Entity Disambiguation and Linking in Catullan Poetry: Toward an Enriched Digital Edition and a Knowledge Graph

Carlos Javier Nusch [1,2] 0000-0003-1715-4228 Luciana Tanevitch[3] 0000-0002-5322-9314 Diego Torres[3] 0000-0001-7533-0133, Gimena del Rio Riande [4] 0000-0002-8997-5415 Leticia Cecilia Cagnina [5] 0000-0001-7825-2927 Leandro Antonelli [3,6] 0000-0003-1388-0337 Marcelo Luis Errecalde [5] 0000-0001-5605-8963

[1] Dirección PREBI-SEDICI, Universidad Nacional de La Plata, Argentina

[2] CESGI, Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, Argentina

[3] LIFIA, Facultad de Informática, Universidad Nacional de La Plata, Argentina

[4] IBICRIT, Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

[5] LIDIC, Facultad de Ciencias Físico Matemáticas y Naturales, Universidad Nacional de San Luis, Argentina

[6] CAETI, Facultad de Tecnología Informática, Universidad Abierta Interamericana, Argentina

carlosnusch@prebi.unlp.edu.ar

Abstract. This work forms part of a broader project aimed at studying love discourse through computational methods. In this paper, we focus on a corpus of Classical Latin poetry, namely the works of Gaius Valerius Catullus, whose semantic and historical density makes the identification and disambiguation of references to people, places, and groups especially relevant. Given the nature of poetic texts, such entities often appear in highly allusive contexts, with morphological variation, semantic ambiguity, and stylistic devices such as hyperbaton, which make their automatic processing particularly challenging. To address this problem, we comparatively evaluate two entity linking approaches over Wikidata: a modular pipeline based on named entity recognition, candidate generation, and semantic ranking, and a generative approach based on GPT-4.1, used both for mention detection and for final entity selection. The evaluation relies on a manually curated ground truth and distinguishes between named entity recognition and semantic linking. The results show complementary behaviors: the traditional pipeline offers greater traceability and broader candidate-space coverage, whereas the generative approach achieves higher precision in mention detection and better performance in final entity selection, albeit with a more conservative strategy. Overall, the paper provides a comparative evaluation of methods for Latin poetry and lays the groundwork for enriching an XML-TEI digital edition with auditable identifiers and for Tdesigning a literary knowledge graph.

Keywords: named entity recognition; entity linking; large language models; Latin poetry; Gaius Valerius Catullus; Wikidata; XML-TEI; knowledge graph.

Desambiguación y enlazado de entidades en la poesía de Catulo: hacia una edición digital enriquecida y un grafo de conocimiento

Resumen. Este trabajo se inscribe en un proyecto más amplio orientado al estudio del discurso amoroso mediante herramientas computacionales. En esta oportunidad, nos concentramos en un corpus de poesía latina del período clásico, la obra de Cayo Valerio Catulo, cuya densidad semántica e histórica vuelve especialmente relevante la identificación y desambiguación de referencias a personas, lugares y grupos. Dada la naturaleza de los textos poéticos, estas entidades suelen aparecer en contextos altamente alusivos, con variación morfológica, ambigüedad semántica y recursos estilísticos como el hipérbaton, lo que dificulta su tratamiento automático. Para abordar este problema, evaluamos comparativamente dos enfoques de *entity linking* sobre Wikidata: un pipeline modular basado en reconocimiento de entidades, generación de candidatos y ranking semántico, y un enfoque generativo basado en GPT-4.1, utilizado tanto para la detección de menciones como para la selección final de entidades. La evaluación se apoya en un *ground truth* manual y distingue entre reconocimiento de entidades y enlazado semántico. Los resultados muestran comportamientos complementarios: el pipeline tradicional ofrece mayor trazabilidad y mejor cobertura del espacio de candidatos, mientras que el enfoque generativo alcanza mayor precisión en detección y mejor desempeño en la selección final de entidades, aunque con una estrategia más conservadora. En conjunto, el trabajo aporta una evaluación comparativa de métodos para poesía latina y sienta las bases para el enriquecimiento de una edición digital en XML-TEI con identificadores auditables y el diseño de un grafo de conocimiento literario.

Palabras clave: reconocimiento de entidades nombradas; enlazado de entidades; modelos de lenguaje de gran tamaño; poesía latina; Cayo Valerio Catulo; Wikidata; XML-TEI; grafo de conocimiento.

1 Introducción

El presente trabajo forma parte de un proyecto de investigación doctoral más amplio orientado al estudio del discurso amoroso a lo largo de la tradición occidental y surge de la discusión de las ideas de C. S. Lewis sobre la influencia del amor cortés y la literatura occitana en la configuración del imaginario amoroso moderno (Lewis, 1936; Nusch et al., 2024b, 2024a). En esta línea, se han identificado correspondencias entre la poesía latina del siglo I a.C. y la tradición occitana. El objetivo general de esta investigación consiste en detectar patrones textuales que permitan delimitar una posible continuidad en los tópicos amorosos desde la antigüedad. Para ello, se adopta un enfoque comparativo que combina métodos de lectura cercana con técnicas de lectura distante basadas en herramientas computacionales (Jockers, 2013; Moretti, 2013; Ramsay, 2011).

En el contexto de las humanidades digitales, el desarrollo reciente del Procesamiento del Lenguaje Natural (PLN) ha permitido avanzar en la automatización de tareas de anotación, curaduría y exploración de corpus complejos. En particular, su integración con estándares como XML-TEI posibilita la estructuración del texto y su enriquecimiento semántico mediante la incorporación de identificadores persistentes y vínculos a bases de conocimiento externas. En este marco, la detección de entidades nombradas y la posterior vinculación con recursos como Wikidata constituyen un paso clave hacia la construcción de ediciones digitales que funcionan no solo como objetos de lectura, sino como infraestructuras de datos capaces de soportar consultas complejas y modelado relacional.

Sin embargo, los textos líricos presentan una ambigüedad particular: la flexión morfológica, la libertad sintáctica, el hipérbaton y la superposición entre referencias históricas, geográficas y míticas dificultan tanto la detección como la interpretación de las entidades. Por ello, identificar una mención no basta; es necesario determinar su referente concreto en una base de conocimiento y evaluar la pertinencia del enlace según el contexto.

