



Universidad Nacional de San Luis
RECTORADO

"2026 - Año de la Grandeza Argentina"

"150° Aniversario de la Creación
de la Escuela Normal Juan Pascual Pringles"

"50 años por la Memoria, la Verdad y la Justicia. Nunca más"



SAN LUIS, 21 de abril de 2026.-

VISTO:

El EXPE: 4442/2025, en el cual obran las actuaciones vinculadas a la Pasantía de Posgrado "*Construcción de un Grafo de Conocimiento mediante Procesamiento del Lenguaje Natural y Aprendizaje Automático para el estudio de la quimiorresistencia en Cáncer de Ovario*"; y

CONSIDERANDO:

Que se presentó la propuesta de Pasantía de Posgrado del Lic. Horacio Jesús THOMPSON, estudiante del Doctorado en Ciencias de la Computación a realizarse en el Departamento de Patología de la Universidad McGill (Montreal, Canadá), bajo la dirección del Dr. Carlos TELLERIA.

Que la Comisión Asesora de Investigación (CAI) actuando como Comisión de Posgrado, en su Sesión del 8 de mayo de 2025, luego de analizar la propuesta, recomendó al Consejo de Posgrado aprobar el Plan de Pasantía propuesto y autorizar su realización.

Que el Consejo de Posgrado de la Universidad Nacional de San Luis, en su sesión del día 20 de mayo de 2025, evaluó la propuesta, considerándola aprobada.

Que a fs. 66-82 del expediente de referencia obra el Informe de Pasantía de Posgrado.

Que la Comisión Asesora de Investigación, en su sesión del día 26 de febrero de 2026, luego del análisis del presente expediente, aconseja la aprobación del Informe Final presentado por el Lic. Horacio Jesús THOMPSON.

Que el Consejo de Posgrado de la Universidad Nacional de San Luis, en su sesión del 16 de marzo de 2026, luego del análisis, acuerda la aprobación del informe final.

Que corresponde su protocolización.

Por ello, y en uso de sus atribuciones:



Universidad Nacional de San Luis
RECTORADO

"2026 - Año de la Grandeza Argentina"

"150° Aniversario de la Creación
de la Escuela Normal Juan Pascual Pringles"

"50 años por la Memoria, la Verdad y la Justicia. Nunca más"



EL RECTOR DE LA UNIVERSIDAD NACIONAL DE SAN LUIS

RESUELVE:

ARTÍCULO 1°.- Dar por aprobada la Pasantía de Posgrado "*Construcción de un Grafo de Conocimiento mediante Procesamiento del Lenguaje Natural y Aprendizaje Automático para el estudio de la quimiorresistencia en Cáncer de Ovario*" realizada por el Licenciado en Ciencias de la Computación Horacio Jesús THOMPSON, DU N° 37214321, estudiante del Doctorado en Ciencias de la Computación, llevada a cabo en el Departamento de Patología de la Universidad McGill (Montreal, Canadá), bajo la dirección del Dr. Carlos TELLERÍA, DU N° 16918530.

ARTÍCULO 2°.- Comuníquese, notifíquese, publíquese en el Digesto Administrativo de la Universidad Nacional de San Luis, insértese en el Libro de Resoluciones y archívese.

RC

Documento firmado digitalmente según Ordenanza Rectoral N° 15/2021 por: Rector GIL, Raúl Andrés - Secretaria Académica, de Innovación Educativa y Posgrado LORENZO, Rosa Alejandra.



Universidad Nacional de San Luis
RECTORADO

2026 – “Año de la Grandeza Argentina”

“150° Aniversario de la Creación de la Escuela Normal Juan Pascual
Pringles”

“50 años por la Memoria, la Verdad y la Justicia. Nunca más”

ANEXO

PASANTÍA DE POSGRADO

APELLIDO Y NOMBRE DEL PASANTE: Horacio Jesús THOMPSON

CARRERA DE POSGRADO: Doctorado en Ciencias de la Computación

UNIVERSIDAD DONDE CURSA LA CARRERA: Universidad Nacional de San Luis (UNSL)

DENOMINACIÓN DE LA PASANTÍA: *“Construcción de un Grafo de Conocimiento mediante Procesamiento del Lenguaje Natural y Aprendizaje Automático para el estudio de la quimiorresistencia en Cáncer de Ovario”*

