



PHILOLOGIA HISPALENSIS

ESTUDIOS LITERARIOS

2025 | VOL. XXXIX 2

PHILOLOGIA HISPALENSIS

AÑO 2025
VOL. XXXIX/2

ESTUDIOS LITERARIOS



FACULTAD DE FILOLOGÍA
UNIVERSIDAD DE SEVILLA

EVALUACIÓN DE ORIGINALES: Los originales se someten a una evaluación ciega, un proceso anónimo de revisión por pares, siendo enviados a evaluadores externos y también examinados por los miembros del Consejo de Redacción y/o los especialistas del Consejo Asesor de la Revista.

PERIODICIDAD: Anual en formato tradicional y en formato electrónico.

PUBLICACIÓN EN INTERNET: <<https://editorial.us.es/es/revistas/philologia-hispalensis>>, <<https://revistascientificas.us.es/index.php/PH>>.

BASES DE DATOS: *Philologia Hispalensis* se encuentra indexada en CARHUS Plus+2018, CIRC (grupo A), DIALNET, DOAJ, Dulcinea, Index Islamicus, Latindex 2.0 (100% de los criterios cumplidos), MIAR (ICDS 2022 = 10), MLA, REDIB, SCOPUS, ERIHPLUS y ANVUR (Clase A). Asimismo, cuenta con el sello de calidad de la FECYT (8ª edición, 2023, renovado en 2024 y válido hasta 2025) en los campos de conocimiento Lingüística y Literatura dentro de la modalidad Humanidades.

ENVÍO DE ORIGINALES Y SUSCRIPCIONES: Las colaboraciones deben enviarse a través de <<https://revistascientificas.us.es/index.php/PH>>.

DIRECCIÓN DE CONTACTO: Secretariado de la Revista *Philologia Hispalensis*, Facultad de Filología, Universidad de Sevilla, C/ Palos de la Frontera, s/n, 41004 Sevilla; o bien al correo electrónico <philhisp@us.es>.

INTERCAMBIOS O CANJES (BIBLIOTECAS UNIVERSIARIAS): Solicitense a Editorial Universidad de Sevilla o al Secretariado de la revista <philhisp@us.es>.

© Editorial Universidad de Sevilla

Financiación: Revista financiada por la Universidad de Sevilla dentro de las ayudas del VII PPIT-US y del Decanato de la Facultad de Filología.

PORTADA: referencias.maquetacion@gmail.com

DEPÓSITO LEGAL: SE-354-1986

ISSN: 1132 - 0265 / eISSN 2253-8321

Maquetación: referencias.maquetacion@gmail.com

IMPRIME: Podiprint

DISTRIBUYE: Editorial Universidad de Sevilla, Porvenir, 27, 41013 Sevilla

Licence Creative Commons Atribución/Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0)



EQUIPO EDITORIAL

Directora: Yolanda Congosto Martín, Universidad de Sevilla, España
Secretaria: Leyre Martín Aizpuru, Universidad de Sevilla, España
Editora: Salomé Lora Bravo, Universidad de Sevilla, España
Coordinadora de Reseñas: M. Amparo Soler Bonafont, Universidad Complutense de Madrid, España
Difusión en redes sociales: Blanca Jiménez Coca (coord.), Universidad de Sevilla, España
Sara Blanco López (ayudante), Universidad Complutense de Madrid, España

Consejo de

Redacción: Gema Areta Marigó, Universidad de Sevilla, España
Elisabetta Carpitelli, Université Stendhal - Grenoble Alpes, France
María Auxiliadora Castillo Carballo, Universidad de Sevilla, España
Antonio Luis Chaves Reino, Universidad de Sevilla, España
Marianna Chodorowska-Pilch, University of Southern California, USA
Yves Citton, Université Paris 8 Vincennes-Saint Denis, France
Ninfa Criado Martínez, Universidad de Sevilla, España
Isabel María Íñigo Mora, Universidad de Sevilla, España
Manuel Maldonado Alemán, Universidad de Sevilla, España
Daniela Marcheschi, Università degli Studi di Perugia, Italia
Pedro Martín Butragueño, Colegio de México, México
Miguel Ángel Quesada Pacheco, Universitetet i Bergen, Norge
Angelica Valentinetti, Universidad de Sevilla, España
Alf Monjour, Universität Duisburg-Essen, Deutschland
María José Osuna Cabezas, Universidad de Sevilla, España
Fátima Roldán Castro, Universidad de Sevilla, España
Antonio Romano, Università degli Studi di Torino, Italia
Juan Pedro Sánchez Méndez, Université de Neuchâtel, Suisse
María Luisa Siguán Boehmer, Universitat de Barcelona, España
José Solís de los Santos, Universidad de Sevilla, España
Modesta Suárez, Université de Toulouse-Le Mirail, France
María Ángeles Toda Iglesia, Universidad de Sevilla, España
José Agustín Vidal Domínguez, Universidad de Sevilla, España
María Jesús Viguera Molins, Universidad Complutense de Madrid, España
Adamantía Zerva, Universidad de Sevilla, España

COMITÉ CIENTÍFICO

Juan Francisco Alcina Rovira, Universitat Rovira i Virgili, España
Gerd Antos, Martin-Luther-Universität Halle-Wittenberg, Deutschland
Gianluigi Beccaria, Università degli Studi di Torino, Italia
Isabel Carrera Suárez, Universidad de Oviedo, España
Carmen Herrero, Manchester Metropolitan University, England
Anna Housková, Univerzita Karlova, Česká Republika
Dieter Kremer, Universität Trier, Deutschland
Xavier Luffin, Vrije Universiteit Brussel, Belgique
Roberto Nicolai, Sapienza - Università di Roma, Italia
Marie-Linda Ortega, Université Sorbonne Nouvelle - Paris 3, France
Deborah C. Payne, American University, USA
Carmen Silva-Corvalán, University of Southern California, USA
Alicia Yllera Fernández, UNED, España

CONSEJO ASESOR

ESTUDIOS ÁRABES E ISLÁMICOS

Eva Lapiedra Gutiérrez, Universidad de Alicante, España

Pablo Beneito Arias, Universidad de Murcia, España

Carmelo Pérez Beltrán, Universidad de Granada, España

FILOLOGÍA ALEMANA

Georg Pichler, Universidad de Alcalá, España

Marta Fernández-Villanueva Jané, Universitat de Barcelona, España

María José Domínguez, Universidade de Santiago de Compostela, España

FILOLOGÍA CLÁSICA - LATÍN

Jesús Luque Moreno, Universidad de Granada, España

José Luis Moralejo Álvarez, Universidad de Alcalá de Henares, España

Eustaquio Sánchez Salor, Universidad de Extremadura, España

FILOLOGÍA CLÁSICA - GRIEGO

Didier Marcotte, Université Sorbonne Paris, France

Maurizio Sonnino, Sapienza-Università di Roma, Italia

Stefan Schorn, Université Catholique de Louvain, Belgique

FILOLOGÍA FRANCESA

Dolores Bermúdez Medina, Universidad de Cádiz, España

Monserrat Serrano Mañes, Universidad de Granada, España

María Luisa Donaire Fernández, Universidad de Oviedo, España

FILOLOGÍA ITALIANA

Giovanni Albertocchi, Universitat de Girona, España

Cesáreo Calvo Rigual, Universitat de València - IULMA, España

Margarita Borreguero Zuloaga, Universidad Complutense de Madrid, España

LENGUA ESPAÑOLA

Emilio Montero Cartelle, Universidade de Santiago de Compostela, España

Antonio Salvador Plans, Universidad de Extremadura, España

Antonio Briz Gómez, Universitat de València, España

LENGUA INGLESA

Emilia Alonso Sameño, Ohio University, USA

Carmen Gregori Signes, Universitat de València, España

Nuria Yanez-Bouza, Universidade de Vigo, España

LINGÜÍSTICA

Ángel López García, Universitat de València, España

Eugenio Martínez Celdrán, Universitat de Barcelona, España

Juan Carlos Moreno Cabrera, Universidad Autónoma de Madrid, España

LITERATURA ESPAÑOLA

Pedro M. Cátedra, Universidad de Salamanca, España

Flavia Gherardi, Università degli Studio di Napoli Federico II, Italia

Leonardo Romero Tobar, Universidad de Zaragoza, España

LITERATURA HISPANOAMERICANA

Teodosio Fernandez, Universidad Autónoma de Madrid, España

Noé Jitrik, Universidad de Buenos Aires, Argentina

Edwin Williamson, Oxford University, Inglaterra

LITERATURA INGLESA

Luis Alberto Lázaro Lafuente, Universidad Alcalá de Henares, España

Ricardo Mairal Usón, UNED, España

Carme Manuel Cuenca, Universitat de València, España

TEORÍA DE LA LITERATURA

José Domínguez Caparrós, UNED, España

Antonio Garrido Domínguez, Universidad Complutense de Madrid, España

Isabel Paraíso Almansa, Universidad de Valladolid, España

REVISORES DEL VOLUMEN 39, NÚMERO 2 (2025), ESTUDIOS LITERARIOS

Han actuado como revisores anónimos para uno o más artículos de este número, tanto los aceptados como los rechazados, los siguientes investigadores:

