

Higher Education Institutions are challenged when administratively consulting large volumes of regulations. Traditionally, information retrieval relies on lexical search engines, which are inefficient for complex queries. This systematic review addresses how Natural Language Processing in Information Retrieval Systems and Retrieval-Augmented Generation architecture overcome this problem. Following the PRISMA methodology, articles in academic databases were analyzed to document the migration from static repositories to Artificial Intelligence assistants powered by Large Language Models and vector databases. Findings demonstrate that semantic queries significantly improve precision and usability. However, considerable challenges remain, such as the loss of legal context due to inadequate document segmentation (chunking), as well as a notable scarcity of works in the Spanish-speaking administrative domain.

Palabras clave: Recuperación de Información, Procesamiento de Lenguaje Natural, Generación Aumentada por Recuperación, Bases de Datos Vectoriales, Educación Superior, Asistentes Virtuales.

Key Words: *Information Retrieval, Natural Language Processing, Retrieval-Augmented Generation, Vector Databases, Higher Education, Virtual Assistants.*

[1] Ing. Edgar David Mariano Ruiz 25600451@uagro.mx ORCID: 0009-0005-6171-1206
[2] Dr. Arnulfo Catalán Villegas catalanvillegas@uagro.mx ORCID: 0009-0001-0391-7960
[3] Dra. Mercedes Hernández de la Cruz mercedes.hd@chilpancingo.tecnm.mx ORCID: 0000-0001-8837-3204
[4] Dr. Rene Edmundo Cuevas Valencia reneecuevas@uagro.mx ORCID: 0000-0001-9528-7603

Prefijo DOI: 10.70417

ISSN 2594-2298
latindex

INTRODUCCIÓN

La comunidad universitaria experimenta dificultades operativas al recuperar normativa específica (como un artículo de un reglamento o cláusula de una beca). Estas Instituciones de Educación Superior (IES) gestionan repositorios dispersos de información con lenguaje altamente burocrático y legal. En la práctica, la búsqueda se limitaba a ingresar palabras exactas en sistemas de búsqueda obsoletos o heredados, resultando en búsquedas fallidas, situación que estudiantes y académicos han vivido o recuperación de documentos irrelevantes o que ya expiraron, todo esto si el usuario desconoce la terminología especializada.

Este es un problema significativo, ya que, en el dominio normativo y administrativo de una IES, hay una conexión entre documentos donde ciertos reglamentos pueden modificar otros o un procedimiento depende de múltiples cláusulas. La búsqueda basada únicamente en palabras clave ha resultado ser insuficiente para este contexto. Por eso resulta necesario evolucionar hacia herramientas con inteligencia artificial, en especial la recuperación de contexto semántico. Gracias al Procesamiento de Lenguaje Natural (NLP, Natural Language Process por sus siglas en inglés) y a arquitecturas avanzadas con la Generación Aumentada por Recuperación (RAG, Retrieval Augmentation Generation por sus siglas en inglés), las búsquedas pueden ir más allá de la coincidencia exacta de caracteres, midiendo la similitud matemática entre la intención del usuario y el significado profundo de los textos institucionales (Hu et al., 2025).

No obstante, implementar estas arquitecturas en el entorno de una IES no es sencillo; es un proceso que supone superar limitaciones, como la dificultad de los modelos computacionales para trabajar con documentos extensos sin perder contexto legal; y, sobre todo, la barrera del idioma, ya que, si bien los modelos funcionan bien en inglés, en el español universitario hay rezago, ya que gran parte de los avances EN NLP jurídicos están enfocados en inglés. Esta revisión se propone examinar cómo la literatura científica reciente aborda estas limitaciones, identificando qué herramientas están definiendo el estándar tecnológico en las IES del siglo XXI.

METODOLOGÍA

Para evitar sesgos y garantizar que la recolección de literatura fuera exhaustiva y transparente, este trabajo se apoyó en los lineamientos de PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). PRISMA no es un simple formato, sino un estándar científico internacional basado en la evidencia. Su estructura obliga a detallar explícitamente cómo se identificaron, seleccionaron, filtraron y evaluaron los estudios primarios, permitiendo que otro investigador pueda replicar el proceso y llegar a conclusiones equivalentes (Aliaga et al., 2025).