FECHA DE INICIO: 16 de junio de 2025

DURACIÓN: 6 meses

APELLIDO Y NOMBRE DEL DIRECTOR: Dr. Carlos TELLERIA

INSTITUCIÓN LABORAL DE ORIGEN: Universidad Nacional de San Luis

LUGAR DE REALIZACIÓN DE LA PASANTÍA: Departamento de Patología de la Universidad McGill, Montreal, Canadá

FUNDAMENTACIÓN

En los últimos años, el avance de la Inteligencia Artificial (IA) y el Procesamiento del Lenguaje Natural (PLN) ha permitido explorar nuevas estrategias para la recuperación y análisis de información a gran escala. La Web es una fuente inagotable de datos que ofrece acceso a contenidos que abarcan desde información estructurada en bases de datos científicas como PubMed Central, ClinicalTrials.gov y DrugBank, hasta datos no estructurados provenientes de redes sociales y foros de discusión como Reddit, Twitter (actualmente conocido como X) y Telegram. Mientras que los artículos científicos siguen esquemas estandarizados que garantizan su organización y validación, las redes sociales brindan una perspectiva dinámica y actualizada de la información, reflejando tendencias emergentes, experiencias directas y discusiones comunitarias sobre distintos temas [1]. Sin embargo, la variabilidad de esta información, influenciada por factores contextuales como edad, sexo biológico y aspectos sociodemográficos, así como fenómenos lingüísticos como abreviaturas, ironía y sarcasmo, dificulta su interpretación y la extracción precisa de conocimiento [2]. A pesar de esto, las redes sociales aportan información valiosa para comprender y detectar eventos y señales tempranas, útiles en dominios críticos como la salud [3,4]. Por lo cual, integrar ambas fuentes constituye una oportunidad para



2026 – “Año de la Grandeza Argentina”

“150° Aniversario de la Creación de la Escuela Normal Juan Pascual Pringles”

“50 años por la Memoria, la Verdad y la Justicia. Nunca más”

Universidad Nacional de San Luis
RECTORADO

mejorar el análisis y la extracción de información, pudiendo incluso potenciar los enfoques actuales de la IA [5,6].

Una de las estrategias más relevantes en la actualidad para representar información de manera estructurada es el uso de los Grafos de Conocimiento (Knowledge Graphs o KGs), los cuales permiten modelar entidades y sus relaciones semánticas, consolidando grandes volúmenes de datos en una estructura unificada y coherente [7]. Esta representación permite la integración de fuentes heterogéneas, mejora la interpretabilidad de la información y posibilita el razonamiento automatizado [8]. Por su parte, el auge de los Grandes Modelos de Lenguaje (Large Language Models o LLMs) ha revolucionado el campo de la IA, al ofrecer notables habilidades para abordar una amplia variedad de tareas en múltiples dominios. En el ámbito médico, por ejemplo, han sido utilizados para evaluar los beneficios y riesgos de su aplicación [9], generar resúmenes de enfermedades [10], identificar interacciones farmacológicas [11], detectar pacientes en riesgo [12], entre otros. Sin embargo, a pesar de sus llamativas habilidades emergentes [13], estos modelos presentan importantes desafíos y limitaciones [14,15], como el elevado costo computacional, preocupaciones acerca de la privacidad y la transparencia, y la frecuente aparición de alucinaciones y sesgos. Modelos como Gemini [16], GPT [17] y DeepSeek [18], accesibles únicamente a través de APIs, junto con otros que pueden ejecutarse en entornos locales como Llama-2 [19] y Mistral [20], conforman el núcleo de gran parte de las investigaciones actuales. Áreas como la interpretabilidad y el razonamiento buscan comprender y explicar las decisiones tomadas por los LLMs, promoviendo el desarrollo de técnicas como RetrievalAugmented Generation (RAG) [21], Chain-of-Thought (CoT) [22] y Graph-of-Thought (GoT) [23]. Estas estrategias permiten combinar la capacidad generativa de los LLMs con conocimiento externo, mitigando así su dependencia de los datos y patrones aprendidos durante el entrenamiento. En este contexto, los KGs surgen como un enfoque atrayente para incorporar conocimiento estructurado a los LLMs. Al representar la información mediante tripletas ‘(entidad-A, relación, entidad-B)’, los KGs facilitan la adaptación de estos modelos para generar respuestas más precisas, justificadas y comprensibles. A su vez, dada la complejidad y el costo para construir y mantener los KGs, los LLMs pueden ser aprovechados para descubrir nuevas entidades, resolver correferencias y extraer relaciones a partir de grandes volúmenes de texto [24, 25]. De esta forma, la combinación de los LLMs y KGs ofrece una estrategia prometedora para la comprensión y el descubrimiento de conocimiento, como así también para asistir en la toma de decisiones, particularmente relevante en dominios críticos como la salud, donde la confiabilidad de las soluciones es un aspecto esencial [26].