Alvite Díez, María Luisa (Universidad de León)
Calarco, Gabriel Alejandro (CONICET, Argentina)
Camprubi Vinyals, Adriana (Universitat de Barcelona)
de Mora Valcárcel, Carmen (Universidad de Sevilla)
Fernández Riva, Gustavo (Heidelberg Universität, Alemania)
Gatica Cote, Paulo Antonio (Universidad Complutense de Madrid)
González Martínez, Déborah (Universidade de Santiago de Compostela)
González Miranda, Emilio (Universidade de Santiago de Compostela)
González Montes, Antonio (Universidad Nacional Mayor de San Marcos / Academia Peruana de la Lengua, Perú)
González Sanz, Marina (Universidad de Sevilla)
Martí Junquera, Sadurní (Universitat de Girona)
Martín Junquera, Imelda (Universidad de León)
Moreno Lago, Eva María (Universidad de Sevilla)
Morrás Ruiz-Falcó, María (Universitat Pompeu Fabra)
Pérez Agustín, María Mercedes (Universidad Complutense de Madrid)
Pérez Ben, Lorena (Universidade de Santiago de Compostela)
Plata Parga, Fernando (Colgate University, EEUU)
Roeder, Torsten (German National Academy Leopoldina, Alemania)
Sánchez Dueñas, Blas (Universidad de Córdoba)
Simó Torres, Meritxell (Universitat de Barcelona)
Vaamonde Dos Santos, Gael (Universidad de Granada)
Vargas Díaz-Toledo, Aurelio (Universidad Complutense de Madrid)

ÍNDICE

| | |
|---|--------|
| Sección Monográfica. Ediciones filológicas digitales: hitos y nuevas perspectivas / <i>Digital Scholarly Editions: Achievements and New Perspectives</i> | 13 |
| Introducción. Ediciones filológicas digitales: hitos y nuevas perspectivas/ <i>Introduction. Digital Scholarly Editions: Achievements and New Perspectives</i> | 15-19 |
| VICTOR MILLET (Universidade de Santiago de Compostela) | |
| Taming the Beast: Transcribing Hernando Colón's <i>Libro de los epítomes</i> / <i>Domando a la bestia: transcripción del Libro de los epítomes de Hernando Colón</i> | 21-39 |
| Matthew Driscoll (University of Copenhagen) | |
| Alessandro Gnasso (University of Copenhagen) | |
| https://dx.doi.org/10.12795/PH.2025.v39.i02.01 | |
| Editing Premodern German Song Texts within the Musicologocial Edition Platform <i>E-Laute</i> . Challenges and Possibilities/ <i>Editar textos de canciones alemanas de los siglos XV y XVI en el contexto de la plataforma de edición digital musicológica e-laute: desafíos y oportunidades</i> | 41-63 |
| Stefan Rosmer (Universität Bayreuth) | |
| David M. Weigl (Universität für Musik und darstellende Kunst Wien) | |
| https://dx.doi.org/10.12795/PH.2025.v39.i02.02 | |
| Edición digital académica de textos áureos no canónicos: la propuesta del proyecto BIDISO/ <i>The Scholarly Digital Editions of Non-Canonical Spanish Golden Age Texts: the Proposal of the BIDISO Project</i> | 65-96 |
| Nieves Pena Sueiro (Universidade da Coruña) | |
| Carlota Fernández Travieso (Universidade da Coruña) | |
| https://dx.doi.org/10.12795/PH.2025.v39.i02.03 | |
| Designing a User Interface for the DSE 2.0: New Opportunities, New Challenges/ <i>Diseño de una interfaz de usuario para la DSE 2.0: Nuevas oportunidades, nuevos retos</i> | 97-115 |
| Roberto Rosselli Del Turco (Università di Torino) | |
| https://dx.doi.org/10.12795/PH.2025.v39.i02.04 | |

Designing Digital Editions for Humans: Insights on User Experience and Usability from the *Digital Dossier: Alexander von Humboldt and Cuba (1800–1830)* / *Diseño de ediciones digitales para humanos: perspectivas sobre la experiencia del usuario y la usabilidad a partir del Dossier digital: Alexander von Humboldt y Cuba (1800–1830)*..... 117-130
 Antonio Rojas Castro (Eberhard Karls Universität Tübingen)
<https://dx.doi.org/10.12795/PH.2025.v39.i02.05>

Un modelo HTR para incunables castellanos / *An HTR Model for Spanish Incunabula*..... 131-177
 José Manuel Fradejas Rueda (Universidad de Valladolid)
 Mario Cossío Olavide (Université de Bretagne Occidentale)
<https://dx.doi.org/10.12795/PH.2025.v39.i02.06>

The Future is already here: Navigating the New Frontiers of Digital Scholarly Editing in an Age of HTR and AI / *El futuro ya está aquí: navegar por las nuevas fronteras de la edición filológica digital en la era de la HTR y la IA*..... 179-199
 James Cummings (University of Newcastle)
<https://dx.doi.org/10.12795/PH.2025.v39.i02.07>

Sección Varia

Trauma, mito y tragedia en la narrativa de Leslie Marmon Silko: de «Old Juana» a los zapatistas / *Trauma, Myth, and Tragedy in Leslie Marmon Silko's Narrative: from "Old Juana" to the Zapatistas*..... 203-218
 José Manuel Correoso Rodenas (Universidad Complutense de Madrid)
<https://dx.doi.org/10.12795/PH.2025.v39.i02.08>

El episodio del alguacil en el tratado séptimo del *Lazarillo de Tormes*: sátira contra el arzobispo de Toledo Juan Martínez Silíceo y los «retraídos» de la catedral primada / *The Episode of the Bailiff in in the Seventh Treatise of Lazarillo de Tormes: Satire against the Archbishop of Toledo Juan Martínez Silíceo and the "withdrawn" of the Primate Cathedral*..... 219-237
 Jesús Fernando Cáseda Teresa (IES Valle del Cidacos - Calahorra, La Rioja)
<https://dx.doi.org/10.12795/PH.2025.v39.i02.09>

El alma del mundo y el tejido cósmico: la poética del territorio del Inca Garcilaso de la Vega / *The Soul of the World and the Cosmic Fabric: the Inca Garcilaso de la Vega's Poetics of Territory*..... 239-257
 Pedro Martín Favaron Peyón (Pontificia Universidad Católica del Perú)
<https://dx.doi.org/10.12795/PH.2025.v39.i02.10>