El proceso de recolección comenzó con el diseño de ecuaciones de búsqueda construidas mediante operadores booleanos (AND, OR, NOT), vinculando conceptos como la inteligencia artificial, la búsqueda semántica y las arquitecturas conversacionales en entornos educativos y legales. La ejecución de estas ecuaciones se llevó a cabo en cuatro de los índices científicos más robustos: Web of Science, Scopus, Science Direct y EBSCO. Al explorar

ISSN 2594-2298
© FEGLININ
| Año 10, No 37, abril - junio 2026 |
https://doi.org/10.70417/S30dE4cnB27

estas plataformas, notamos dinámicas interesantes en los resultados. Por ejemplo, en Scopus y Science Direct se concentró la mayor parte de la literatura "dura"; es decir, artículos detallando la matemática detrás de los algoritmos de recuperación densa o el rendimiento de bases de datos de grafos. Mientras que en EBSCO y Web of Science predominaron estudios con un enfoque más orientado hacia la usabilidad, evaluando cómo los estudiantes interactúan con los chatbots bilingües y cómo estas herramientas mitigan la brecha de alfabetización digital.

El flujo de selección funcionó como un filtro progresivo. En la fase inicial de identificación no se aplicaron restricciones. Posteriormente, aplicamos filtros básicos para acotar el periodo de publicación (2020-2026) y depuramos todos los registros duplicados entre las diferentes bases de datos. Durante la fase de cribado, se excluyeron decenas de artículos que, aunque contenían términos como "procesamiento de lenguaje" o "redes neuronales", pertenecían al ámbito médico o clínico, un dominio que no corresponde al contexto universitario administrativo. Finalmente, en la etapa de elegibilidad, se realizó una lectura a texto completo de los manuscritos restantes, evaluando la calidad de su arquitectura de software, su pertinencia temática y sus metodologías de validación. Como se observa en la Figura 1, el proceso de selección de literatura se estructuró conforme a los lineamientos PRISMA 2020. De los 95 registros identificados en bases de datos científicas, 15 resultaron duplicados y fueron eliminados. Los 80 registros restantes se evaluaron por título y resumen; 45 quedaron fuera por no cumplir los criterios de inclusión. Se analizaron 35 artículos a texto completo, de los cuales 16 fueron descartados por razones metodológicas y de pertinencia temática. Finalmente, se incluyeron 19 estudios en la revisión sistemática.

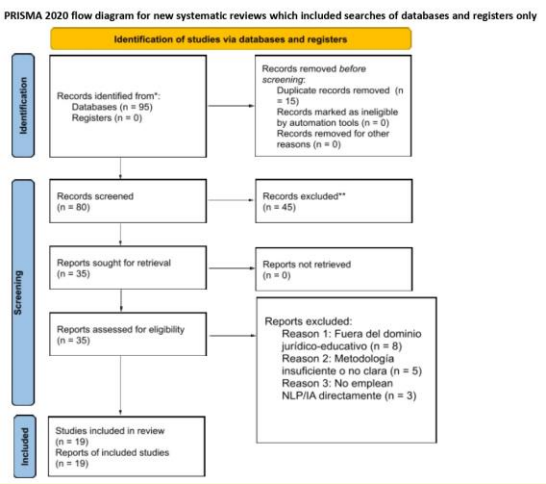


Figura 1. Flujo de selección y depuración de registros bibliográficos conforme a PRISMA 2020.

RESULTADOS Y DISCUSIÓN

Al analizar la literatura seleccionada de acuerdo a los lineamientos PRISMA, se encontraron los siguientes casos de estudio y modelos matemáticos que son el estándar actual en la interacción entre usuarios y documentos de IES, más allá de un listado de autores estos artículos están listando tecnologías que están consolidándose, sin embargo, no han logrado escalar fuera de entornos controlados. La Tabla 1 muestra las contribuciones de los trabajos

Prefijo DOI: 10.70417

ISSN 2594-2298
latindex

más relevantes en esta revisión donde se destaca sus enfoques, herramientas empleadas y hallazgos que inciden directamente en el desarrollo de asistentes en la comunidad universitaria.