Estas dificultades se manifiestan claramente en el corpus de Cayo Valerio Catulo, que constituye el caso de estudio de este trabajo. Su poesía presenta una alta densidad de referencias a personas, lugares y colectivos, en un entramado donde convergen experiencia amorosa, alusividad literaria y tradición cultural. Entidades ambiguas o diversos gentilicios y topónimos requieren decisiones interpretativas que exceden la coincidencia superficial, y obligan a considerar tanto el contexto textual como las características de los recursos de autoridad utilizados.

Las tareas abordadas en este artículo pertenecen a un flujo de trabajo mayor orientado a la automatización parcial de procesos editoriales sobre poesía latina codificada en XML-TEI (Nusch et al., 2026). En particular, el estudio se centra en la detección, desambiguación y vinculación de entidades nombradas con recursos externos, no con el objetivo de reemplazar la intervención filológica, sino de evaluar en qué medida estas técnicas pueden asistirla, al facilitar la generación de candidatos, la reducción de tareas repetitivas y la revisión sistemática de los resultados. En este marco, el problema ya no consiste únicamente en determinar si una entidad puede ser detectada y enlazada, sino también en establecer qué tipo de arquitectura resulta más adecuada para asistir al trabajo editorial en un dominio como el de la poesía latina.

En este contexto, el trabajo realiza cinco aportes principales. En primer lugar, propone y compara dos estrategias de entity linking aplicadas a poesía latina: un enfoque tradicional, modular y auditable, basado en reconocimiento de entidades, generación de candidatos y ranking semántico, y un enfoque generativo apoyado en un large language model, como recurso para abordar la detección y la desambiguación. En segundo lugar, construye un *ground truth* humano específico para el corpus de Catulo, que permite evaluar de manera controlada tanto el reconocimiento de entidades como el enlazado semántico. En tercer lugar, ofrece una evaluación diferenciada de ambos

paradigmas, distinguiendo entre errores de detección y errores de vinculación. En cuarto lugar, analiza las fortalezas y limitaciones de cada enfoque desde una perspectiva a la vez técnica, filológica y de edición textual. Finalmente, proyecta estos resultados hacia la integración entre XML-TEI y Wikidata, con vistas a una edición digital enriquecida y a la futura construcción de un grafo de conocimiento literario.

2 Estado del arte y marco conceptual

El *entity linking* puede definirse como la tarea de identificar menciones de entidades en un texto y vincularlas con una entrada específica en una base de conocimiento externa. A diferencia del reconocimiento de entidades nombradas (NER), que detecta y clasifica menciones en categorías generales como personas, lugares o grupos, el *entity linking* incorpora una etapa de desambiguación semántica orientada a determinar a qué entidad concreta corresponde cada mención (Kolitsas et al., 2018; Nadeau & Sekine, 2007; Wei Shen et al., 2015). Ambas tareas son complementarias: una entidad puede ser correctamente detectada, pero incorrectamente vinculada o no encontrar correspondencia en la base de conocimiento.

La aplicación de estas técnicas a corpora literarios y, en particular, a materiales de humanidades digitales, presenta dificultades adicionales respecto de dominios más estandarizados. La ambigüedad léxica, los usos figurados, la intertextualidad y la polisemia dificultan tanto la detección como la desambiguación, ya que una misma forma puede referirse a una persona histórica, un personaje mítico, un lugar o una personificación (Elson, Dames & McKeown, 2010; Bamman, Underwood & Smith, 2014).

Estas dificultades se intensifican en corpora históricos en latín, la flexión nominal, la variación ortográfica, la libertad sintáctica y la distancia cultural complican la alineación entre texto y base de conocimiento. Además, muchas entidades pertenecen a universos míticos o geográficos antiguos que no siempre están representados de forma homogénea en recursos contemporáneos (Sommerschield et al., 2023).

En relación con el latín, se ha señalado anteriormente la escasez de datos anotados y la complejidad morfológica como obstáculos principales para el NER (Erdmann et al., 2016), mientras que enfoques supervisados han mostrado resultados prometedores en corpus históricos (Aguilar et al., 2016). Más recientemente, modelos contextualizados como LatinBERT han mejorado el desempeño en tareas lingüísticas relevantes (Bamman & Burns, 2020), y herramientas como LatinCy han alcanzado niveles cercanos al 90 % en tareas de NER (Burns, 2023). Paralelamente, el proyecto LiLa: Linking Latin constituye una referencia central para la integración de recursos lingüísticos latinos mediante Linked Open Data. Desarrollado en el CIRCSE de la Università Cattolica del Sacro Cuore, el proyecto propone una base de conocimiento orientada a conectar recursos léxicos, corpus anotados y herramientas de PLN para el latín a partir de un principio articulador: el lema como nodo de interoperabilidad. El

proyecto permite pensar la relación entre forma textual, análisis lingüístico y recursos externos de manera estructurada y reutilizable (Mambrini & Passarotti, 2023; Passarotti et al., 2020). En el contexto argentino pueden mencionarse antecedentes afines en el cruce entre Humanidades Digitales, edición académica, anotación semántica e inteligencia artificial aplicada a corpus antiguos. Por un lado, las iniciativas vinculadas con HD LAB / IIBICRIT / CONICET ofrecen un marco local relevante para pensar ediciones digitales abiertas, anotación, XML-TEI y reutilización de datos en clave de minimal computing (Riande, 2022). Por otro lado, LUMERA constituye un antecedente reciente en la aplicación de técnicas de IA y PLN, especialmente en relación con la detección de conexiones semánticas y morfosintácticas en corpus filosóficos tardoantiguos (Riesgo, 2024).

En los últimos años, el desarrollo de large language models ha introducido un nuevo escenario para tareas tradicionalmente resueltas mediante pipelines modulares. Frente a arquitecturas que separan detección de menciones, generación de candidatos y ranking, los modelos generativos permiten abordar parte de este proceso de forma más integrada, utilizando instrucciones en lenguaje natural y razonamiento contextual (Cao et al., 2021; Xiao et al., 2023). Sin embargo, estos enfoques también plantean nuevos problemas, como la menor trazabilidad de las decisiones, la dificultad para controlar la abstención y el riesgo de producir vinculaciones plausibles pero ontológicamente inadecuadas. Por ello, comparar ambos paradigmas permite evaluar no solo su desempeño, sino también su utilidad como apoyo a la curaduría filológica.

Entre las bases de conocimiento abiertas, Wikidata resulta especialmente adecuada porque ofrece identificadores persistentes, alias, descripciones y relaciones semánticas que permiten evaluar candidatos más allá de la coincidencia nominal (Hogan et al., 2021; Vrandečić & Krötzsch, 2014). Además, facilita la interoperabilidad con otros recursos dentro del ecosistema de datos enlazados (Bizer et al., 2009).