Reseñas de libros

- Javier Salazar Rincón: *De alcaldes y alcaldadas. Trayectoria y significado de un personaje risible en la literatura del Siglo de Oro*. La Seu d'Urgell: Javier Salazar Rincón, 2024, 735 pp. ISBN: 978-84-09-58793-3 261-264
Luis Gómez Canseco (Universidad de Huelva)
- Prue Shaw (Ed.): *Dante Alighieri Commedia. A Digital Edition*. Segunda edición. Florencia: Fondazione Ezio Franceschini, Inkless Editions, Saskatoon, 2021. ISBN: 1-904628-21-4 265-267
Adriana Camprubí Vinyals (Universidad de Barcelona)
- Antonio Reyes Ruiz: *Génesis del Ensanche de Tetuán (Marruecos). El papel de la burguesía comercial catalana*. Sevilla: Editorial Universidad de Sevilla (Colección *Mediterráneo Textos y Estudios*, vol. 1), 2024, 158 pp. ISBN: 978-84-472-2758-7 269-272
Carmelo Pérez Beltrán (Universidad de Granada)
- Verónica Casais Vila (Ed.): Pedro Calderón de la Barca, *Las manos blancas no ofenden* (https://www.calderondelabarca.org/editions/las_manos_blancas.pdf), pp. 1-102; Verónica Casais Vila (Ed.): Pedro Calderón de la Barca, *A secreto agravio, secreta venganza* (https://www.calderondelabarca.org/editions/secreto_agravio.pdf), pp. 1-62; Laura Carbajo Lago (Ed.): Pedro Calderón de la Barca, *Duelos de amor y lealtad* (<https://www.calderondelabarca.org/editions/duelos.pdf>), pp. 1-94. Acceso en línea: https://www.calderondelabarca.org/ediciones_criticas 273-278
Candela Iglesias Balsa (Universidade de Santiago de Compostela)
- Evina Stein y Gustavo Fernández Riva (Eds.): Networks of Manuscripts, Networks of Texts, *Journal of Historical Network Research*, 9(1), 2023, 238 pp. ISSN-e: 2535-8863 279-281
Lorena Pérez Ben (Universidade de Santiago de Compostela)



ESTUDIOS LITERARIOS

THE FUTURE IS ALREADY HERE: NAVIGATING THE NEW FRONTIERS OF
DIGITAL SCHOLARLY EDITING IN AN AGE OF HTR AND AI

EL FUTURO YA ESTÁ AQUÍ: NAVEGAR POR LAS NUEVAS FRONTERAS DE LA EDICIÓN
FILOLÓGICA DIGITAL EN LA ERA DE LA HTR Y LA IA

JAMES CUMMINGS

University of Newcastle

james.cummings@newcastle.ac.uk

 0000-0002-6686-3728

Recibido: 23-12-2023 | Aceptado: 23-04-2024

Cómo citar: Cummings, J. (2025). The Future Is already here: Navigating the New Frontiers of Digital Scholarly Editing in an Age of HTR and AI. *Philologia Hispalensis*, 39(2), 179-199.
<https://dx.doi.org/10.12795/PH.2025.v39.i02.07>

ABSTRACT

This article explores the evolving landscape of digital scholarly editing, focusing especially on the impact for scholarly editing of recent advancements in Artificial Intelligence (AI) and Handwritten Text Recognition (HTR) as a transformative technology that significantly accelerates the transcription of historical documents and promises to expand access to archives that were previously accessible only in person by those with palaeographic training. However, the article argues for caution against uncritical adoption of AI in scholarly editing workflows, emphasising the crucial role of human editors in ensuring accuracy, addressing inherent biases, and maintaining a human-centred approach to digital scholarship. The article investigates both the potential and limitations of AI in scholarly tasks such as Named Entity Recognition (NER), text parsing, collation of multiple witnesses, analysis of allusion and re-use and document summarisation, arguing for a collaborative model where human expertise complements, rather than is replaced by, AI tools. The article advocates for a thoughtful integration of AI into scholarly workflows, prioritising accuracy, transparency, and the preservation of essential editorial judgement.

Keywords: Digital Scholarly Editing, HTR, AI.

RESUMEN

Este artículo explora el panorama en evolución de la edición filológica digital, centrándose especialmente en el impacto que tienen para la edición académica los recientes avances

en Inteligencia Artificial (IA) y Reconocimiento de Texto Manuscrito (HTR) como una tecnología transformadora que acelera significativamente la transcripción de documentos históricos y promete ampliar el acceso a archivos que antes sólo eran accesibles en persona por aquellos con formación paleográfica. No obstante, el artículo aboga por la cautela frente a la adopción acrítica de la IA en los flujos de trabajo de edición académica, haciendo hincapié en el papel crucial de los editores humanos a la hora de garantizar la precisión, abordar los sesgos inherentes y mantener un enfoque centrado en el ser humano en la filología digital. El artículo investiga tanto el potencial como las limitaciones de la IA en tareas académicas como el reconocimiento de entidades nombradas (NER), el análisis sintáctico de textos, el cotejo de múltiples testigos, el análisis de alusiones y reutilización y el resumen de documentos, abogando por un modelo colaborativo en el que la experiencia humana complementa a las herramientas de IA, en lugar de sustituirlas. El artículo aboga por una integración meditada de la IA en los flujos de trabajo académicos, dando prioridad a la precisión, la transparencia y la preservación del juicio editorial esencial.

Palabras clave: Edición filológica digital, HTR, AI.

1. INTRODUCTION

The development of digital scholarly editions has been shaped by a long and dynamic history, inextricably linked to the progress of computing technologies, the evolution of the web, and the emergence of standards for the critical representation of historical documents.¹ The history of not only digital humanities, but human technological development, can be traced through the approaches to, issues with, and anxieties concerning digital scholarly editions over time. While early digital editions often relied on simple so-called “plain e-texts”,² limited by the technologies available at the time, the field progressed with the development of more sophisticated markup languages and encoding standards. The guidelines of the Text Encoding Initiative (TEI) long ago emerged as the de facto standard for encoding digital scholarly editions, providing a robust framework for representing textual features, editorial interventions, and a community to support this.³ This widespread adoption of the TEI might suggest that the challenges of representing the interpretations of editors might have been resolved. However, the reality is, as one might expect, more complex. At

¹ This article has been influenced by my presentation on “*Encoding Critical Apparatus in an Age of AI*” at a workshop on *New Perspectives on Critical Editions* sponsored by the Groupe de Recherche sur les Éditions critiques en contexte Numérique (GREN) and the Centre de recherche interuniversitaire sur les humanités numériques at the Université de Montréal.

² The imposition of a caveat “so-called” to “plain e-texts” is because all too often the critical use of forms of markup is elided, even in the cases of truly “plain text”, these still often use forms of markup even if only language-specific punctuation and spacing. Increasingly, what is viewed as “plain text” by some includes incredibly complex XML Markup structures such as Word DocX files.

³ For the TEI guidelines themselves see: *TEI: Guidelines for Electronic Text Encoding and Interchange*, <https://tei-c.org/release/doc/tei-p5-doc/en/html/index.html>

its heart, the TEI is a highly customizable framework, allowing projects to tailor the application of the TEI guidelines to their specific needs. They do this through the customisation of the TEI scheme using a TEI ODD customisation format, itself authored in TEI, that is a form of meta-schema from which both documentation and specific schemas can be generated (Cummings, 2014). This inherent flexibility, while a strength for accommodating diverse projects, also introduces potential complexities and inconsistencies. Simultaneously, more recent developments in Artificial Intelligence (AI) are creating new opportunities and challenges for those editors (and research teams) who create digital scholarly editions. This article engages with the topics of HTR and scholarly editing, and whether technologies like this democratise digital scholarly editing. It looks at AI and scholarly editing methodologies (including: named entity recognition; text parsing, collation, and analysis; document summarisation) before considering the state of AI and scholarly editing knowledge and presenting some tentative conclusions. (Tentative because any consideration of AI and digital scholarly editing is currently an immense field of shifting sands). Overall, this article reflects upon the ongoing effects of recent developments, such as sophisticated Handwritten Text Recognition (HTR) and AI, from the perspective of those creating and publishing digital scholarly editions. In scouting the current terrain of new frontiers in digital scholarly editing, this article advocates for a future which retains a human-centered approach to using these technologies.

2. HANDWRITTEN TEXT RECOGNITION AND SCHOLARLY EDITING

Continued development of Machine Learning (ML) methodologies are transforming numerous fields, and the creation of digital scholarly editions and their associated research, is certainly one of these. While multiple ML techniques, such as the development of GPT-based Large Language Models (LLMs) and the AI Chat-Bots based on them, are a significant forward step, we must not overlook Handwritten Text Recognition (HTR), which is a very significant development for scholarly editing. HTR uses sophisticated, model-based machine learning algorithms to pattern-match handwritten text and transform such documents into machine-processable digital text (Terras *et al.*, 2024). Just as Optical Character Recognition (OCR) revolutionised the digitisation of printed materials by reliably converting clearly printed text into digital form, HTR is now doing the same not only for handwritten texts but also that print which is harder to decipher with an OCR-based approach. However, due to the complexities inherent in deciphering diverse and often internally-inconsistent handwriting styles, the process of using HTR is considerably more involved than the application of OCR to the image of a text and thus is most advantageous for projects with substantial handwritten resources to transcribe. That said, the HTR processes are becoming easier as more detailed models grow and are shared publicly.

