

PROBING LLMs' 'INTERPRETATION' OF
CHAUCER'S *CANTERBURY TALES*:
REFLECTIONS ON THE USE OF GENERATIVE AI
FOR DIGITAL SCHOLARLY EDITING & THE
UNIVERSITY CLASSROOM

Sophie Whittle
Digital Humanities Institute
University of Sheffield

sophie.whittle@sheffield.ac.uk
@sophiewhittle.bsky.social

On behalf of the *C21 Editions* project, funded by the AHRC-
IRC/UK-Ireland collaboration in the Digital Humanities.
AH/W001489/I and IRC/W001489/I

Questions

1. What are the issues when using LLMs to respond to medieval texts, from an editorial and pedagogical perspective?
2. Who should have their say on the development of LLMs?

Key take-home messages:

1. Collaboration between disciplines and people with varying expertise
2. Accessibility for wider audiences
3. Interrogation as a solution



Portrait of Chaucer by Thomas Hoccleve in the Regiment of Princes (1412), Public Domain, Wikimedia Commons.

WHAT ARE YOUR CONCERNS ABOUT
GENERATIVE AI?



GO TO:
MENTI.COM



TYPE IN CODE:
7173 0185

NORBOT TAKES OVER GROMIT'S GARDEN!
WALLACE & GROMIT: A VENGEANCE MOST
FOWL



YouTube: Netflix After School

LARGE LANGUAGE MODELS & CRITICAL AI

- Recent hype & fear around AI (Goodlad 2023, Stone et al. 2024)
 - LLMs are black boxes: training data & sources, billions of parameters, homogenised responses based on probabilistic scoping
 - Reinforcement Learning from Human Feedback (RLHF)
- Ethics: authorship/plagiarism, environmental impact, human labour exploitation, hallucinations & stereotypes
- Importance of sociolinguistics (Schneider 2024)
- Recommend: Katie Conrad's "[A Blueprint for an AI Bill of Rights for Education](#)", Emily Bender et al.'s (2021) "[On the dangers of stochastic parrots: Can language models be too big?](#)" & the [Algorithmic Justice League](#) led by Joy Buolamwini along with her TED Talk (2017)



Adapted from Stone et al. (2024), figure 1



Prompted by Erin Bow (@erinbowbooks.bsky.social)

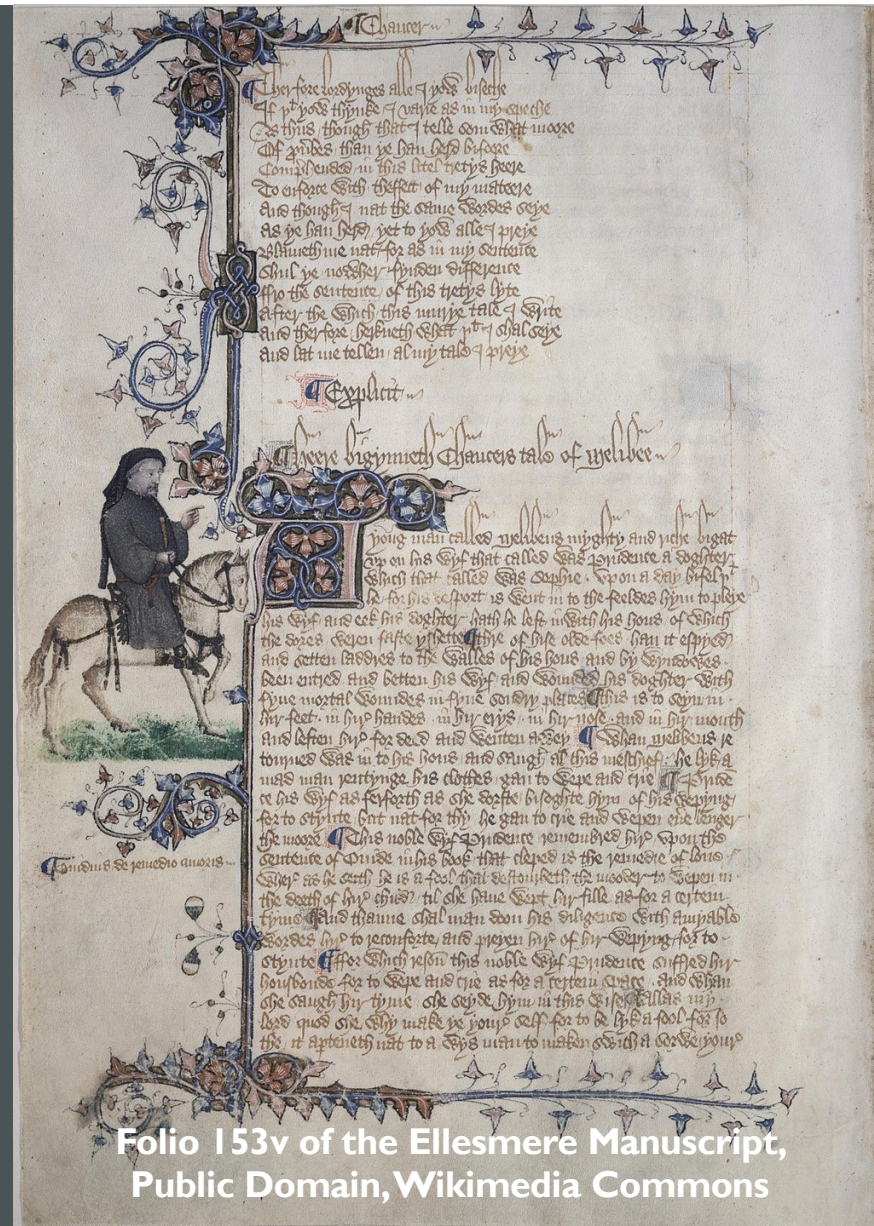
IMPACT ON INTERPRETATION OF MEDIEVAL TEXTS