La motivación para realizar esta pasantía de posgrado se basa en el tema de investigación doctoral del pasante, centrado en la detección anticipada de riesgos en la Web en contextos vinculados a la salud, particularmente en problemáticas de salud mental (como depresión, anorexia y suicidio), y recientemente en el estudio de enfermedades ginecológicas. El objetivo es aplicar técnicas de PLN y AA para analizar datos textuales provenientes de la Web en el dominio de estudio del cáncer de ovario (CO), una de las neoplasias ginecológicas más letales a nivel mundial [27, 28]. A pesar de los avances en el diagnóstico y tratamiento, esta enfermedad suele detectarse en etapas avanzadas, lo que limita las opciones terapéuticas y reduce la tasa de supervivencia a 5 años por



2026 – “Año de la Grandeza Argentina”

“150° Aniversario de la Creación de la Escuela Normal Juan Pascual Pringles”

“50 años por la Memoria, la Verdad y la Justicia. Nunca más”

Universidad Nacional de San Luis
RECTORADO

debajo del 40% [29]. Si bien la cirugía y la quimioterapia con platino pueden ser inicialmente efectivas, más del 70% de las pacientes desarrollan resistencia, conduciendo a un cuadro extremadamente difícil de tratar [30, 31], lo que subraya la necesidad urgente de comprender en mayor profundidad los mecanismos de quimioresistencia (MQR). Para esto, se propone construir un KG integrado por evidencia científica estructurada y testimonios de pacientes, con el fin de enriquecer la comprensión del CO y encontrar nuevas opciones diagnósticas y terapéuticas, centrándose en los MQR. De esta manera, la pasantía se desarrollará como un trabajo interdisciplinario que contribuirá tanto a las Ciencias de la Computación como a la Biomedicina.

OBJETIVOS:

Desarrollar un KG que integre evidencia científica con experiencias de pacientes extraídas de la Web. Se utilizarán técnicas avanzadas de PLN y AA para recopilar, extraer y analizar datos de diversas fuentes, con el fin de construir de forma incremental un KG sobre el CO y los MQR.

Desarrollar métodos explicativos y exploratorios para la comprensión del CO y los MQR. Con el KG obtenido, se aplicarán enfoques predictivos de análisis de datos y razonamiento para comprender de manera integral esta patología y explorar nuevas hipótesis que permitan identificar posibles intervenciones diagnósticas y terapéuticas innovadoras.

PROGRAMA DETALLADO:

Introducción a la Biomedicina, al CO y a los MQR.

Revisión de la literatura científica relevante.

Identificación y selección de fuentes de información en la Web.

Definición de los ejes de interés (CO y MQR, tratamientos actuales, efectos adversos, consecuencias en la salud física, social y mental de las pacientes, entre otros).

Recolección y preprocesamiento de datos utilizando técnicas de PLN.

Extracción de información de fuentes científicas y redes sociales.

Construcción inicial del KG integrando datos heterogéneos.

Evaluación continua del KG (calidad, cobertura y consistencia).

Análisis del conocimiento adquirido.

Enriquecimiento semántico del KG mediante ontologías biomédicas.



2026 – “Año de la Grandeza Argentina”

“150° Aniversario de la Creación de la Escuela Normal Juan Pascual Pringles”

“50 años por la Memoria, la Verdad y la Justicia. Nunca más”

Universidad Nacional de San Luis
RECTORADO

Desarrollo de modelos predictivos y aplicación de técnicas de razonamiento sobre el KG. Detección de patrones y relaciones complejas.

Análisis del conocimiento generado y su impacto en la comprensión del dominio. Identificación de nuevas hipótesis, conceptos y potenciales tratamientos.