HTR relies heavily on the construction of predictive language models that can interpret individual words and lines of handwritten text sequentially. The effectiveness of these models hinges on the availability of comprehensive and accurately transcribed training datasets (Retsinas *et al.*, 2022: 248). There are many potential sources for such data, some of which have aligned transcribed text with source images, in large text corpora projects and text archives.⁴ As the models are provided with more examples of handwriting, they can decipher increasingly complex scripts and adapt to variations in style (*cfr.* Stokes & Kiessling, 2025). Consequently, projects with larger manuscript corpora, which offer a greater range of handwriting samples for training within the same script, are likely to derive the most benefit from their application of HTR technology. Recent advancements in AI and ML, particularly in the realm of deep neural networks, have propelled HTR from a computationally challenging problem to a largely solved one as Nockels *et al.* (2024) discuss:

The aspiration of recognising handwriting from people of various backgrounds, nationalities, professions and education, with equal competence, accuracy and speed, has largely been realised. (Nockels *et al.*, 2024)

While a couple of decades ago the dream of HTR was very much still a promising fantasy (but significant research area), it has rapidly developed alongside ML techniques. It has developed from only being able to recognise extremely clear handwriting with a very large training corpus, through successive stages of improvement, and can now recognise quite difficult handwritten texts with significantly smaller amounts of training data. These improvements are evident in the increasing accuracy, speed, and diversity of handwritten texts with which HTR systems can now process, often with very small Character Error Rates (CER). The wisdom of past years that suggested HTR would only ever be useful for keyword searching through handwritten corpora to then present a surrogate image of the document, has been truly refuted by such developments (Nockels *et al.*, 2024). HTR is now of significant benefit to many editorial project workflows.

The increasing accessibility and accuracy of HTR will have profound implications for most fields of historical research over the next few decades. As technologies such as this continue to mature and develop, we should expect an avalanche (in this case slow at the start, but rapidly gaining momentum and material) in the

⁴ A reviewer of this article notes that in the Hispanic Digital Humanities field there are corpus projects such as *CHARTA*, *ODE*, *Post Scriptum*, *CORDICan*, *Panépica Digital*, *Biblia Medieval*, or *CAREXIL* whose data might be useful in providing training material for HTR processes from a variety of historical periods and contexts. Indeed, there are many such corpora and international archives of digital texts that could be useful in model building and expansion. The provision of greater and more diverse training materials, especially where these are transcriptions that are linked to the word-zones of their surrogate images, can only improve the HTR models available.

availability of previously inaccessible archival texts, opening up new avenues for various forms of historical analysis across datasets, archives, and corpora. In this new world, more projects that are editing historical documents will include “automatic text acquisition in their data processing chain” (Pinche and Stokes, 2024: 1). Moreover, as this creates significantly larger collections of machine-processable texts originating from more diverse linguistic, historical, and cultural contexts than previously, this novel data can be leveraged to train, develop, and enhance existing models, thus further accelerating research in digital textual studies. Ultimately, HTR has the potential to revolutionise not only how we access, interrogate, and interact with historical documents themselves, but also more broadly how we understand and interpret the past itself.

This fundamental achievement presents ramifications for the scholarly editing community that textual scholars must consider in order to understand the epistemic affordances of this developing technology at the juncture of machine learning and online technologies. (Terras *et al.*, 2024)

This handwritten text paradigm shift will be first noticeable in forms of archival research, benefitting from sources that were previously obscured, searching across entire archives for particular records of historical interest, and the greater inclusion in academic discourse of more accurate transcriptions of handwritten written materials in a more diverse range of languages and historical sources. However, as one might expect, the integration of new technologies into the process of creating and publishing digital scholarly editions faces a number of significant, non-technical, obstacles ranging from the availability of appropriate funding, through the need for training in these areas, to the understandable reticence of some of the human editors themselves. While advancements like HTR and AI show immense potential, their effective implementation is sometimes constrained by human factors such as a lack of training or the development of tools that editors find truly fit for their tasks. While it may not be necessary for editors of historical texts to understand the inner workings of such assistive technologies fully, there is the need for both some degree of skills training and also for those creating tools in this area to involve the highly-skilled editors they seek to assist in the development of these tools (Cummings, 2019: 190-191).

If new technologies are going to be successful in the creation of digital scholarly editions, then these need to be fully adopted by editors and integrated into their research workflows. “The technology necessary to produce digital text editions must be present in the text management practice of the specialists who work on the research from the start” (Kecskeméti, 2023: 5). However, this needs to be done in a balance that consciously avoids a technological dependence which excludes or even outsources the human from the scholarly editing process.

The need for “humans in the loop” is not always recognised by those working creating or studying digital scholarly editions.⁵ Indeed, it seems problematic when students seem to uncritically claim that automated “software has the potential to render a bespoke edition model instantly” or indeed that such an edition “could be created at the click of a button, removing the restraints of time, money, and training” (Trowsdale, 2023: 36). While carefully constructed software can lessen or shift the responsibility of such restraints –but never truly remove them– this is not where the bulk of editorial work is enacted. Such systems rely on a previously-edited digital edition file as the source for an automatically-generated website to display that edition. As usual, the human-centric editing really resides in the intellectual content of the master source file –the underlying data is what is crucial, rather than the generation of derivative views of it. However, the provision of out-of-the-box edition software can also undermine the “interface-as-argument” approach where choices in how to display an edition inherently reflect the arguments that the edition is making (*cfr.* Andrews and van Zundert, 2018: 3). What needs to be foregrounded in training the next generation is that we should be hesitant to call something an “edition”, and certainly not a “digital scholarly edition”, if the human knowledge and experience of an editor has not been applied directly to the critical representation of those texts and only done so in providing the setting of options for processing the texts automatically. The human editor is a necessary component of digital scholarly editing, otherwise we are shirking the duties of an editor.⁶

While it is undeniable that the integration of HTR into scholarly editing workflows can present a number of challenges, the potential benefits are equally undeniable, especially for projects with a large number of handwritten historical documents. The main advantage lies in the significant acceleration of the transcription process. Once a robust model has been developed, based either on an individual scribe’s handwriting or more generally using a model of an applicable type of hand, HTR can efficiently process large volumes of material, often achieving remarkably low CER. Where projects involve transcribing a significant number of textual sources, the benefits of successful HTR implementation are likely to outweigh any associated difficulties. While proofreading still remains necessary, through delegating the repetitive and error-prone task of transcription to a semi-automated

⁵ In this case the term “humans in the loop” is used in the manner considered by the field of Human Computer Interaction (HCI), where humans are a necessary component of the training and validation of results or decisions, but any consideration of this in the domain of AI is indebted to the early science fiction explorations of these issues by authors such as Asimov, Dick, and Clarke.

⁶ Part of the duty of an editor is to present a reliable text needing “accuracy with respect to texts, adequacy and appropriateness with respect to documenting editorial principles and practice, consistency and explicitness with respect to methods”. MLA Guidelines for Editors of Scholarly Editions, <https://www.mla.org/Resources/Guidelines-and-Data/Reports-and-Professional-Guidelines/Guidelines-for-Editors-of-Scholarly-Editions>

process, HTR can mitigate the risk of human error and substantially reduce the overall time required for transcription. This results in the ability of a single project to transcribe a much larger volume of material within a shorter timeframe (Terras *et al.*, 2024). Furthermore, HTR can enhance the accuracy and consistency of transcriptions, particularly in the case of complex or ambiguous handwriting. By providing a machine-generated initial transcription, HTR can serve as a valuable tool for human editors to then review and correct errors, thereby improving the overall quality of the final product. Additionally, HTR can facilitate the identification of patterns and trends within large datasets, enabling scholars to gain new insights into historical and literary texts (Nockels *et al.*, 2024).