La Tabla 1 Contribuciones de los trabajos más relevantes

Autor(es) y Año	Enfoque del Estudio	Herramientas / Tecnologías Clave	Observaciones y Resultados Relevantes
Neupane et al. (2024)	Chatbot (BARKPLUG V.2) para recursos universitarios.	RAG, LLMs (GPT-3.5-turbo), Chroma DB, LangChain, Scrapy.	Lograron una precisión y relevancia superior al 95%. Se comprobó que el uso de Chroma DB como base vectorial elimina las "alucinaciones" del modelo al forzarlo a basarse solo en textos de la universidad.
Hsain & El Housni (2024)	Asistente conversacional bilingüe para estudiantes internacionales.	RAG, GPT-3.5, GPT-4 (como verificador), LangChain.	Demostraron que el enfoque de "doble prompt" (uno genera la respuesta y otro la audita) es crucial en temas burocráticos. La validación garantizó que el sistema no inventara normativas migratorias o académicas.
Li et al. (2023)	Modelo pre-entrenado para recuperación de casos legales (SAILER).	Arquitectura Transformer sensible a la estructura, encoder asimétrico.	Evidenciaron que tratar un documento legal como texto plano destruye su lógica. SAILER codifica separadamente "hechos", "razonamientos" y "decisiones", logrando un rendimiento superior en comprensión de dependencias largas.
Tang et al. (2023)	Reformulación de entradas para recuperación de casos (PromptCase).	Modelos pre-entrenados (SAILER, BERT), Prompts instructivos.	Comprobaron que inyectar "resúmenes de hechos jurídicos" directamente en el codificador previene la pérdida de contexto cuando el documento excede el límite de tokens del modelo.
Cha et al. (2023)	Recuperación de conocimiento mediante grafos para especificaciones de compras.	Bases de datos de Grafos, Neo4j, lenguaje Cypher.	Demostraron que los grafos (Knowledge Graphs) modelan jerarquías y reglas lógicas mejor que los vectores aislados, reduciendo los errores humanos en el análisis de normativas técnicas de contratación.
Aldama García et al. (2024)	Creación de corpus legal en español (3CEL).	Extracción de Información, modelado span, RoBERTalex, MEL.	Su análisis revela un vacío tecnológico preocupante. Al intentar entrenar modelos en español, notaron que el desbalance de clases y el vocabulario ambiguo ralentizan drásticamente el aprendizaje de la IA.
González-González et al. (2024)	Explicabilidad en la clasificación automática de sentencias en España.	Machine Learning, Random Forest, Árboles de Decisión, XAI.	Un aporte vital: los sistemas jurídicos no pueden ser "cajas negras". Implementaron plantillas en lenguaje natural que le explican al usuario exactamente qué palabras usó la máquina para tomar una decisión.

Tabla 1. Resumen del Estado del Arte en Recuperación de Información y NLP.

ISSN 2594-2298
© FEGLININ
| Año 10, No 37, abril - junio 2026 |
https://doi.org/10.70417/S30dE4cnB27

ARQUITECTURAS RAG Y BASES DE DATOS VECTORIALES

Al analizar implementaciones reales en IES, se observa cómo la arquitectura RAG ha desplazado los antiguos chatbots basados en flujos preprogramados, pero resulta un problema en este desplazo, los LLMs puros (como ChatGPT o Gemini) tienden a generar alucinaciones fácticas o producir información apócrifa cuando no cuentan con cierta información compleja en sus entrenamientos, para solucionar esto las IES han optado por adoptar ecosistemas basados en recuperación vectorial. En la Mississippi State University se cuenta con un sistema BARKPLUG V.2 el cual es un ejemplo de esta adopción (Neupane et al., 2024), el sistema usa la base de datos vectorial Chroma DB, el flujo es el siguiente: primero se extraen miles de reglamentos, manuales y normativa alojados en la web de la universidad, posteriormente a través de "embeddings", cada párrafo se convierte en un vector denso de información el cual representa su significado semántico en un espacio multidimensional y se almacena en Chroma DB.

Cuando alguien de la comunidad universitaria pregunta algo complejo, el sistema convierte esta petición en un vector con los mismos embeddings con los cuales se almacena la información previamente recopilada y mediante una búsqueda de "similitud del coseno" dentro de Chroma DB. La base de datos recupera en cuestión de segundos la información que matemáticamente está más cercana a la pregunta, estamos hablando de una recuperación semántica, dejando de lado la búsqueda por palabras exactas. Es solo entonces cuando el LLM (orquestado a través de librerías como LangChain) recibe esta instrucción y mediante Ingeniería de Prompts se le formula la petición para que no se salga del contexto y se le envían los resultados obtenidos mediante la similitud del coseno, esta inyección factual blinda al sistema de alucinaciones.