Evaluación y presentación de resultados.

METODOLOGÍA SEGUIDA EN EL DESARROLLO DE LA PASANTÍA:

Introducción al dominio y definición del marco de trabajo: se abordarán los fundamentos de la Biomedicina, centrándose en el CO y los MQR. Se llevarán a cabo reuniones con el grupo de investigación para definir los ejes de interés, que incluirán: el estado actual del conocimiento sobre los MQR en CO, los desafíos clínicos, su relación con otras enfermedades, factores físicos, mentales y sociales que afectan a las pacientes, tratamientos actuales, biomarcadores, tipos de resistencia, causas y efectos de la quimioresistencia y alternativas terapéuticas. Esto permitirá delimitar el alcance del estudio y guiará la recopilación de información. Además, se realizará un análisis de las fuentes de datos disponibles, incluyendo literatura científica y testimonios de pacientes provenientes de redes sociales.

Recopilación y extracción de información: se implementará una herramienta de software para recopilar y estructurar datos provenientes de la Web. Esto incluirá: el desarrollo de scrapers para extraer información de diversas fuentes; el preprocesamiento de los datos para la limpieza, normalización y estructuración del contenido; y la aplicación de técnicas de PLN y AA para realizar un análisis exploratorio de los datos, clasificar y categorizar la información de acuerdo con los ejes de interés y diseñar un método de integración para los datos heterogéneos.

Construcción y validación del Grafo de Conocimiento (KG): Una vez recopilada la información, se procederá a la construcción de un KG inicial que integre de manera consistente los datos heterogéneos obtenidos. Este proceso será iterativo, abordando tareas complejas como la extracción de entidades y relaciones, así como la resolución de correferencias, donde se aprovechará el potencial de los LLMs para asistir en la comprensión del lenguaje biomédico y en la desambiguación semántica. Se llevarán a cabo evaluaciones continuas para garantizar aspectos como la calidad (¿El KG abarca todos los aspectos relevantes del CO y de los MQR?), la cobertura (¿Incluyen todos los ejes de interés definidos?) y la consistencia (¿La información es correcta y confiable? ¿Existen errores de integración, semánticos, sintácticos o lógicos?). Estas evaluaciones se complementarán con revisiones realizadas por expertos en el dominio, quienes validarán la información contenida en el KG. Además, se podrá llevar a cabo un enriquecimiento semántico del KG mediante la incorporación de datos provenientes de ontologías biomédicas reconocidas, como DisGeNET [32] (genes y enfermedades), DrugBank [33] (fármacos y mecanismos de acción), UniProt [34] (proteínas), ChEBI [35] (compuestos químicos), y CTD



2026 – “Año de la Grandeza Argentina”

“150° Aniversario de la Creación de la Escuela Normal Juan Pascual Pringles”

“50 años por la Memoria, la Verdad y la Justicia. Nunca más”

Universidad Nacional de San Luis
RECTORADO

[36] (interacción entre sustancias químicas, genes y enfermedades). Este proceso iterativo de integración y validación permitirá refinar progresivamente el KG hacia una representación más contextualizada del CO y los MQR.

Análisis de relaciones y patrones en el KG: a partir del KG construido, se aplicarán técnicas de AA y razonamiento para detectar patrones, relaciones y correlaciones que profundicen los ejes de interés del estudio. En particular, se emplearán métodos que integran el poder de los LLMs con el conocimiento estructurado de los KGs. Se explorará el uso de técnicas como RAG y CoT, para generar contenido con información relevante y razonamientos lógicos en múltiples pasos; Knowledge Augmented Generation (KAG) [37] y GraphRAG [38], orientadas al razonamiento híbrido entre los LLMs y los KGs; e incluso enfoques de razonamiento multimodal [39], que pueden ser útiles para descubrir conexiones no evidentes y patrones complejos en los datos. Los análisis previstos incluyen: el estudio de la relación entre los MQR y los tratamientos actuales, el impacto de la salud física sobre el estado mental y social de las pacientes, la identificación de enfermedades similares al CO según criterios biomédicos y la detección de nuevas líneas terapéuticas emergentes. Este enfoque no solo contribuirá a generar conocimiento fundado sobre esta patología, sino que también permitirá identificar nuevas hipótesis y líneas de investigación, para el diagnóstico y desarrollo de nuevas opciones terapéuticas centradas en los MQR.