The efficacy of HTR systems is usually measured by CER (Character Error Rate) but when HTR is used in conjunction with a LLM it might be better to use a metric of a “Word Error Rate” as this is demonstrably contingent upon the quality and representativeness of the training data. Instances where HTR encounters difficulties, and CER increases, arise when there are insufficient training examples for specific character forms, ligatures, or words. This challenge is particularly prevalent when significant variation and inconsistency in writing styles exist, when text is deformed to fit unconventional spaces, or special cases like abbreviated word-forms. Moreover, the physical condition of the source material directly impacts the reliability of HTR. The degradation or damage sustained by the physical object, and consequently any digital surrogates, can significantly affect the accuracy of HTR output (Parker *et al.*, 2019: 15). This issue is further compounded when the digital surrogate employed has image quality problems such as low resolution. The inherent ambiguity in the representation of characters in certain writing systems (such as Hebrew), presents an additional layer of complexity. Differentiating between characters in such systems often relies heavily on contextual understanding, making accurate classification solely based on visual features challenging. In such circumstances, the pairing of HTR with LLMs trained on a corpus of similar texts may be a more fruitful solution for research projects. The interplay of all these factors underscores the necessity of robust training datasets that encompass the diverse characteristics of handwritten materials to ensure optimal HTR performance.

3. DEMOCRATISATION OF DIGITAL SCHOLARLY EDITING

The potential for HTR to democratise digital scholarly editing is often suggested by those involved with its creation. Proponents of HTR systems will highlight its capacity to provide individual scholars, as well as institutions with limited resources, access to the infrastructure required for publication. While HTR undeniably lowers certain barriers to entry, the assertion that this technology might “democratize the practice of digital edition creation, making it easier to publish editions of underrepresented voices in the scholarly canon” (Terras *et al.*, 2024) should be approached

with some caution. While the potential certainly exists for HTR to enable access to previously underrepresented texts and marginalised voices, there is no guarantee that these will not merely be swamped by the torrent of archival documents already of a mainstream nature, and thus still remain obscured. Moreover, acquiring the technical skills necessary for substantial use of HTR technologies with difficult sources can also present a new barrier, and overcoming it might offset any perceived gains in general accessibility to the infrastructure. In many cases there are barriers in the access to this backend infrastructure –since the majority of any sophisticated form of HTR currently does not usually happen on a user’s computer itself, but the ML-based pattern recognition is farmed out to a remote data centre and hence is limited to a number of pages or other premium models to sustain the service (*cfr.* Stokes & Kiessling, 2025). In some cases not having access to modify server parameters or processes directly, having limited page processing, or only having a few viable competitors in HTR provision can reinforce systemic biases baked into the provision of the services. The fields of study for those underrepresented texts which represent these marginalised voices are correspondingly also those least likely to have easy and well-supported access to such new technologies. It is primarily those who are funding and creating the infrastructures that provide HTR who need to ensure that their systems at least attempt to counterbalance historical privileges, but large digitisation projects, and scholarly editors enhancing those transcriptions also need to play a role in any such democratisation:

Those who digitize content at scale, those who maintain HTR-based edition infrastructure, and those who undertake scholarly editing all need to scrutinize which material can be served by this technology, to ensure that traditionally marginalized voices can have the chance to benefit from it. (Terras *et al.*, 2024)

Moreover, the historical lack of representation for marginalised or non-canonical voices (and thus editions of their texts) within mainstream academic discourse stems from a complex combination of factors, extending far beyond the mere resource capacity to automatically transcribe texts. The absence of readily available and well-documented source material, coupled with the entrenched biases within academic publishing, and the standard approaches of research funding structures, present more significant obstacles than the limited access to HTR infrastructure alone. While the efficient generation of digital text through HTR is a contribution which can certainly be leveraged for digital scholarly editing, the more general institutional, governmental, and systemic issues of disparity in research opportunity throughout academia in general may remain more of a barrier (*cfr.* Terras *et al.*, 2024).

The use of open international standards goes some way to helping people overcome some of these barriers. However, this is not to suggest that such standards

are without their limitations or problematic histories –often current digital standards reflect an understanding of digital text which is predominantly from the anglo-american-european context. Incorporating a diversity of voices in their creation or updating can go some way to counteracting this. The TEI framework, rightly *de facto* standard for digital scholarly editions, is an open community with members from around the globe and is extensible at both a project and community level. While the TEI actively expands its recommendations to take on the respective interests of the community that creates it, the TEI community still faces some difficulties in broadening its inclusivity in a number of areas. One of these is the underlying pedagogical outreach which is a catalyst for greater diversity:

Also, one of the biggest drawbacks of teaching TEI in this context is the use of proprietary software in workshops. Although there are free XML editors, working with them is more challenging for novice Spanish students, requiring extra help for the markup activity and translation of tags. Proprietary licensing fees constrain the use of that software to short-term training opportunities. (Allés-Torrent & del Río Riande, 2019: para. 31)

The use of proprietary software in contexts which can afford it may indeed cause harm for the adoption of non-proprietary software in other locations. It may be possible to reduce these barriers through development of non-proprietary software and the internationalisation of these into a variety of languages.⁷

4. AI AND THE METHODOLOGIES OF SCHOLARLY EDITING

The use of HTR, as already discussed, is a very specific application of AI, based on supervised machine learning, which constructs models capable of interpreting distinct handwriting styles by learning from transcribed training data. This technology stands as a prominent example of AI's application that will affect humanities scholarship substantially over the next couple of decades. It directly addresses the challenges faced by scholars when engaging with vast archives of handwritten historical sources that are often opaque to technological interrogation. This is especially true in cases where the resulting data outputs are complex and fragile to technological shifts (Cummings, 2023). HTR can now significantly expedite the time-consuming task of transcribing larger corpus of documents, providing access to the data they contain, where the sheer number of them may have made mass-transcription efforts unlikely.

⁷ For example, TEI XML+RDF semantic text editors such as LEAF-Writer (see <https://leaf-writer.leaf-vre.org>).

One of the other side-effects of AI-based HTR technologies is not so much the creation of transcriptions themselves but the data which arises from the process of modelling scribal hands. Work in this area has been so focused on the creation and application of these models –and now the creation of standards of interchange for HTR models– that there has been less focus on the aggregate effects of a world where even many more of these models exist and archives have been transcribed. For example, there is potential for using these to assist in scribal identification internationally, by measuring the similarity weightings between an unknown scribe and those used in training of a model. Similarly, digital interrogation of individual graphemes, created by similar technological processes, are leading to improvements in digital palaeography and computational codicology. Approaches such as these both enable and benefit from studies across larger datasets internationally, and such methodologies will only improve as more HTR is undertaken with more data and models shared publicly. Overall, data generated from such processes has a significant value in itself, and unfortunately this inherent value can inhibit the desire for some projects (especially in the commercial world) to share it openly. Whether modelling data is easily available or not has a direct influence on the feasibility of new developments for such applications (Stokes & Kiessling, 2025). The biases of academic research, as mentioned earlier, mean that HTR and related technologies have been overly-focused on scripts written in the Latin alphabet, as there is already a wealth of training data in these scripts, but over time HTR proponents believe this will also lead to an increase in study of minority historical scripts and writing systems.

Another application of AI is in relation to source texts that cannot be viewed directly because they are too degraded or fragile. An example of this is the carbonised papyri from Herculaneum, which were illegible and unable to be physically unrolled, but which were digitally unrolled through use of micro-CT by a team led by Brent Seales (Parker *et al.*, 2019). While micro-CT scans work with iron gall ink, they don't really work with carbon-based ink, but the application of deep learning and 3D convolutional neural networks has enabled developments in this area of study as well. Increasingly, ongoing research and improvements in digital reading of otherwise inaccessible texts such as these are driving humanities imaging scholarship. While this may make some otherwise inaccessible texts readable, nevertheless it still raises substantial questions for scholarly editors –how do you make arguments about the nature of a text from an object you have never seen? If these “born-virtual”⁸ texts (as James H. Brusuelas characterises them) are to be trusted, the entire process must first be understood:

⁸ Here “born-virtual” should be understood as carefully distinguished from “born digital”. A born digital text is merely one created entirely digitally (for example, this article), whereas a “born virtual”

To ensure trust in the born-virtual text before us, we need to understand its virtual birth. We need to understand the data, i.e. the structured data describing and visualizing the entire process from start to finish. (Brusuelas, 2021: 68)

To promote equity of progress in AI-driven improvements in these areas, a collaborative approach to data sharing is essential. It is only by the pooling of resources and expertise in the academic domain that researchers can avoid redundant efforts, ensure reproducibility, and develop more robust models while ensuring academic concerns (rather than corporate ones) are foregrounded in this research. However, this requires the use of standardised practices: for example, in the case of digital palaeography, this means guidelines for transcription, including the handling of abbreviations, punctuation, and variations in spelling and capitalization. As it makes most sense for these to be stored in a structured form, and given its flexibility and diverse community support, it would make sense for this to be done using the TEI framework. Some of the basic tasks where AI might help scholarly editors includes the recognition and disambiguation of named entities, the parsing of texts for specific tasks, the collation of multiple textual witnesses, the analysis of textual content (especially for re-use and allusion), and the summarisation of documents. All of these will be considered in the sections below.