El mismo enfoque se aplicó a la Universidad de Al Akhawayn para atender a estudiantes extranjeros, en este estudio además de usar repositorios vectoriales, el equipo de desarrollo incorporó a GPT-4 como "inspector" (Hsain & El Housni, 2024). Dentro del flujo de trabajo, el primer agente redactó la respuesta basándose en la similitud del coseno y en una segunda etapa este agente "inspector" auditaba la respuesta generada antes de mostrarla al estudiante, desechando cualquier consejo u orientación que no estuviera literalmente respaldada por el reglamento institucional.

LONGITUD DOCUMENTAL Y MODELADO ESTRUCTURAL

La literatura actual sostiene un debate técnico, en la documentación jurídica y administrativa los textos son largos y complejos, aunque la recuperación vectorial parece la solución definitiva, es en esta fase del proceso donde radica la principal limitación técnica, ya que a pesar de que los motores de comprensión profunda más populares, procesan correctamente bloques de 512 tokens (400 palabras aproximadamente), colapsan ante un estatuto universitario de cincuenta páginas. La solución que adopta la industria actual es la segmentación (chunking) por bloques de 500 palabras. Tang et al. (2023) advierten que cortar arbitrariamente un texto destruye las referencias cruzadas, como ejemplo una excepción legal mencionada en la página cinco de un estatuto pierde su significado semántico si se separa de la regla general expuesta en la página dos. Como se observa en la Figura 2, la segmentación

Prefijo DOI: 10.70417

ISSN 2594-2298
latindex

tradicional fragmenta el documento arbitrariamente, separando el contenido sin respetar sus componentes jurídicos, en contraste, el modelado estructural divide el contenido respetando estos componentes y preservando las dependencias semánticas.

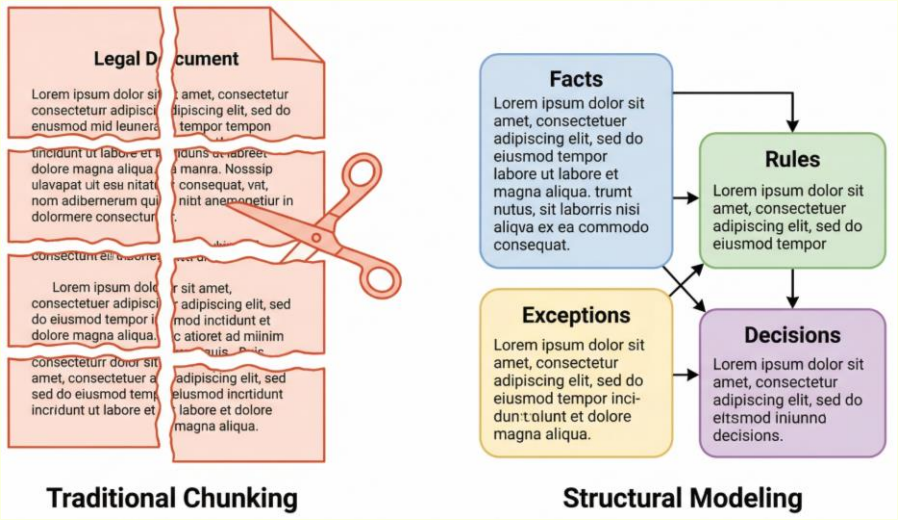


Figura 2. Evolución de la segmentación de texto.

Para resolver este desafío, se propone un "modelado consciente de la estructura", uno de los avances encontrados en esta revisión es la arquitectura SAILER, desarrollada por Li et al. (2023), en lugar de leer un documento extenso, se decodifica asimétricamente desarmándolo en sus componentes lógicos, se separan los hechos de los razonamientos y de las decisiones. Al obligar a una red neuronal a entrenarse respetando la arquitectura de la normativa, el modelo aprende la dependencia semántica sin requerir anotaciones manuales, otra herramienta propuesta por Tang et al. (2023) es PromptCase, que aísla los componentes jurídicos antes de inyectarlos a la base de datos, previniendo así que el ruido contamine la búsqueda.

VACÍO HISPANO Y LA NECESIDAD DE GRAFOS Y EXPLICABILIDAD

Estos documentos técnicos requieren recursos idiomáticos fundacionales, es aquí donde la revisión sistemática advierte que, para nuestro contexto, los modelos multilingües actuales presentan deficiencias al procesar el léxico legal en español. Investigadores del Instituto de Ingeniería del Conocimiento advierten sobre la carencia de un corpus hispanohablante en tareas de extracción de información (Aldama García et al., 2024). Aunque el desarrollo del corpus 3CEL (Cláusulas de contratos legales en español) y los ensayos con modelos como RoBERTalex y MEL representan un esfuerzo ante esta brecha, los experimentos confirman que la naturaleza técnica de nuestro idioma propone un entrenamiento que muchas IES no están preparadas para asumir debido al costo monetario que supone reentrenar un modelo, sin contar la infraestructura con la que el contexto latinoamericano no dispone.