Por su parte, el estándar XML-TEI permite representar la estructura textual y documental con alto grado de detalle, al integrar múltiples capas de anotación y posibilitar la vinculación de entidades mediante atributos como @ref o @type (Pierazzo & Driscoll, 2016). La combinación entre TEI y bases como Wikidata permite articular la localización precisa de las menciones con identificadores externos estables y sienta las bases para la construcción de grafos de conocimiento que integren texto, entidades y relaciones semánticas (Ciotti et al., 2014).

3 Configuración experimental

3.1 Corpus y recursos

El caso de estudio está constituido por poemas de Cayo Valerio Catulo (Merrill, 1893), procesados a partir de archivos de texto plano y convertidos posteriormente a una representación estructurada en XML-TEI que provienen de la edición empleada

en el proyecto mayor[1]. Esta codificación conserva identificadores por poema, encabezado y segmentación por versos, lo que permite mantener la trazabilidad entre el texto fuente, la anotación editorial y los resultados del procesamiento automático. La unidad básica de análisis es el poema individual, tratado como documento autónomo mediante identificadores del tipo *Catulo_Carmen_005*.

Las entidades consideradas pertenecen a tres categorías: PERSON, que incluye personas históricas, personajes míticos o figuras individualizadas; LOC, que agrupa lugares geográficos; y NORP, que abarca pueblos, colectivos y gentilicios. Como base de conocimiento principal se utilizó Wikidata, aprovechando sus identificadores persistentes, etiquetas, descripciones, alias, propiedades tipológicas —en particular *instance of* (P31)— y relaciones de equivalencia, entre ellas *said to be the same as* (P460), para generar y evaluar candidatos. Como recurso de reconocimiento de entidades se empleó el modelo *la_core_web_lg* de LatinCy integrado en spaCy (Burns, 2023). Además, para el enfoque no generativo se utilizaron embeddings multilingües basados en LaBSE (Feng et al., 2022) con el objetivo de estimar similitud contextual entre la mención en latín y la información textual de los candidatos en Wikidata, generalmente en inglés. Su buen desempeño para tareas afines aplicadas al latín se ha documentado previamente (Craig et al., 2023).

3.2 Ground truth y protocolo de evaluación

La evaluación se apoyó en un *ground truth* manual construido a partir de varias ediciones comentadas, ediciones críticas y traducciones (Fordyce, 1961; Galán, 2013; Merrill, 1893; Soler Ruiz, 2000). La unidad de anotación fue la mención textual tal como aparece en el poema, independientemente de su forma —flexionada, gentilicio, teónimo[2] o referencia colectiva—. A cada mención se le asignó un tipo de entidad (PERSON, LOC, NORP) y, cuando existía una correspondencia clara en Wikidata, un identificador QID con su etiqueta asociada. Cuando la identificación resultaba ambigua o incierta, la mención se conservó sin QID para evitar decisiones artificiales y reflejar la dificultad real de la tarea. El recurso resultante reúne 723 menciones distribuidas en 108 poemas y se organizó en una tabla con las columnas *poem_id*, *mention*, *char_start*, *char_end*, *type*, *qid* y *wikidata_label_en*.

La anotación requirió criterios explícitos para casos problemáticos. Las menciones ambiguas se resolvieron en función del contexto poético inmediato; por ejemplo, *Lesbia* fue tratada como entidad personal en función de su rol en el corpus, aun cuando pudiera asociarse formalmente a un topónimo o gentilicio. Las entidades sin correspondencia clara en Wikidata se mantuvieron sin QID, y las referencias míticas,

[1] https://github.com/KharolusIII/digital-edition-draft
[2] Un teónimo es el nombre propio de una deidad o dios

geográficas o colectivas se clasificaron según su función textual: teónimos y personificaciones como PERSON cuando actúan como entidades individualizadas, y gentilicios o pueblos como NORP. La evaluación se organizó en dos niveles: reconocimiento de entidades nombradas y *entity linking*. Para NER se emplearon *precision*, *recall* y F1-score. Para *entity linking*, la comparación se realizó sobre menciones alineadas con el *ground truth*: en el enfoque basado en ranking se utilizaron Accuracy@1, Recall@10 y MRR, mientras que en el enfoque generativo se reportó Accuracy@1 sobre la entidad final predicha.

4 Metodología de *entity linking*

El pipeline de *entity linking*, tradicionalmente, comprende: (1) detección de entidades[3], (2) generación de candidatos, (3) ranking (Moro et al., 2014; Sawada et al., 2026). En este trabajo comparamos dos estrategias sobre Wikidata. La primera, un enfoque basado en arquitecturas consolidadas para la tarea (Kolitsas et al., 2018; Moreno et al., 2017; Shen et al., 2021), sigue un pipeline modular y auditable, que combina NER para detectar menciones, consultas a Wikidata para recuperar candidatos y un ranking semántico basado en LaBSE. La segunda utiliza un enfoque generativo con un LLM, aplicado tanto a la detección de menciones como a la selección final de entidades entre los candidatos recuperados. Esta comparación permite evaluar las ventajas y los límites de un enfoque tradicional, más trazable, frente a un modelo generativo, potencialmente más flexible pero menos transparente.

4.1 Sistema basado en NER + LaBSE

Para la etapa de detección de entidades (1) se utilizó el modelo *la_core_web_lg* provisto por la biblioteca LatinCy. Este modelo permite identificar menciones en textos en latín, asignando etiquetas de tipo como PERSON, LOC y NORP, además de proveer la lematización de cada mención detectada. Esta información resulta fundamental para las etapas posteriores, ya que tanto la superficie original como el lema son utilizados en la generación de candidatos. En la etapa de generación de candidatos (2) se emplea la API de MediaWiki para recuperar posibles entidades de Wikidata. Esta decisión se basa en la necesidad de realizar búsquedas por

[3] Aunque la literatura distingue entre detección de menciones, resolución de correferencia y enlazado semántico, en este trabajo se restringe el alcance a la detección de entidades nombradas explícitas y a su posterior vinculación con una base de conocimiento externa, sin abordar la resolución de menciones nominales o pronominales. Esta distinción resulta especialmente relevante en el ámbito de la filología clásica: en el corpus de Catulo, la diferencia entre una entidad nombrada como *Lesbia* y una mención anafórica como *illa puella* (“aquella muchacha”) ilustra cómo una referencia puede ser central para la interpretación sin constituir una entidad nombrada directamente enlazable.