What might genAI be able to do for digital editing and pedagogy?

- Automated transcription (e.g. ATR) and translation
- Summarisation of salient themes and topics
- Free up editors time for creativity & ideas for interactive critical apparatus/tech development

But: what are the issues/challenges?

- Training primarily based on modern data
- Researchers aware of many of the nuances already
- Not suitable for those lacking in experience to probe responses



Folio 153v of the Ellesmere Manuscript,
Public Domain, Wikimedia Commons



Tete d'une Femme Lisant, Pablo Picasso

C2I EDITIONS: INITIAL AIMS

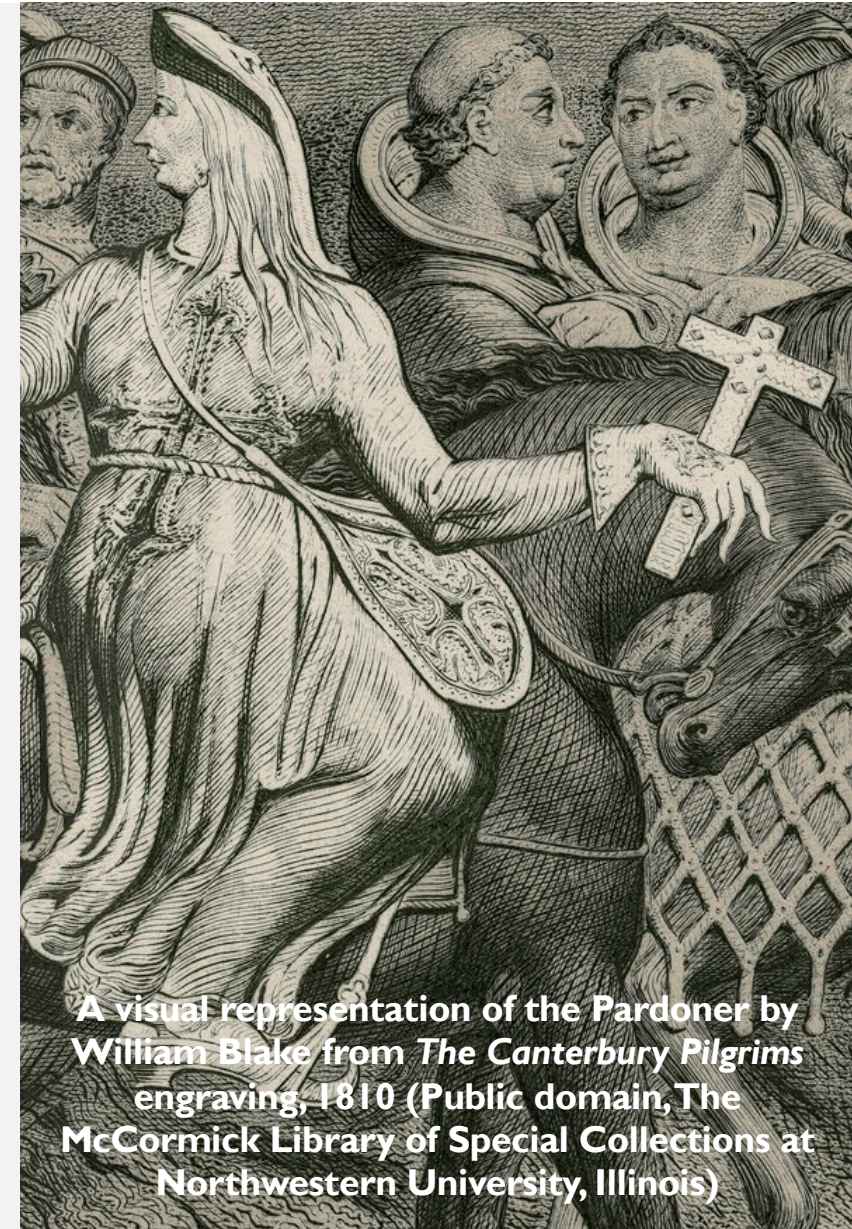
- How Machine Learning and Natural Language Processing can improve the process of edition-making
- Open access edition, drawing on open access knowledge
- Reproducible framework
- Inclusive design workshops/focus groups
- Later: focus on generative AI and impact on digital pedagogical editing