Aparte de las barreras de la segmentación y el idioma, hay una última parte en esta evolución que la literatura señala: los grafos de conocimiento y la explicabilidad (XAI). Los vectores son excelentes buscando similitudes de significado, pero las bases de datos de grafos (Knowledge Graphs) estructuran la información como una red de nodos y relaciones reales,

ISSN 2594-2298
© FEGLININ
| Año 10, No 37, abril - junio 2026 |
https://doi.org/10.70417/S30dE4cnB27

un grafo sabe exactamente qué relación guarda la regla del "Estatuto A" está subordinada por el "Estatuto B", permitiendo una recuperación basada en razonamiento lógico y no solo en cercanía semántica. Investigaciones como el modelo POKREM, operando sobre plataformas como Neo4j, han demostrado que es posible mapear jerarquías complejas (Cha et al., 2023).

Siguiendo esta misma línea, de nada sirve un sistema inteligente si el usuario no puede auditar sus decisiones, como ejemplo si un estudiante es dado de baja o pierde una beca por la interpretación del sistema, el sistema debe justificar el porqué de su respuesta. González-González et al. (2024) para resolver esto han integrado sistemas híbridos de Machine Learning (Random Forest y árboles de decisión) para rastrear el camino que tomó el algoritmo y generar reportes en lenguaje natural que explican al usuario qué palabras y pesos umbrales desencadenan esa respuesta normativa específica.

RESULTADOS ESPERADOS

De acuerdo a esta revisión de la literatura, se espera que el despliegue de arquitecturas RAG en las IES reduzca el tiempo de recuperación de información y asimismo se logre una mitigación de alucinaciones propias de los LLMs. A nivel tecnológico un resultado crítico es desarrollar un corpus especializado en el dominio administrativo universitario en idioma español, así como modelos de segmentación (chunking) que preserven las jerarquías y dependencias lógicas dentro de los documentos, superando así la fragmentación tradicional.

CONCLUSIONES

La transición de ecosistemas semánticos apoyados por inteligencia artificial no es un capricho tecnológico, sino una respuesta al colapso burocrático actual en IES. Esta revisión deja claro que las arquitecturas RAG y las bases de datos vectoriales tienen la capacidad de democratizar la normativa institucional, a su vez en su forma más avanzada se están estudiando los grafos de conocimiento. Sin embargo, el despliegue de estos ecosistemas en el contexto latinoamericano debe hacerse con cautela y rigor ya que su adopción exitosa depende de que se supere la práctica ineficiente de fragmentar los textos, respetando la normativa en su estructura y jerarquía, asimismo, se hace indispensable un esfuerzo regional para nutrir corpus especializados en derecho administrativo hispanohablante y exigir que el despliegue de estos sistemas cuente con módulos de explicabilidad transparentes, solo así logramos que la tecnología sirva verdaderamente a la comunidad universitaria.

REFERENCIAS

Aldama García, N., Marsà Morales, P., Betancur Sánchez, D., Barbero Jiménez, Á., Guerrero Nieto, M., Haya Coll, P., Martín Chozas, P., & Montiel Ponsoda, E. (2025). 3CEL: A corpus of legal Spanish contract clauses. arXiv preprint arXiv:2501.15990. https://doi.org/10.48550/arXiv.2501.15990.

Aliaga, M. R., Llacctahuaman, J. S., Esquivel, C. N., & Saboya, N. (2025). Mobile applications that incorporate AI for information search and recommendation: A systematic literature review. International Journal of Advanced Computer Science and Applications, 16(8). https://doi.org/10.14569/IJACSA.2025.0160863.

Prefijo DOI: 10.70417

ISSN 2594-2298
latindex

Althammer, S., Askari, A., Verberne, S., & Hanbury, A. (2021). DoSSIER@COLIEE 2021: Leveraging dense retrieval and summarization-based re-ranking for case law retrieval. arXiv preprint arXiv:2108.03937. https://doi.org/10.48550/arXiv.2108.03937.

de Arriba-Pérez, F., García-Méndez, S., González-Castaño, F. J., & González-González, J. (2023/2024). Explainable machine learning multi-label classification of Spanish legal judgements. Journal of King Saud University - Computer and Information Sciences. https://doi.org/10.1016/j.jksuci.2022.10.015.