5. NAMED ENTITY RECOGNITION

There are numerous applications of AI which could significantly assist or enhance the work of scholarly editors. These include technologies that have been around for a while, but have recently improved significantly, such as Named Entity Recognition (NER). This makes it possible to automatically identify and label specific individuals, organisations, locations, creative works, and other named entities within a collection of texts. The evolution of NER has evolved in several distinct stages. Initial approaches relied heavily on complex rule-based systems which were usually then augmented by specialised lexicons for the general research area and languages in question. Subsequent developments included the integration of statistical analysis techniques, employing algorithms such as Hidden Markov Models and Conditional Random Fields to predict entity classifications with greater accuracy. However, the most profound expansion in NER capabilities emerged with the advent of deep neural networks, resulting from the developments in machine learning (*cfr.* Humbel *et al.*, 2021). Having been provided with vast datasets of training material, this has created sophisticated models that are capable of automatically discerning named entities from usual prose. While there are still limitations, especially

text is created by the virtual reassembly, generation, and differentiation of the text through opaque processes that are purporting to represent an analogue text which is unable to be seen directly.

for texts in languages for which less training has been done, this has provided a remarkable improvement in NER performance. The integration of deep learning has enabled NER to become a standard part of AI-driven text analysis, enabling the efficient and accurate extraction of key information from a wide array of textual sources. The growing sophistication and accessibility of NER technologies, especially where these are made to work in concert with LLMs, promises to significantly impact scholarly editing. As with any improvements created by AI, the automation of more time-consuming processes such as identifying and classifying named entities is claimed by its proponents to give editors more time to focus on higher-level tasks, such as textual interpretation, critical analysis, and the provision of editorial annotations. However, scholarly editors should be wary of just this kind of promise of removed drudgery with more time to focus on other tasks – the recent history of technological advances teaches us that this rarely comes without some significant downsides. The wholesale removal of scholarly editors from the process as a whole must strenuously be guarded against.

6. TEXT PARSING, COLLATION, AND ANALYSIS

One of the most time-consuming tasks for scholarly editors, particularly those producing critical editions of texts with many textual witnesses, is the collation of variant sources. Given the potential for AI to be used with NER and other forms of text analysis, it seems reasonable to assume that AI would help in the collation of parallel texts, identifying where the texts have similarities and where there are differences at any site of textual variance. Digital scholarly editors have long developed tools to facilitate such collation.⁹ However, these tools are often limited when working at scale, with complex, or problematically-variant texts. Moreover, they usually have difficulty with variation that extends over larger textual distances, for example that which results from fragmentation and reorganisation of a text between editions. Consequently, many scholars, such as Beshero-Bondar, create bespoke tools to address their specific collation needs for the texts that they are working with.

Even when it is machine-assisted, document collation is tiring, tedious work. It is one thing to prepare an algorithm for comparison and apply it to good, adaptable software for the purpose, but it is quite another to have to correct the output. That is where the real challenge begins – the intellectual challenge, mental discipline, or “self-psych-out” of “machine-assisted” collation: When do you give up trying to refine the software algorithm, and when do you crack and resort to hand-correcting problematic outputs? (Beshero-Bondar, 2023)

⁹ See, for example CollateX <https://collatex.net/>, which although it has well-known certain limitations is still considered a solid approach.

Frustrated by the challenges of developing a program to assist with the collation of materials for the *Frankenstein Variorum*, Beshero-Bondar faced the persistent dilemma that digital scholarly editors face i.e. when to refine the algorithm of a custom processing script that undertakes some semi-automated process, or just manually correct the output that it produces. Part of that tension comes from knowing that any manual corrections will need to be repeated if a substantial error in the work done by the script is discovered at a later date. Beshero-Bondar explored the potential of LLM-based chatbots to undertake some of the collation tasks. Using OpenAI's ChatGPT as her primary tool, she conducted a series of experiments, presenting the model with sections from *Frankenstein* and, in later tests, lines from *The Rime of the Ancient Mariner*. However, chat-interfaces to GPT-based LLMs currently turned out to be inadequate for such text-processing tasks. While the LLM could generate text that was coherent and similar to the source material, it struggled to accurately identify, align texts, or correct errors, particularly where there were any forms of complex textual variation. The chatbot often made mistakes in collation, occasionally mistaking minor variations as substantive differences but more usually failing to recognise significant textual variants at all. As Beshero-Bondar notes:

They are predictably unreliable, and never once did I see a response without errors. I also tried simplifying the task and asking the AI directly only to diff some strings, wondering if that word might be more familiar to the language model. But this made no difference and I have yet to see an accurate response to a prompt requesting a comparison of two or more strings. (Beshero-Bondar, 2023)

Despite these limitations, research such as this highlights the potential of AI-powered tools to assist in scholarly editing. As LLM technology continues to evolve, it is likely that future iterations will be better equipped to handle the complexities of textual collation. The current limitations of LLMs in textual analysis are understandable as they are rooted in the fundamental architecture of how GPT-based LLMs function. Improvements are likely to be based on infrastructural changes to the attention layers, hybrid forms of retrieval-augmented generation, or the integration of more sophisticated text parsing and analysis tools into chat interfaces. Explorations by Sewunetie and Kovács (2024: 38–802) suggest that the most beneficial approach is where “Hybrid Parser-based methods integrate multiple NLP components, combining rule-based and machine-learning techniques, to extract and represent semantic relationships from text”.

Beshero-Bondar's work on collation argues for the potential benefits of using specific declarative markup, such as TEI XML, for texts that are going to be collated. By embedding explicit human-defined semantics within the markup, editors could eventually control hybrid collation extensions to GPT-based AI, leading to more accurate and meaningful text comparisons and parsing. Such a strategy could help

mitigate the inherent limitations of current LLM models and enhance their ability to process complex textual data.

As the basic LLM interfaces have improved significantly since Beshero-Bondar undertook her work, I attempted similar experiments (in April 2025) using a variety of medieval textual witnesses of *Piers Plowman* and speeches from the late-medieval *The Digby Conversion of Saint Paul*. These resulted in similar kinds of issues when testing: text parsing and the provision of automated markup, glossing of hard words for a student edition, collation of multiple textual witnesses, the detection and explanation of text re-use and allusion, or the wholesale provision of editorial notes. As with Beshero-Bondar, I felt the tension between writing a (transparent, preservable) script to undertake the work and the results from LLMs. A final test was for the LLM to generate an XSLT stylesheet for processing the TEI XML I had given it, and provide links to the *Middle English Dictionary* for any words an undergraduate student would find difficult to understand. Although this kind of test was undertaken using over a dozen different approaches in prompting, in every case the resulting XSLT stylesheet would have needed modification by someone who already understood XSLT well enough to write such a stylesheet. A conclusion of this investigation might be that the trust one places in the output of LLMs should be directly proportional to one's ability to have created that output manually oneself (Cummings, 2025).