coincidencia de etiquetas a partir de distintas formas de la mención: superficie original, lema y variantes obtenidas mediante stemming. El uso de stemming permite aumentar la cobertura en casos donde la lematización, si bien es correcta, no coincide con ninguna entrada en Wikidata. Esto ocurre, por ejemplo, con gentilicios, donde la forma presente en el texto no tiene una entidad propia, pero sí puede vincularse a la entidad del lugar correspondiente. Dado que este tipo de casos está contemplado en el *ground truth*, se incorporan estas variantes para ampliar el conjunto de candidatos. A partir de este proceso se obtienen 15 candidatos por mención. Sobre estos candidatos se realiza una etapa de enriquecimiento mediante consultas a Wikidata, con el objetivo de recuperar información adicional como label, descripción y, en el caso de personas, propiedades asociadas a la edad (por ejemplo, fechas de nacimiento y muerte). En estas consultas se prioriza el uso del latín, pero también se incorpora el inglés, dado que una gran proporción de entidades en Wikidata carece de etiquetas o descripciones en latín, y esta información adicional resulta útil para la desambiguación. Finalmente, en la etapa de ranking (3), los candidatos se ordenan en función de un puntaje de relevancia. Para ello, se construyen representaciones vectoriales (embeddings) tanto de los candidatos (a partir de sus labels y descripciones) como de la mención en su contexto. La similitud entre estos vectores constituye una de las principales señales del modelo. Para este cálculo se utiliza un codificador basado en LaBSE, dado que las descripciones de los candidatos pueden estar en latín o en inglés, mientras que los contextos se encuentran en latín, por lo que resulta necesario un modelo multilingüe que permita comparar representaciones en distintos idiomas. A esta señal se le asigna un peso parcial dentro del puntaje total, que se complementa con otras características, tales como la compatibilidad entre el tipo de entidad en Wikidata y la etiqueta asignada por el modelo de detección, y la similitud semántica entre la mención y las distintas formas alternativas del candidato (*aliases* o *sameAs*). El resultado de esta etapa es una lista ordenada de hasta 10 candidatos por mención. Se selecciona como entidad final aquella ubicada en la primera posición del ranking, siempre que su puntaje supere un umbral fijo de 0.3. En caso contrario, la mención se considera no enlazada.

4.2 Sistema basado en GPT

Para el enfoque basado en modelos generativos, el pipeline mantiene la misma estructura general, pero delega tanto la detección de entidades como la desambiguación en un modelo de lenguaje (GPT 4.1[4]) con prompts específicos para cada etapa. En la etapa de detección de entidades (1), se utiliza un prompt que instruye al modelo a identificar menciones de tipo PERSON, LOC y NORP en textos en latín, y devuelve para cada una su forma superficial, lema, etiqueta y offsets a nivel

[4] https://developers.openai.com/api/docs/models/gpt-4.1

de carácter. Adicionalmente, en el caso de gentilicios (NORP), se solicita inferir la entidad geográfica o colectiva asociada (por ejemplo, “Romanus” → “Rome”). En la etapa de generación de candidatos (2) se utiliza el mismo procedimiento que en el enfoque anterior, basado en consultas a la API de MediaWiki. Finalmente, en la etapa de ranking (3), se utiliza nuevamente el modelo generativo en un esquema tipo RAG (Retrieval-Augmented Generation). Se construye un prompt que incluye el texto (truncado), las menciones detectadas y los candidatos asociados a cada una. El modelo debe seleccionar, para cada mención, la entidad más adecuada entre los candidatos provistos, respetando restricciones explícitas (por ejemplo, priorizar entidades históricamente relevantes para el corpus). También se incorporan reglas específicas para el tratamiento de gentilicios y se privilegian entidades de tipo grupo o pueblo o, en su defecto, el lugar asociado.

4.3 Decisiones heurísticas y reproducibilidad del pipeline

Las decisiones heurísticas incorporadas en el pipeline modular deben entenderse como criterios operativos orientados a producir un enlazado trazable y revisable, antes que como una optimización exhaustiva del espacio de parámetros. En particular, el sistema limita la recuperación inicial a un conjunto acotado de candidatos por mención; consulta Wikidata a partir de la forma superficial, el lema y variantes morfológicas simples del latín; contempla alternancias gráficas como *u/v* y *ae/e*; y enriquece los candidatos con etiquetas, descripciones, aliases multilingües, tipos *instance of* y, cuando están disponibles, fechas biográficas. La ordenación de candidatos combina similitud contextual basada en embeddings, compatibilidad entre la etiqueta NER y los tipos de Wikidata, similitud léxica entre mención, lema y aliases, una bonificación débil asociada a descripciones latinas y un ajuste histórico para entidades con cronología compatible con el autor. Finalmente, solo se acepta automáticamente al candidato mejor rankeado cuando supera un umbral mínimo de confianza. Estos criterios afectan directamente el equilibrio entre cobertura y precisión: configuraciones más permisivas pueden aumentar el número de enlaces propuestos, pero también el riesgo de asociaciones espurias, mientras que criterios más restrictivos reducen ese riesgo a costa de dejar más menciones sin enlazar. Por ello, los resultados presentados deben interpretarse como una primera configuración controlada y reproducible del sistema, no como una calibración definitiva. Un análisis sistemático de sensibilidad sobre el límite de candidatos, las ponderaciones del ranking y el umbral de aceptación constituye una línea prioritaria de trabajo futuro.

5 Resultados

Evaluamos tanto el *pipeline* basado en LaBSE como el enfoque basado en GPT en las tareas de detección de menciones y enlazado de entidades.

El primer enfoque alcanza un desempeño sólido en la detección de menciones, con una precisión de 0.92, un *recall* de 0.83 y un puntaje F1 de 0.85, lo que indica un equilibrio confiable entre precisión y cobertura en la identificación de menciones. En el linking, el rendimiento es más limitado, con un *Accuracy@1* de 0.29, un *Recall@10* de 0.40 y un MRR de 0.32. Estos resultados indican que, aunque la entidad correcta a veces aparece en posiciones altas del ranking, la obtención general de candidatos correctos sigue siendo limitada.

GPT-4.1 alcanza una precisión similar (0.94), con un *recall* algo más bajo (0.75), lo que da como resultado un puntaje F1 de 0.83. Esto refleja una estrategia más conservadora en la detección de menciones. En cuanto al linking, GPT alcanza una exactitud de 0.46, pero aún es un puntaje bajo.

Tabla 1. Resultados de evaluación de NER.