Evaluación de resultados y reporte final: se discutirá el conocimiento adquirido con el grupo de investigación para evaluar su impacto en la comprensión del CO y los MQR, así como en la identificación de nuevas hipótesis y estrategias terapéuticas. Finalmente, se elaborará un informe final y se discutirá sobre trabajos futuros y la publicación de los resultados en congresos y revistas científicas.

CRONOGRAMA

MES	Actividades
16 al 30 de junio	Introducción a la Biomedicina, al CO y a los MQR. Reuniones iniciales con el grupo de investigación. Seminario del pasante: presentación del tema del doctorado y objetivos de la pasantía.



2026 – “Año de la Grandeza Argentina”

“150° Aniversario de la Creación de la Escuela Normal Juan Pascual Pringles”

“50 años por la Memoria, la Verdad y la Justicia. Nunca más”

Universidad Nacional de San Luis
RECTORADO

Julio	<p>Búsqueda bibliográfica relacionada con PLN/AA aplicadas a Biomedicina.</p> <p>Revisión y selección de posibles fuentes de información (científicas y redes sociales). Discusión de objetivos y definición de ejes de interés. Participación en seminarios del grupo. Nota: este mes puede incluir un receso de hasta dos semanas, sujeto a la agenda del grupo de investigación y a las políticas de la Universidad donde se llevará a cabo la pasantía.</p>
Agosto	<p>Diseño del pipeline para la recolección, extracción y procesamiento de la información. Desarrollo de scrapers para distintas fuentes. Integración de fuentes heterogéneas. Construcción inicial del KG.</p>
Septiembre	<p>Evaluación de calidad y cobertura del KG, chequeo de consistencias y detección de errores. Enriquecimiento semántico y refinamiento del KG usando ontologías biomédicas. Validación con expertos del dominio.</p>
Octubre	<p>Diseño y desarrollo de un prototipo para aplicar modelos predictivos usando KGs. Primeras pruebas de inferencias. Ajustes sobre los modelos y el KG en función de los resultados preliminares. Aplicación de técnicas de razonamientos sobre el KG. Detección de patrones y relaciones relevantes.</p>
Noviembre	<p>Revisión conjunta de resultados con expertos.</p> <p>Análisis del conocimiento generado y su impacto en la comprensión del dominio.</p> <p>Identificación de nuevas hipótesis, conceptos y potenciales tratamientos.</p>
Diciembre Fin: 31 de diciembre	<p>Redacción del informe final.</p> <p>Presentación de resultados al grupo de investigación. Discusión sobre trabajos futuros y publicación en congresos y revistas científicas.</p>



2026 – “Año de la Grandeza Argentina”

“150° Aniversario de la Creación de la Escuela Normal Juan Pascual Pringles”

“50 años por la Memoria, la Verdad y la Justicia. Nunca más”

Universidad Nacional de San Luis
RECTORADO

BIBLIOGRAFÍA:

- [1] Zhang J, Meng Y. Research on Media Text Translation Based on Information Retrieval. *International Journal of E-Collaboration* 2025;21:1–14. <https://doi.org/10.4018/IJeC.370405>.
- [2] Zachlod C, Samuel O, Ochsner A, Werthmüller S. Analytics of social media data – State of characteristics and application. *Journal of Business Research* 2022;144:1064–76. <https://doi.org/10.1016/j.jbusres.2022.02.016>.
- [3] Correia RB, Wood IB, Bollen J, Rocha LM. Mining Social Media Data for Biomedical Signals and Health-Related Behavior. *Annu Rev Biomed Data Sci* 2020;3:433–58. <https://doi.org/10.1146/annurevbiodatasci-030320-040844>.
- [4] Park A, Conway M. Tracking Health Related Discussions on Reddit for Public Health Applications. *AMIA Annu Symp Proc* 2018;2017:1362–71.
- [5] Chen Q, Zhu X, Ling Z-H, Inkpen D, Wei S. Neural Natural Language Inference Models Enhanced with External Knowledge. In: Gurevych I, Miyao Y, editors. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Melbourne, Australia: Association for Computational Linguistics; 2018, p. 2406–17. <https://doi.org/10.18653/v1/P18-1224>.
- [6] Liu S, Cai T, Li X. Representation-Enhanced Neural Knowledge Integration with Application to LargeScale Medical Ontology Learning 2024. <https://doi.org/10.48550/arXiv.2410.07454>.
- [7] Ji S, Pan S, Cambria E, Marttinen P, Yu PS. A Survey on Knowledge Graphs: Representation, Acquisition and Applications. *IEEE Trans Neural Netw Learning Syst* 2022;33:494–514. <https://doi.org/10.1109/TNNLS.2021.3070843>.
- [8] Neural, symbolic and neural-symbolic reasoning on knowledge graphs - ScienceDirect n.d. <https://www.sciencedirect.com/science/article/pii/S2666651021000061> (accessed April 9, 2025).
- [9] Lee P, Bubeck S, Petro J. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. *New England Journal of Medicine* 2023;388:1233–9. <https://doi.org/10.1056/NEJMs2214184>.
- [10] Nayak A, Alkaitis MS, Nayak K, Nikolov M, Weinfurt KP, Schulman K. Comparison of History of Present Illness Summaries Generated by a Chatbot and Senior Internal Medicine Residents. *JAMA Intern Med* 2023;183:1026–7. <https://doi.org/10.1001/jamainternmed.2023.2561>.
- [11] Han K, Cao P, Wang Y, Xie F, Ma J, Yu M, et al. A Review of Approaches for Predicting Drug-Drug Interactions Based on Machine Learning. *Front Pharmacol* 2021;12:814858. <https://doi.org/10.3389/fphar.2021.814858>.



2026 – “Año de la Grandeza Argentina”

“150° Aniversario de la Creación de la Escuela Normal Juan Pascual Pringles”

“50 años por la Memoria, la Verdad y la Justicia. Nunca más”

Universidad Nacional de San Luis
RECTORADO

- [12] Beaulieu-Jones BK, Yuan W, Brat GA, Beam AL, Weber G, Ruffin M, et al. Machine learning for patient risk stratification: standing on, or looking over, the shoulders of clinicians? NPJ Digit Med 2021;4:62. <https://doi.org/10.1038/s41746-021-00426-3>.
- [13] Wei J, Tay Y, Bommasani R, Raffel C, Zoph B, Borgeaud S, et al. Emergent Abilities of Large Language Models 2022. <https://doi.org/10.48550/arXiv.2206.07682>.
- [14] Hoffmann J, Borgeaud S, Mensch A, Buchatskaya E, Cai T, Rutherford E, et al. Training ComputeOptimal Large Language Models 2022. <https://doi.org/10.48550/arXiv.2203.15556>.
- [15] Li H, Ai Q, Chen J, Dong Q, Wu Z, Liu Y, et al. BLADE: Enhancing Black-box Large Language Models with Small Domain-Specific Models 2024.
- [16] Team G, Anil R, Borgeaud S, Alayrac J-B, Yu J, Soricut R, et al. Gemini: A Family of Highly Capable Multimodal Models 2024. <https://doi.org/10.48550/arXiv.2312.11805>.
- [17] Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, et al. Language Models are Few-Shot Learners. Advances in Neural Information Processing Systems, vol. 33, Curran Associates, Inc.; 2020, p. 1877–901.
- [18] Liu A, Feng B, Xue B, Wang B, Wu B, Lu C, et al. Deepseek-v3 technical report. arXiv Preprint arXiv:241219437 2024.
- [19] Touvron H, Martin L, Stone K, Albert P, Almahairi A, Babaei Y, et al. Llama 2: Open Foundation and Fine-Tuned Chat Models 2023. <https://doi.org/10.48550/arXiv.2307.09288>.
- [20] Jiang AQ, Sablayrolles A, Roux A, Mensch A, Savary B, Bamford C, et al. Mixtral of Experts 2024. <https://doi.org/10.48550/arXiv.2401.04088>.
- [21] Wang X, Wei J, Schuurmans D, Le Q, Chi E, Narang S, et al. Self-Consistency Improves Chain of Thought Reasoning in Language Models 2023. <https://doi.org/10.48550/arXiv.2203.11171>.
- [22] Wei J, Wang X, Schuurmans D, Bosma M, Ichter B, Xia F, et al. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models 2023. <https://doi.org/10.48550/arXiv.2201.11903>.
- [23] Besta M, Blach N, Kubicek A, Gerstenberger R, Podstawski M, Gianinazzi L, et al. Graph of Thoughts: Solving Elaborate Problems with Large Language Models. AAAI 2024;38:17682–90. <https://doi.org/10.1609/aaai.v38i16.29720>.
- [24] Luo L, Li Y-F, Haffari G, Pan S. Normalizing Flow-based Neural Process for Few-Shot Knowledge Graph Completion. Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, New York, NY, USA: Association for Computing Machinery; 2023, p. 900–10. <https://doi.org/10.1145/3539618.3591743>.