WHY CHAUCER AND THE PARDONER?

- Chaucer: The father of English literature?
- Pardoner and his tale: transgressive, resisting and disrupting the norms of medieval society & the Catholic church
- Sarah Baechle & Carissa M. Harris (2021): ethical challenges of Chaucerian scholarship
- Ruth Evans (2022): reimagining Chaucer's literature

Possible approaches:

- Queer DH (Jason Boyd and Bo Ruberg 2022): Estranging, skeptical, subversive, social, sensual
- Critical race theory & intersectionality (Rahul Gairola 2022): Need to account for colonialist histories and decolonial approaches to probe AI



A visual representation of the Pardoner by William Blake from *The Canterbury Pilgrims* engraving, 1810 (Public domain, The McCormick Library of Special Collections at Northwestern University, Illinois)

METHODS

Conducted initial preliminary focus groups and interviews (49 theorists/practitioners, 13 students, 6 teachers) to feed into development of edition.

Different features of edition: translation, annotation, **background/topic guide and summary generation**, recording of the tale, manuscript recognition.

For each feature, we asked the overarching question: “Does the use of AI technology enrich the experience of digital editing and digital pedagogy?” from the perspectives of editors, teachers and students.

STUDENT & TEACHER INSIGHTS

"I try not to underestimate how curious and intelligent my students are because often they come up with brilliant questions and really interesting insights, and that really happens when you give the floor to them and when you give them things to work on."

"I also think it's part of a larger historical conversation about this movement from investigating people that are not typically investigated and making them part of the narrative."

"It is vital I tell my students about ways in which, in the 19th Century American/British literature curriculum, Black students and Black educators were trying to find ways of engaging with Chaucer, and were trying to counter some of the narratives of 'Anglo-Saxon' whiteness..."

"There's a lot of this unquestioned inherited knowledge that I find that my students can often just push through. We look up a word in the Oxford English Dictionary and they're like 'why can't it mean this instead?'"

"It's going against the grain of the idea of pilgrimage. There are religious characters, there's a prioress, there are nuns, and there's a parson – so there are religious representations – but they're going against the norm of pilgrimage. Doesn't that immediately put them into a queer space?"

"Medieval literature is actually a lot more disruptive than you would think and has a lot more to say about present times than we think."

"It gives you such insight into all different corners of everyday medieval life in terms of giving him a voice to all these radically different, socially stratified characters"

STUDENT & TEACHER INSIGHTS

“There’s a valuable way of learning in reverse engineering basically. And you’re taking on that typical marking role and you’re having to understand everything from a different perspective.”

I guess it’s about reliability. Is it representative? [...] But I wonder if that’s what we’ll arrive at— ‘this is what a human could do, isn’t that nice?’”