Evaluación de NER			
	Precision	**Recall**	**F1-Score**
LatinCy	0.92	0.83	0.85
GPT-4.1	0.94	0.75	0.83

Tabla 2. Resultados de evaluación de EL.

Evaluación de EL			
	Accurracy@1	**Recall@10**	**MRR**
LatinCy + LaBSE	0.29	0.40	0.32
GPT-4.1	0.46	-[5]	-

6 Análisis cualitativo y de errores

Los resultados cuantitativos permiten identificar patrones de error recurrentes y diferencias sistemáticas entre los enfoques evaluados. En la detección de menciones, se observan comportamientos opuestos. GPT-4.1 adopta una estrategia más conservadora, reflejada en una mayor tasa de falsos negativos, especialmente en entidades mitológicas (como *Iuppiter*, *Diana* o *Cybele*), geográficas antiguas (como *Pontus* o *Brixia*) y formas latinas menos frecuentes. Este patrón sugiere que el modelo

[5] No aplica, porque el modelo devuelve una selección final y no un ranking completo comparable

prioriza la coherencia semántica por sobre la cobertura. En contraste, LatinCy tiende a la sobredetección, es decir, identificar como entidades expresiones no referenciales, entre ellas verbos (*lugete*, *amabo*) y formas léxicas comunes, lo que indica una mayor dependencia de regularidades superficiales del texto y una menor capacidad de discriminación semántica. Asimismo, en esta etapa se registran tres perfiles principales de error: omisión de *spans*, *spans* espurios y asignación incorrecta del tipo de entidad.

Entre los casos de omisión, aparecen entidades presentes en el *ground truth* que el sistema no detecta, como *Tempe* (Carmen 64) y *Hamadryades* (Carmen 61). En otros casos, la mención es detectada, pero clasificada de manera incorrecta: *Sirmio* (Carmen 31), anotado como LOC, puede ser interpretado como PERSON, posiblemente influido por el contexto (*salve, o venusta Sirmio*), mientras que *Romuli* (Carmen 34), anotado como PERSON, pero identificado como NORP. También resultan problemáticos los adjetivos de procedencia y formas derivadas. En *Sestianus* (Carmen 44), por ejemplo, la forma deriva del nombre propio *Sestius* y remite a una relación personal o doméstica, no a un grupo o colectivo, aunque superficialmente pueda asemejarse a un gentilicio.

En la fase de *entity linking*, los errores se concentran en varios perfiles recurrentes. Un primer grupo corresponde a fallos en la generación de candidatos: la entidad correcta no siempre es recuperada o no está representada de forma adecuada en Wikidata. Un caso ilustrativo es *Cytorio / Cytore* (Carmen 4), donde el referente textual es el Monte Citorio, mientras que Wikidata privilegia la ciudad de *Cytorus*. Un segundo grupo incluye errores de ranking: el candidato correcto está presente, pero no es priorizado adecuadamente. En *Castoris* (Carmen 4), por ejemplo, el referente esperado es Pólux (*Castoris gemelle*), pero la forma superficial induce naturalmente la asociación con Cástor. Aquí la similitud superficial del significante resulta insuficiente puesto que la resolución depende de inferencias contextuales.

Un tercer grupo reúne ambigüedades ontológicas entre persona, lugar y entidad mítica. En *Tempe*, ambos modelos reconocen correctamente la mención como LOC, pero difieren en el enlazado: el pipeline tradicional la vincula con el municipio contemporáneo, mientras que GPT selecciona el *Vale of Tempe*, de carácter histórico-mitológico. Algo semejante ocurre con *Diae* (Carmen 64), que puede interpretarse como la isla Dia o como Naxos. En este caso existe una ambigüedad interpretativa genuina entre los filólogos especializados. A ello se suman desajustes ontológicos en Wikidata: *Orci* (Carmen 3) puede enlazarse como deidad, aunque en el poema funciona más bien como referido al inframundo, y *Sarapim* (Carmen 10) puede vincularse con *Serapis*, aunque en el contexto latino remite al templo (*ad Sarapim*). Se trata de enlaces razonables, pero no plenamente ajustados al uso poético.

Otro conjunto importante de problemas deriva de fenómenos morfológicos y literarios propios del latín poético. Formas como *Pontica* o *Hibera* (Carmen 29) pueden funcionar como gentilicios o como modificadores geográficos, mientras que

Troiugenum (Carmen 64) oscila entre NORP y LOC. Además, el corpus presenta numerosas personificaciones y superposiciones ontológicas. En *Austri* (Carmen 26), la anotación como PERSON responde a la personificación del viento y no al punto cardinal; en Carmen 64, entidades como *Scylla*, *Charybdis* y *Syrtis* activan simultáneamente dimensiones míticas y geográficas. También aparecen alusiones poéticas cuya resolución exige cadenas interpretativas complejas, como en *Memnonis* y *Arsinoes* (Carmen 66). Estos fenómenos muestran que la lematización y la similitud superficial son necesarias, pero insuficientes para una desambiguación robusta.
Finalmente, el latín poético introduce dificultades estructurales adicionales, como el hipérbaton y las entidades discontinuas o internamente complejas. En la secuencia *Cinna est Gaius* (Carmen 10), por ejemplo, los componentes del nombre propio aparecen separados por el verbo, lo que complica la identificación automática del *span*. Del mismo modo, expresiones como *Marrucine Asini* (Carmen 12), anotadas como PERSON, incluyen elementos cuya forma superficial remite a otras categorías, como los gentilicios. En conjunto, estos resultados muestran que una parte sustancial del error no proviene solo de limitaciones técnicas, sino de la interacción entre la flexión, la derivación, el orden libre, la polisemia y la granularidad de las ontologías utilizadas. En este sentido, el *entity linking* en poesía latina se configura como una tarea híbrida: parcialmente automatizable, pero todavía dependiente de conocimiento filológico para resolver ambigüedades estructurales, culturales y semánticas. Como líneas de mejora, se perfilan el desarrollo de reglas más finas para adjetivos de procedencia, una tipología más detallada, un mejor aprovechamiento del contexto en el ranking, la incorporación de recursos filológicos complementarios y el fortalecimiento de la revisión editorial como parte integral del proceso.

7 Discusión

Los resultados obtenidos permiten sostener que el *entity linking* en poesía latina es viable como herramienta de asistencia editorial, aunque no como procedimiento completamente automático, exhaustivo y filológicamente fiable. En este marco, el *ground truth* humano constituye un aporte metodológico central, ya que permite evaluar el desempeño sobre una base explícita y distinguir entre errores de detección, tipificación y enlazado.