7. DOCUMENT SUMMARISATION

The increase in HTR-transcribed archival documents enables opportunities for both the archives and researchers using these documents. The careful cataloguing, summarisation of archival documents, and the production of relevant metadata that professional archival and special collections librarians create is a time-consuming but academically worthwhile process. However, LLMs excel at certain types of tasks, and while it is not always unproblematic, the summarisation of documents seems to be one of those that produces generally acceptable results. Currently, the main problems with LLM summarisation come usually in the LLM not recognising briefly-mentioned but important aspects as significant, or more usually failing to notice the absence of content which a researcher might expect to see. Nevertheless, AI can be used to generate summaries of transcribed archival documents, which could then be incorporated into library catalogue records. This can provide researchers with a more detailed and informative overview of the content of archival materials, making it easier to find relevant sources. While LLMs can generate catalogue records, it is still rare that it does so unproblematically (Moulaison-Sandy & Coble, 2024: 381). However, analysing a full text transcription and assigning classification metadata for discrete fields, such as Library of Congress Subject Headings, is an assistive approach which may be beneficial to the cataloguing community. However, when doing so, the potential for bias in LLMs must be recognised, alongside the lack of transparency

and control over their functions. Any AI-produced catalogue output will still require either careful human review and correction, or more likely (given resource-implications) at least strong flagging so that readers understand the provenance and unreliability of the inclusion of this alongside human-created metadata. The real question for collection cataloguers is whether the inherent lack of reliability is so much of a detriment that it outweighs the benefit of having the summary for readers. Where the HTR text of a document has been fed to an LLM in order to return a synopsis, the real sources of the statistical information that enables this summary are sometimes hard to disentangle because of the commercial service-based nature:

When the calculations and training capacities of a large language model are subject to rapid change with the next month's update, and when developers of generative language models conceal their sources for commercial reasons and do not share their transformer architectures openly, we would do well to inspect our tools and research methods for brittle dependencies. (Beshero-Bondar, 2023)

Part of the issue of any AI-based solution is the lack of transparency, not only in its sources of information, but also how the answer is derived. If we can't understand how the system processes data and arrives at its results, then use of it in any academic manner is problematic. The inherent opacity of LLMs poses a significant challenge to ensuring the accuracy and reliability of AI-generated material, including cataloguing data. The idea that this will be a panacea for all the digital hurdles of under-resourced library systems is just as controversial as its use in any academic endeavour:

Yes, finding new ways in which AI can support the work of librarians, especially technical services librarians like catalogers, will be critical to future success. However, given the rise of AI and the perception that it is able to solve specialized problems in cataloging easily, with the click of a button, if only the right prompt is created, is problematic to perpetuate – at least for now. (Moulaison-Sandy & Coble, 2024: 382)

8. AI AND SCHOLARLY EDITING KNOWLEDGE

Given the types of tasks that GPT-based LLMs appear to be better suited for, it is tempting to want to incorporate AI into the workflows of those undertaking scholarly editing. This is certainly useful for discrete tasks such as HTR and NER which benefit from being paired with LLMs. However, other methodologies such as digital palaeography do not naturally arise from the use of LLMs and would benefit from a model of the editorial decision-making process to be mimicked by the AI:

To be able to train an artificial intelligence system that can assist manuscript scholars in doing their research we need to provide the computer with the expert's knowledge: the so-called ground truth. (Busch, 2020)

In the case of scholarly editing the decision-making process is multi-variant, complex, and based in many different fields of learning. Editorial tasks range from the basic transcription (which might be assisted by HTR), through deep knowledge of the textual transmission, variation, and collation of the textual witness, to creating an edition that is a critical representation of the text suitable for the intended audience with glosses, textual notes, editorial notes and more. If the decision-making processes of scholarly editors were successfully modelled, then the AI-assistance would need to span this wide range of tasks for any text type encountered. With enough modelling data this may be possible, but it is difficult to gather given the varied nature of catalysts for an editor needing to make a decision. One way LLM-based approaches may be made more helpful is in the curation of a domain-specific model to provide more accurate responses influenced by the ingestion of specific texts.¹⁰ The more sophisticated version of this is the promising development of Retrieval Augmented Generation (RAG), where relevant information is retrieved from external knowledge bases to augment the LLM's dataset and thus potentially generate more accurate and contextually-relevant responses to user queries. However, for its use in the creation of digital scholarly editions, the creation and control of those external knowledge bases, should remain in academic not corporate control. Indeed, there has been some success in projects supplementing LLMs with substantial datasets of project-specific information which remains under their control. This approach could be used in the creation of a tool which tracks and records editorial decision-making, but to be successful it would need to be indispensable to the editorial workflow (and thus be used a lot) in order to gather sufficient data. Even so, there would be an understandable suspicion that scholarly editors might be helping to create a tool which then those unfamiliar with the complexity of their tasks mistakenly feel can replace them.

Indeed, if a large enough corpus of practical editorial decisions (not just what was decided, but how and why) was collected through an assistance interface then a semi-supervised learning AI approach could model the work of a scholarly editor. If such a model was used to enhance the hypothetical tool then the help given for editorial decisions would improve and feed back into a better model. While it is a mistake in my view if the development of AI tools for editorial assistance would eventually lead to scholarly editors being fully replaced by the very learning machines they help to create, the hiding of the underlying text encoding implied by the very existence of these tools also has its drawbacks. (Cummings, 2022: 155)

¹⁰ At its heart, this is the approach of tools like NotebookLM by Google, where a number of texts can be ingested and queried with footnotes given back to the sources of the answers in the provided corpus. See <https://notebooklm.google.com/> for more information.

The use of AI-based tools to develop timesaving assistive technologies for scholarly editors should not necessarily be resisted where it is beneficial to them, however, editors must be fully aware of the potential outcomes of their provision of data to such model-building activities. As already mentioned, the complexity of deep neural networks raises concerns about their inherent opacity. Even the creators of these systems often struggle to fully comprehend their inner workings and decision-making processes. The lack of model transparency poses challenges for the incorporation of AI into any scholarly editing workflow as it can be difficult to assess the reliability and potential biases of systems whose operations are not fully understood.

A tool like ChatGpt will certainly impact the way we organize and conduct teaching and research. What Bender *et al.* call the fundamental “unfathomability” of the deep neural architectures upon which most current AI systems and models such as GPT rely will continue to be problematic, but they are humanistic problems, as they concern the nature of discourse and the cultural biases and negotiations that shape our linguistic reality. (Mischke & Ohge, 2023: 57-58)

At the time of writing, it seems to remain general consensus (within the field of digital humanities at least) that AI tools should still need to retain the human as part of the workflow for such activities. It is the collaboration between human editors and AI tools which will be most beneficial, not the replacement of the “human in the loop”.

9. CONCLUSIONS

Digital scholarly editing projects often focus on a single manuscript or set of witnesses to a textual work. Sometimes they are valued for their textual or artistic contents, codicological importance, or cultural significance. However, HTR is potentially transformative not only for providing transcription of individual special manuscripts, but also first-stage transcriptions of complete collections of more general archival documents. Having searchable HTR-transcripts would help significantly with historical research of all sorts, even if the text is not error-free. Historical editing projects which excerpt and edit content relevant only to their collections would be faced with a much more straightforward task. For example, consider the work of the *Records of Early English Drama* project (REED), which has since 1976 sought out, transcribed, edited, and published (first in print, more recently digitally) records of “drama, secular music, and other popular entertainment in England from the Middle Ages until 1642”.¹¹ For such projects, archives for which a majority of the applicable manuscripts had been already transcribed would indeed be a boon. If

¹¹ See REED website: <https://ereed.org/>.

that was the state of archives in 1976, the REED project would have had a very different form. By the time the REED project completes its work (if indeed it ever does), a later revision of its work may be able to link REED project records back to their original context. We need to keep a critical eye on these developments and how scholarly editors' workflows may change and the possibilities that might enable.

It is possible to envisage a future of supervised machine learning that can aid in the editing of digital editions as well as their creation, assisting scholarly editors in their task, and expanding their purviews and the volume and scale of what they can achieve. (Terras *et al.*, 2024)

Nevertheless, it is important to distinguish carefully between sophisticated HTR text –presented with images, and perhaps with automated annotations of one sort or another– and actual digital scholarly editions where those have been provided by a human editor with critical and careful judgement. This may become more difficult to distinguish in the near future. It is one thing to accept the assistance of machine learning and generative AI tools, but quite another to count their outputs as editions, unless of course a human editor has subsequently edited them. While such tools may be useful, it is essential to approach them critically, recognising their limitations and the need for human oversight to ensure accuracy and reliability.