Aumiller, D., Almasian, S., Lackner, S., & Gertz, M. (2022). EUR-Lex-Sum: A multi- and cross-lingual dataset for long-form summarization in the legal domain. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. https://doi.org/10.18653/v1/2022.emnlp-main.519.

Bai, X., & Li, J. (2021). Applied research of knowledge in the field of artificial intelligence in the intelligent retrieval of teaching resources. Scientific Programming. https://doi.org/10.1155/2021/9924435.

Bithel, S., & Malagi, S. S. (2021). Unsupervised identification of relevant prior cases. arXiv preprint arXiv:2107.08973. https://doi.org/10.48550/arXiv.2107.08973.

Cha, H.-J., Choi, S.-W., Lee, E.-B., & Lee, D.-M. (2023). Knowledge retrieval model based on a graph database for semantic search in equipment purchase order specifications for steel plants. Sustainability, 15, 6319. https://doi.org/10.3390/su15076319.

Esteva, A., Kale, A., Paulus, R., Hashimoto, K., Yin, W., Radev, D., & Socher, R. (2021). COVID-19 information retrieval with deep-learning based semantic search, question answering, and abstractive summarization. npj Digital Medicine, 4, 68. https://doi.org/10.1038/s41746-021-00437-0.

González-González, J., de Arriba-Pérez, F., García-Méndez, S., Busto-Castiñeira, A., & González-Castaño, F. J. (2024). Automatic explanation of the classification of Spanish legal judgments in jurisdiction-dependent law categories with tree estimators. Journal of King Saud University - Computer and Information Sciences. https://doi.org/10.1016/j.jksuci.2023.101634.

Henriquez, C., Sánchez-Torres, G., & Salcedo, D. (2021). Tashi-Bot: A intelligent personal assistant for users in an educational institution. Journal of Management Information and Decision Sciences, 24(S6), 1–11. https://doi.org/10.20944/preprints202108.0380.v1.

Hsain, A., & El Housni, H. (2024). Large language model-powered chatbots for internationalizing student support in higher education. arXiv preprint arXiv:2403.14702. https://doi.org/10.48550/arXiv.2403.14702.

Hu, B., Malik, I., Chen, Q., Xie, H., & Sohail, N. (2025). Bridging digital literacy gaps with AI-driven semantic search technologies. International Journal on Semantic Web and Information Systems, 21(1), 1–17. https://doi.org/10.4018/IJSWIS.380355.

Isaev, R., Gumerov, R., Esenalieva, G., & Mekuria, R. R. (2022). HIVA, holographic intellectual voice assistant. arXiv preprint arXiv:2307.05501. https://doi.org/10.48550/arXiv.2307.05501.

Krasadakis, P., Sakkopoulos, E., & Verykios, V. S. (2024). A survey on challenges and advances in natural language processing with a focus on legal informatics and low-resource languages. Electronics, 13, 648. https://doi.org/10.3390/electronics13030648.

ISSN 2594-2298

© FEGLININ

| Año 10, No 37, abril - junio 2026 |
https://doi.org/10.70417/S30dE4cnB27

Li, H., Ai, Q., Chen, J., Dong, Q., Wu, Y., & Liu, Y. (2023). SAILER: Structure-aware pre-trained language model for legal case retrieval. In Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. https://doi.org/10.1145/3539618.3591761.

Neupane, S., Hossain, E., Keith, J., Tripathi, H., Ghiasi, F., Golilarz, N. A., Amirlatifi, A., Mittal, S., & Rahimi, S. (2024). From questions to insightful answers: Building an informed chatbot for university resources. arXiv preprint arXiv:2405.08120. https://doi.org/10.48550/arXiv.2405.08120.

Rabelo, J., Goebel, R., Kim, M.-Y., Kano, Y., Yoshioka, M., & Satoh, K. (2024). Overview and discussion of the competition on legal information extraction/entailment (COLIEE) 2021. The Review of Socionetwork Strategies, 18(1), 27–47. https://doi.org/10.1007/s12626-023-00152-0.

Tang, Y., Qiu, R., & Li, X. (2023). PromptCase: Effective input reformulation for legal case retrieval. arXiv preprint arXiv:2309.02962. https://doi.org/10.48550/arXiv.2309.02962.

Prefijo DOI: 10.70417

F.G.D.P

FEDERACIÓN
GLOBAL
DE PROFESIONES