La comparación entre ambos enfoques muestra comportamientos complementarios. El pipeline modular ofrece mayor trazabilidad: al separar detección, generación de candidatos y ranking, facilita la identificación de errores y la intervención editorial sobre etapas específicas del proceso. El enfoque basado en GPT, en cambio, muestra mayor capacidad para integrar contexto y producir decisiones plausibles en la selección final, aunque con una estrategia más conservadora en la detección y menor transparencia en sus decisiones. Por ello, el primero funciona mejor como herramienta

auditable y ajustable, mientras que el segundo opera como un asistente más flexible, pero menos controlable desde el punto de vista filológico.

Las principales limitaciones observadas derivan de la falta de alineación entre tres niveles: la representación del conocimiento, los modelos semánticos y las propiedades del dominio textual. Wikidata ofrece una cobertura amplia y valiosa, pero no fue diseñada específicamente para el mundo clásico ni para textos poéticos; por ello, algunas entidades aparecen modeladas de forma demasiado general, moderna o insuficientemente diferenciada. El impacto de Wikidata debe analizarse en términos de cobertura y granularidad: en algunos casos, el referente esperado no aparece entre los candidatos recuperados o carece de etiquetas y descripciones informativas; en otros, la entidad existe, pero su modelado no coincide plenamente con el uso poético, mítico, geográfico o colectivo activado por el texto. A esto se suma que modelos como LaBSE no han sido entrenados sobre latín poético, y que los modelos generativos, aunque integran mejor el contexto, no están optimizados para entity linking filológico.

Estos problemas se vuelven especialmente visibles en gentilicios, teónimos, personificaciones y alusiones poéticas, donde la dificultad no es solo técnica, sino también interpretativa. Por ello, los errores de linking no deben atribuirse únicamente al ranking, sino también a la forma en que la base de conocimiento organiza y describe los referentes antiguos. Esta observación refuerza la necesidad de combinar Wikidata con recursos especializados para el mundo clásico, como gazetteers antiguos, bases prosopográficas o infraestructuras lingüísticas como LiLa, y de mantener una capa de revisión filológica sobre los enlaces propuestos.

8 Hacia un grafo de conocimiento: TEI + Wikidata

En este trabajo, el *entity linking* no se entiende solo como una tarea de desambiguación, sino como una etapa intermedia entre la anotación textual en XML-TEI y una futura representación relacional del corpus. TEI aporta la organización textual, la localización precisa de las menciones y la trazabilidad editorial; Wikidata, por su parte, incorpora identificadores persistentes, descripciones, equivalencias y relaciones externas. Sobre esa base, el grafo podría modelar nodos como *Poema*, *Verso*, *Entidad*, *Tópico*, *Taxonomía* e *Ítem de Wikidata*, y relaciones como *mentionsEntity*, *hasTopic*, *linkedTo*, *sharesEntityWith* y *coOccursWith*, con la idea de articular la evidencia textual con una red de conocimiento más amplia.

La edición digital enriquecida resultante puede ser útil para distintos públicos. Para investigadores especializados, los identificadores externos permiten formular consultas sobre recurrencias, coocurrencias, redes de personajes, lugares y colectivos,

así como comparar patrones de referencia entre poemas o autores. Para estudiantes, la vinculación entre mención textual, información contextual y recursos externos facilita la lectura guiada de un corpus culturalmente denso, en el que muchos nombres propios, gentilicios y alusiones mitológicas no son transparentes. Para lectores generales, la edición puede funcionar como una interfaz de exploración que conecte el texto poético con mapas, descripciones, genealogías míticas y relaciones históricas. En consecuencia, el valor del entity linking no se limita a la asignación técnica de QIDs, sino que reside en su capacidad para transformar la edición en una infraestructura de consulta, enseñanza y reutilización del conocimiento.

```
  </teiHeader>
  <standOff>
    <spanGrp type="topics">
      <span xml:id="span-INVITACIÓN_AL_DISFRUTE_VITAL" from="1" to="6" ana="#INVITACIÓN_AL_DISFRUTE_VITAL"/>
      <span xml:id="span-BESOS" from="7" to="13" ana="#BESOS"/>
      <span xml:id="span-AOJAMIENTO_DE_LOS_AMANTES" from="7" to="13" ana="#AOJAMIENTO_DE_LOS_AMANTES"/>
      <span xml:id="span-ENVIDIA_HACIA_LOS_AMANTES" from="12" to="13" ana="#ENVIDIA_HACIA_LOS_AMANTES"/>
    </spanGrp>
  </standOff>
  <text>
    <body>
      <l n="1" ana="#INVITACIÓN_AL_DISFRUTE_VITAL">Vivamus, mea <placeName>Lesbia</placeName>, atque
      amemus,</l>
      <l n="2" ana="#INVITACIÓN_AL_DISFRUTE_VITAL">rumoresque senum severiorum</l>
      <l n="3" ana="#INVITACIÓN_AL_DISFRUTE_VITAL">omnes unius aestimemus assis.</l>
      <l n="4" ana="#INVITACIÓN_AL_DISFRUTE_VITAL">soles occidere et redire possunt:</l>
      <l n="5" ana="#INVITACIÓN_AL_DISFRUTE_VITAL">nobis, cum semel occidit brevis lux,</l>
      <l n="6" ana="#INVITACIÓN_AL_DISFRUTE_VITAL">nox est perpetua una dormienda.</l>
      <l n="7" ana="#BESOS #AOJAMIENTO_DE_LOS_AMANTES">da mi basia mille, deinde centum,</l>
      <l n="8" ana="#BESOS #AOJAMIENTO_DE_LOS_AMANTES">dein mille altera, dein secunda centum,</l>
      <l n="9" ana="#BESOS #AOJAMIENTO_DE_LOS_AMANTES">deinde usque altera mille, deinde centum,</l>
      <l n="10" ana="#BESOS #AOJAMIENTO_DE_LOS_AMANTES">dein, cum milia multa fecerimus,</l>
      <l n="11" ana="#BESOS #AOJAMIENTO_DE_LOS_AMANTES">conturbabimus illa, ne sciamus,</l>
      <l n="12" ana="#BESOS #AOJAMIENTO_DE_LOS_AMANTES #ENVIDIA_HACIA_LOS_AMANTES">aut ne quis malus invidere
      possit,</l>
      <l n="13" ana="#BESOS #AOJAMIENTO_DE_LOS_AMANTES #ENVIDIA_HACIA_LOS_AMANTES">cum tantum sciat esse
      basiorum.</l>
    </body>
  </text>
</TEI>
```

Fig. 1. Sector del archivo XML TEI donde se encuentra el cuerpo del texto y las etiquetas.