As for AI, identifying it and bringing it into our editions may or may not be necessary at this time. It is probably not the type of AI about which our science fiction induced imaginations think and dream. It is an intelligence that makes predictions. (Brusuelas, 2021: 68)

The “intelligence” of so-called “Artificial Intelligence” is inherently problematic. Currently, generative AI provides answers that are not always fit for purpose, but this will surely change as the commercial businesses behind generative AI solutions target new types of improvement. Moreover, as more sources and indeed more diverse kinds of sources (perhaps enabled by the HTR-driven handwritten textual paradigm shift), the LLMs which power AI solutions will inevitably get better at certain sorts of tasks. Nevertheless, at the time of writing, AI's mimicry of expected text in answer to prompting is so problematic that scholarly editors should remain very careful with the output from any discrete tasks. Beshero-Bondar noted this in her study of attempts at collation.

Perhaps we should not expect anything better. Today dialogue with generative language-based AI gives us the opportunity to declare and inquire with the voice of reason, but the stochastic outputs we receive sometimes contradict themselves and frequently miscalculate and misrepresent. We understand that

prompt generation is based on statistical predictions of what might be the best-fit, reasonable next tokens of text to supply in sequence, and that this makes generative language models not intelligent at all but rather stochastic machines. (Beshero-Bondar, 2023)

That gets to the heart of the limitations that we currently face when navigating the new frontiers of scholarly editing with LLM-based solutions. The problems of “how to mitigate the harms of LMs used as stochastic parrots” (Bender *et al.*, 2021: 619) is outside the domain of experience of most scholarly editors. As LLMs can easily misrepresent the texts being edited, this limitation currently means that the care which needs to be taken in their application often outweighs the benefits they provide. Scholarly editors –who are experts in the careful handling of text– need to continue to be part of the conversation around the use of LLMs in helping to critically understand historical documents. While the future is already here, it should be navigated with scepticism and care, so that this future continues to develop for the benefit of all.

REFERENCES

- Allés-Torrent, S., & del Rio Riande, G. (2019). The Switchover: Teaching and Learning the Text Encoding Initiative in Spanish. *Journal of the Text Encoding Initiative*, (12), 1–29. <https://doi.org/10.4000/jtei.2994>
- Andrews, T. L., & van Zundert, J. J. (2018). What Are You Trying to Say? The Interface as an Integral Element of Argument. In R. Beier, M. Bürgermeister, H. W. Klug, F. Neuber, & G. Schneider (Eds.), *Digital Scholarly Editions as Interfaces* (pp. 3–33). Universität zu Köln., <https://kups.ub.uni-koeln.de/9106/>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 2. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)* (pp. 610–623). Association for Computing Machinery. <https://doi.org/https://doi.org/10.1145/3442188.3445922>
- Beshero-Bondar, E. (2023). Declarative Markup in the Time of “AI”: Controlling the Semantics of Tokenized Strings. *Proceedings of Balisage: The Markup Conference*, 28. Advance online publication. <https://doi.org/10.4242/BalisageVol28.Beshero-Bondar01>
- Brusuelas, J. (2021). Scholarly Editing and AI: Machine Predicted Text and Herculaneum Papyri. *Magazén*, 2(1), 45–70. <https://doi.org/10.30687/mag/2724-3923/2021/03/002>
- Busch, H. (2020). *An Artificial Eye for Palaeography. Applying Deep Machine Learning for the Study of Medieval Latin Scripts*. Schoenburg Symposium 2020. <https://doi.org/10.5281/zenodo.4302209>
- Cummings, J. (2014). The Compromises and Flexibility of TEI Customisation. In C. Mills, M. Pidd, & E. Ward (Eds.), *Proceedings of the Digital Humanities Congress 2012*. Studies in the Digital Humanities. HRI Online Publications., <https://www.dhi.ac.uk/books/dhc2012/compromises-and-flexibility-of-tei-customisation/>

- Cummings, J. (2019). Opening the Book: Data Models and Distractions in Digital Scholarly Editing. *International Journal of Digital Humanities*, 1, 179–193. <https://doi.org/10.1007/s42803-019-00016-6>
- Cummings, J. (2022). The Present and Future of Encoding Text(s). In J. O'Sullivan (Ed.), *The Bloomsbury Handbook to the Digital Humanities* (pp. 147–157). Bloomsbury Publishing., <https://doi.org/10.5040/9781350232143.ch-14>
- Cummings, J. (2023). Academics Retire and Servers Die: Adventures in the Hosting and Storage of Digital Humanities Projects. *Digital Humanities Quarterly*, 17(1). <https://dhq.digitalhumanities.org/vol/17/1/000669/000669.html>
- Cummings, J. (2025) Encoding Critical Apparatus in an Age of AI, slides from presentation at New Perspectives on Critical Editions. GREN-CRIHN workshop. <https://doi.org/10.5281/zenodo.15128308> as part of <https://crihn.openum.ca/nouvelles/2025/03/13/workshop-new-perspectives-on-critical-editions-part-2/>
- Humbel, M., Nyhan, J., Vlachidis, A., Sloan, K., & Ortolja-Baird, A. (2021). Named-Entity Recognition for Early Modern Textual Documents: A Review of Capabilities and Challenges with Strategies for the Future. *The Journal of Documentation*, 77(6), 1223–1247. <https://doi.org/10.1108/JD-02-2021-0032>
- Kecskeméti, G. (2023). Humanist Texts in a Digital Age. *Camoenae Hungaricae*, 8, 5–20. <http://real.mtak.hu/id/eprint/185420>
- Mischke, D., & Ohge, C. (2023). Digital Melville and Computational Methods in Literary Studies. *Leviathan*, 25(2), 35–60. <https://doi.org/10.1353/lvn.2023.a904374>
- Moulaison-Sandy, H., & Coble, Z. (2024). Leveraging AI in Cataloging: What Works, and Why? *Technical Services Quarterly*, 41(4), 375–383. <https://doi.org/10.1080/07317131.2024.2394912>
- Nockels, J., Gooding, P., & Terras, M. (2024). The Implications of Handwritten Text Recognition for Accessing the Past at Scale. *The Journal of Documentation*, 80(7), 148–167. <https://doi.org/10.1108/JD-09-2023-0183>
- Parker, C. S., Parsons, S., Bandy, J., Chapman, C., Coppens, F., & Seales, W. B. (2019). From Invisibility to Readability: Recovering the Ink of Herculeaneum. *PLoS One*, 14(5). Advance online publication. <https://doi.org/10.1371/journal.pone.0215775>
- Pinche, A., & Stokes, P. (2024). Historical Documents and Automatic Text Recognition: Introduction. *Journal of Data Mining & Digital Humanities*. <https://doi.org/10.46298/jdmdh.13247>
- Retsinas, G., Sfikas, G., Gatos, B., & Nikou, C. (2022). Best Practices for a Handwritten Text Recognition System. In S. Uchida, E. Barne & V. Eglin (Eds.), *Document Analysis Systems, Proceedings of 15th IAPR International Workshop* (pp. XX-XX). <https://doi.org/10.1007/978-3-031-06555-2>
- Sewunetie, W. T., & Kovács, L. (2024). Exploring Sentence Parsing: OpenAI API-Based and Hybrid Parser-Based Approaches. *IEEE Access : Practical Innovations, Open Solutions*, 12, 38801–38815. <https://doi.org/10.1109/ACCESS.2024.3360480>
- Stokes, P., & Kiessling, B. (2025). Sharing Data for Handwritten Text Recognition (HTR). In C. Crompton, L. Estill, R. J., Lane & R. Siemens (Eds.), *The Companion to Digital Humanities in Practice* (pp. XX-XX). Routledge <https://www.taylorfrancis.com/books/oa-edit/10.4324/9781003327677/companion-digital-humanities-practice-ray-siemens-laura-estill-constance-crompton-richard-lane?context=ubx&refId=efic5dcc-6016-4aaf-8d7f-57a2e64637b0>

- Terras, M., Nockels, J., Ames, S., Gooding, P., Stauder, A., & Mühlberger, G. (2024). On Automating Editions: The Affordances of Handwritten Text Recognition Platforms for Scholarly Editing. *Scholarly Editing*, 41. Advance online publication. <https://doi.org/10.55520/W257A74E>
- Trowsdale, E. (2023). 'This Cloud Had Been Written upon': Investigating Digital Editing Methods for Anne Bathurst's Visionary Writing [MSc Dissertation]. University of Oxford.



FACULTAD DE FILOLOGÍA
UNIVERSIDAD DE SEVILLA