2026 – “Año de la Grandeza Argentina”

“150° Aniversario de la Creación de la Escuela Normal Juan Pascual Pringles”

“50 años por la Memoria, la Verdad y la Justicia. Nunca más”

Universidad Nacional de San Luis
RECTORADO

[25] Zhong L, Wu J, Li Q, Peng H, Wu X. A Comprehensive Survey on Automatic Knowledge Graph Construction 2023. <https://doi.org/10.48550/arXiv.2302.05019>.

[26] Pan S, Luo L, Wang Y, Chen C, Wang J, Wu X. Unifying Large Language Models and Knowledge Graphs: A Roadmap. *IEEE Trans Knowl Data Eng* 2024;36:3580–99. <https://doi.org/10.1109/TKDE.2024.3352100>.

[27] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin* 2022;72:7–33. <https://doi.org/10.3322/caac.21708>.

[28] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021;71:209–49. <https://doi.org/10.3322/caac.21660>.

[29] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin* 2022;72:7–33. <https://doi.org/10.3322/caac.21708>.

[30] Christie EL, Bowtell DDL. Acquired chemotherapy resistance in ovarian cancer. *Ann Oncol* 2017;28:viii13–5. <https://doi.org/10.1093/annonc/mdx446>.

[31] Pignata S, Pisano C, Di Napoli M, Cecere SC, Tambaro R, Attademo L. Treatment of recurrent epithelial ovarian cancer. *Cancer* 2019;125 Suppl 24:4609–15. <https://doi.org/10.1002/cncr.32500>.

[32] Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, et al. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res* 2020;48:D845–55. <https://doi.org/10.1093/nar/gkz1021>.

[33] DrugBank 6.0: the DrugBank Knowledgebase for 2024 - PubMed n.d. <https://pubmed.ncbi.nlm.nih.gov/37953279/> (accessed April 8, 2025).

[34] UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* 2023;51:D523–31. <https://doi.org/10.1093/nar/gkac1052>.

[35] Hastings J, Owen G, Dekker A, Ennis M, Kale N, Muthukrishnan V, et al. ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Res* 2016;44:D1214-1219. <https://doi.org/10.1093/nar/gkv1031>.

[36] Davis AP, Wiegiers TC, Johnson RJ, Sciaky D, Wiegiers J, Mattingly CJ. Comparative Toxicogenomics Database (CTD): update 2023. *Nucleic Acids Res* 2023;51:D1257–62. <https://doi.org/10.1093/nar/gkac833>.



2026 – “Año de la Grandeza Argentina”

“150° Aniversario de la Creación de la Escuela Normal Juan Pascual Pringles”

“50 años por la Memoria, la Verdad y la Justicia. Nunca más”

Universidad Nacional de San Luis
RECTORADO

[37] Liang L, Sun M, Gui Z, Zhu Z, Jiang Z, Zhong L, et al. KAG: Boosting LLMs in Professional Domains via Knowledge Augmented Generation 2024. <https://doi.org/10.48550/arXiv.2409.13731>.

SISTEMA DE EVALUACIÓN:

Seminario con presentación de resultados obtenidos durante la pasantía.

FINANCIAMIENTO DE LA PASANTÍA

COSTOS: La pasantía no requiere gastos significativos. Se prevé el uso de equipamiento personal para llevar a cabo las actividades de la pasantía.

FUENTES DE FINANCIAMIENTO: La Universidad que recibe al pasante proveerá los recursos necesarios para la realización de la pasantía, que incluyen acceso a Internet, espacios de trabajo adecuados y acceso a datos y/o plataformas que el grupo de investigación utiliza en sus proyectos. No se prevé la necesidad de financiamiento adicional.

Hoja de firmas