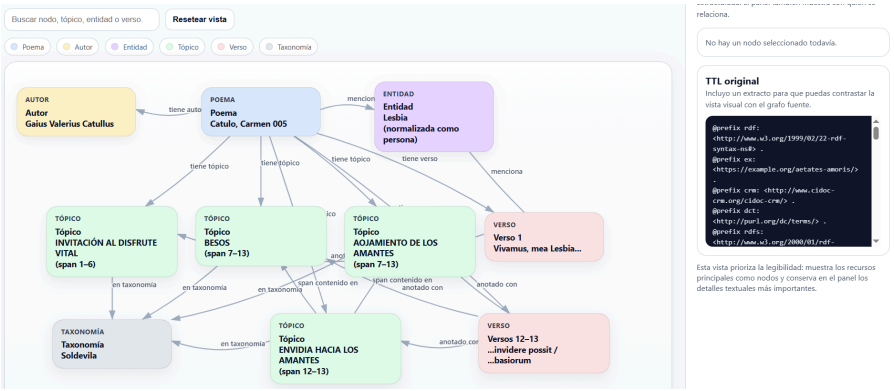

Fig. 2. Propuesta de grafo extraído de los archivos XML-TEI..

9 Conclusiones y trabajo futuro

Este trabajo se propuso evaluar la viabilidad del *entity linking* aplicado a poesía latina, tomando como caso de estudio un corpus de Catulo e integrando la tarea en un flujo más amplio de edición digital semánticamente enriquecida. El problema abordado no

fue únicamente el reconocimiento de entidades, sino su vinculación con una base de conocimiento externa en un dominio particularmente complejo, caracterizado por alta densidad alusiva, variación morfológica y ambigüedad entre referentes históricos, míticos y geográficos.

Para ello, se compararon dos enfoques de entity linking sobre Wikidata aplicados al corpus de Catulo: por un lado, un pipeline modular basado en reconocimiento de entidades, generación de candidatos y ranking semántico; por otro, un enfoque generativo basado en GPT-4.1, utilizado tanto para la detección de menciones como para la selección final de entidades. El *ground truth* construido constituye uno de los principales aportes metodológicos del artículo. Además de permitir una evaluación controlada, ofrece un recurso reutilizable para futuras comparaciones y mejora de modelos.

Desde la perspectiva de las humanidades digitales, el valor del enfoque no se limita a la asignación de QIDs, sino que radica en su capacidad para contribuir a una infraestructura semántica más amplia. El *entity linking* actúa como un puente entre la codificación XML-TEI y una futura representación en forma de grafo de conocimiento, en la que texto, entidades y relaciones puedan integrarse en una misma arquitectura.

En este sentido, el trabajo sugiere que la alternativa más prometedora no consiste necesariamente en optar de manera excluyente por uno u otro enfoque, sino en explorar arquitecturas híbridas. Un sistema de este tipo podría combinar la capacidad de control y auditabilidad del pipeline modular con la mayor flexibilidad contextual de los modelos generativos, aprovechando las fortalezas de cada paradigma en diferentes momentos del proceso editorial.

Como líneas de trabajo futuras, se plantea ampliar el corpus hacia otros autores de la poesía amorosa latina, como Tibulo y Propercio, con el fin de evaluar la generalización del enfoque. También será necesario profundizar el ajuste de prompts y restricciones para el modelo generativo, desarrollar estrategias más robustas de abstención y reconocimiento de NIL, y ampliar la comparación con enfoques recientes de entity linking, como modelos de recuperación densa y re-ranking, enfoques generativos autoregresivos y arquitecturas que combinan LLMs instruidos con mecanismos de recuperación. Esto permitiría situar con mayor precisión los resultados obtenidos frente a baselines actuales y evaluar su adaptación a un dominio filológico caracterizado por latín poético, ambigüedad semántica y cobertura desigual de Wikidata. También resultará relevante avanzar en una tipología más diferenciada que distinga con mayor precisión entre personas históricas, figuras míticas, deidades, lugares y colectivos. Finalmente, la integración más profunda entre el pipeline de entity linking y la edición TEI aparece como una línea de desarrollo central. Sobre esta base, la construcción de un grafo de conocimiento TEI + Wikidata se presenta como una continuación natural del trabajo, orientada a consolidar y explotar semánticamente las distintas capas de información producidas por la edición digital.

References

Aguilar, S. T., Tannier, X., & Chastang, P. (2016). Named entity recognition applied on a data base of Medieval Latin charters. The case of chartae burgundiae. *3rd International Workshop on Computational History (HistoInformatics 2016)*. https://hal.science/hal-02407159/

Bamman, D., & Burns, P. J. (2020). *Latin BERT: A Contextual Language Model for Classical Philology* (arXiv:2009.10053). arXiv. https://doi.org/10.48550/arXiv.2009.10053

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked Data—The Story So Far. *International Journal on Semantic Web and Information Systems (IJSWIS)*, *5*(3), 1–22. https://doi.org/10.4018/jswis.2009081901

Burns, P. J. (2023). *LatinCy: Synthetic Trained Pipelines for Latin NLP* (arXiv:2305.04365). arXiv. https://doi.org/10.48550/arXiv.2305.04365

Cao, N. D., Izacard, G., Riedel, S., & Petroni, F. (2021). *Autoregressive Entity Retrieval* (arXiv:2010.00904). arXiv. https://doi.org/10.48550/arXiv.2010.00904

Ciotti, F., Lana, M., & Tomasi, F. (2014). TEI, Ontologies, Linked Open Data: Geolat and Beyond. *Journal of the Text Encoding Initiative*, (Issue 8), Article Issue 8. https://doi.org/10.4000/jtei.1365

Craig, C., Goyal, K., Crane, G., Shamsian, F., & Smith, D. A. (2023). Testing the limits of neural sentence alignment models on classical greek and latin texts and translations. *Proceedings CEUR-WS*, *1613*, 0073. https://ceur-ws.org/Vol-3558/paper6193.pdf

Erdmann, A., Brown, C., Joseph, B., Janse, M., Ajaka, P., Elsner, M., & de Marneffe, M.-C. (2016). Challenges and Solutions for Latin Named Entity Recognition. En E. Hinrichs, M. Hinrichs, & T. Trippel (Eds.), *Proceedings of the Workshop on Language Technology Resources and Tools for Digital Humanities (LT4DH)* (pp. 85–93). The COLING 2016 Organizing Committee. https://aclanthology.org/W16-4012/

Feng, F., Yang, Y., Cer, D., Arivazhagan, N., & Wang, W. (2022). *Language-agnostic BERT Sentence Embedding* (arXiv:2007.01852). arXiv. https://doi.org/10.48550/arXiv.2007.01852

Fordyce, C. J. (1961). *Catullus: A Commentary*. Oxford University Press.

Galán, L. (2013). *Catulo, Cayo Valerio. Poesía Completa*. Colihue.

Hogan, A., Blomqvist, E., Cochez, M., D'amato, C., Melo, G. D., Gutierrez, C., Kirrane, S., Gayo, J. E. L., Navigli, R., Neumaier, S., Ngomo, A.-C. N., Polleres, A., Rashid, S. M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., & Zimmermann, A. (2021). Knowledge Graphs. *ACM Comput. Surv.*, *54*(4), 71:1-71:37. https://doi.org/10.1145/3447772

Jockers, M. L. (2013). *Macroanalysis: Digital Methods and Literary History*. University of Illinois Press.

Kolitsas, N., Ganea, O.-E., & Hofmann, T. (2018). End-to-End Neural Entity Linking. En A. Korhonen & I. Titov (Eds.), *Proceedings of the 22nd Conference on Computational Natural Language Learning* (pp. 519–529). Association for Computational Linguistics. https://doi.org/10.18653/v1/K18-1050

Lewis, C. S. (1936). *La alegoría del amor: Un estudio sobre tradición medieval* (2015a ed.). Encuentro.

Mambrini, F., & Passarotti, M. C. (2023). The LiLa Lemma Bank: A Knowledge Base of Latin Canonical Forms. *Journal of Open Humanities Data*, *9*, 28. https://doi.org/10.5334/johd.145

Merrill, E. T. (with Robarts - University of Toronto). (1893). *Catullus; edited by Elmer Truesdell Merrill*. Boston Ginn. http://archive.org/details/catulluseditedby00catuuoft

Moretti, F. (2013). *Distant Reading*. Verso.

Moro, A., Raganato, A., & Navigli, R. (2014). Entity Linking meets Word Sense Disambiguation: A Unified Approach. *Transactions of the Association for Computational Linguistics*, *2*, 231–244. https://doi.org/10.1162/tacl_a_00179

Nadeau, D., & Sekine, S. (2007). A survey of named entity recognition and classification. *Lingvisticæ Investigationes*, *30*(1), 3–26. https://doi.org/10.1075/li.30.1.03nad

Nusch, C. J., Calarco, G. A., Del Rio Riande, G., Cagnina, L. C., Antonelli, R. L., & Errecalde, M. L. (2026). De Catulo a Wikidata: Automatización de tareas de codificación utilizando modelos de lenguaje, esquemas de metadatos y ontologías para un borrador de edición digital con el estándar XML-TEI. *Journal of the Text Encoding Initiative*, *19*.

Nusch, C. J., del Rio Riande, G., Cagnina, L. C. C., Errecalde, M. L., & Antonelli, L. (2024a). Initial Explorations for Document Clustering Tasks in Latin Elegiac Poets. *Decisioning*. Decisioning 2024.

Nusch, C. J., del Rio Riande, M. G., Cagnina, L., Errecalde, M. L., & Antonelli, R. L. (2024b). Clustering Tasks and Decision Trees with Augustan Love Poets: Cohesion and Separation in Feature Importance Extraction. *CEUR Workshop Proceedings*, *3834*. http://sedici.unlp.edu.ar/handle/10915/175050

Passarotti, M., Mambrini, F., Franzini, G., Cecchini, F. M., Litta, E., Moretti, G., Ruffolo, P., & Sprugnoli, R. (2020). Interlinking through Lemmas. The Lexical Collection of the LiLa Knowledge Base of Linguistic Resources for Latin. *Studi e Saggi Linguistici*, *58*(1), 177–212. https://doi.org/10.4454/ssl.v58i1.277

Pierazzo, E., & Driscoll, M. J. (2016). *Digital Scholarly Editing: Theories and Practices*. Open Book Publishers. http://archive.org/details/43d96298-a683-4098-9492-bba1466cb8e0

Ramsay, S. (2011). *Reading Machines: Toward and Algorithmic Criticism*. University of Illinois Press.

Riande, G. D. R. (2022). Humanidades Digitales o las Humanidades en la intersección de lo digital, lo público, lo mínimo y lo abierto. *Publicaciones de la Asociación Argentina de Humanidades Digitales*, *3*, e038–e038. https://doi.org/10.24215/27187470e038

Riesgo, G. Á. (2024). Dionisio Areopagita y el legado de Jámblico. *Patristica et Mediævalia*, *45*(2), 117–136. https://doi.org/10.34096/petm.v45.n2.16116

Sawada, Y., Fujita, T., Sakai, Y., & Watanabe, T. (2026). entity-linkings: A Unified Library for Entity Linking. En D. Croce, J. Leidner, & N. S. Moosavi (Eds.), *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (Volume 3: System Demonstrations)* (pp. 591–601). Association for Computational Linguistics. https://doi.org/10.18653/v1/2026.eacl-demo.42

Soler Ruiz, A. (2000). *Catulo. Poemas. Tibulo. Elegías*. Gredos.

Sommerschield, T., Assael, Y., Pavlopoulos, J., Stefanak, V., Senior, A., Dyer, C., Bodel, J., Prag, J., Androutsopoulos, I., & de Freitas, N. (2023). Machine Learning for Ancient Languages: A Survey. *Computational Linguistics*, 1–44. https://doi.org/10.1162/coli_a_00481

Vrandečić, D., & Krötzsch, M. (2014). Wikidata: A free collaborative knowledgebase. *Commun. ACM*, *57*(10), 78–85. https://doi.org/10.1145/2629489

Wei Shen, Jianyong Wang, & Jiawei Han. (2015). Entity Linking with a Knowledge Base: Issues, Techniques, and Solutions. *IEEE Transactions on Knowledge & Data Engineering*. https://ieeexplore.ieee.org/document/6823700

Xiao, Z., Gong, M., Wu, J., Zhang, X., Shou, L., Pei, J., & Jiang, D. (2023). *Instructed Language Models with Retrievers Are Powerful Entity Linkers* (arXiv:2311.03250). arXiv. https://doi.org/10.48550/arXiv.2311.03250