I definitely could see, especially in a translation sort of environment [...] how that [critiquing genAI] could be really illuminating for the student. If they were to do that, I think it would foster a lot of critical thinking skills about the text [e.g. Chaucer] which is a very difficult one to actually get into.”

“Maybe incorporating the Middle English Dictionary into ChatGPT rather than just opening up a glossary—I think your AI is the glossary itself.”

I think he [the teacher] is really curious about the implications of AI and how that’s going to change how we are able to convey information with one another, how we process information.”

“...all the stuff about ChatGPT and its usefulness really does emphasise how much you need a human to be able to sort through these things. But if I were teaching translation of Chaucer in the classroom, I think a great thing to do would be to actually use ChatGPT and then ask the students to think about why these translations are inadequate.”

EDITOR INSIGHTS: C2I INTERVIEWS

“Simply by taking the formidable power of deep learning methods and techniques, and AI techniques you will be able to do much more interesting things with your digital text than before. [...] But once this is possible say in five years or even sooner, would you really, as a scholar, still want to be doing that work manually? Again, there will always be these tiny corner cases that will require this kind of manual labour. But in all other cases I think you want to use your time for more challenging aims in that case.” (Joris van Zundert)

“I think it’s that thing about in DH of let machines do what machines can do well and let humans do what humans do and creative work, creating a textual environment where the automatic transcription can then be corrected and improved. But then simultaneously that leaves the scholar to focus on the task of devising more interesting and intensive and hybridised and creative ways of interpreting the text.” (James Louis Smith)

“NLP approaches mixed with HTR and stuff are about to make a second revolution on the way we do things, and there are colleagues of mine here in France theorising about computational philology, computational editing and not digital editing which is a completely different approach.” (Elena Pierazzo)

EDITOR INSIGHTS: C2I INTERVIEWS

“I do worry about machine learning from an environmental perspective; it's hugely resource intensive and I have to wonder how important it is in some contexts to do machine learning just for a kind of 'oh isn't this interesting' sort of result when you're burning up fossil fuels to generate these things. [...] editing itself is kind of computational in that it's always looking for patterns and outliers and doing quantitative analysis.” (Christopher Ohge)

“What do you gain in saved cost versus what you lose in the by-products of human transcription, in the sense of new research questions being formed by the person transcribing, new knowledge being imparted to the person transcribing?” (Ben Brumfield)

“Every new technology is immediately welcomed and embraced: how can we benefit from this innovation and improve our research? Can we address new research questions to the material, to new material available?” (Franz Fischer)

“I just think it's very important that we as a scholarly community are not too keen on simply adopting software that automates or processes, because if we [...] try to critically analyse the different transformations that our text undergoes when it's being processed by such a tool, we can learn so much.” (Elli Bleeker)

ISSUES WITH LLM RESPONSES FOR CHAUCER/MEDIEVAL TEXTS

Input (Bing Chat, 9 October 2023): Provide the meaning and etymology of the following words from Chaucer's Pardoner's Prologue and Tale, to a university level standard, using information from the Oxford English Dictionary and the Middle English Dictionary. Ensure to include information on its meaning in the 1300s and within the context of the tale, and details on the origins of the word and whether it was borrowed from another language:

ioly wenche (And haue a ioly wenche in euery toun)

Output: I'll do my best to provide the meaning and etymology of the words you have requested from Chaucer's Pardoner's Prologue and Tale.

ioly wenche: In Middle English, "ioly wenche" referred to a pretty young woman. [In the context of the Pardoner's Tale, it is used in reference to a woman who is kept by one of the three rioters as a mistress.](#) The word "ioly" has its roots in Old French, where it was used to refer to something that is cheerful or merry.

ISSUES WITH LLM RESPONSES FOR CHAUCER/MEDIEVAL TEXTS

Input (ChatGPT, version 4, 14 August 2023): Can you summarise the theme of gender and sexuality in Chaucer's Pardoner's Prologue and Tale, to a university level standard?

Output:

Gender and sexuality in Chaucer's Pardoner's Prologue and Tale are complex and multifaceted. The Pardoner himself is a controversial figure whose ambiguous gender and sexual identity have long been subjects of critical debate.

The Pardoner's **ambiguous gender presentation** is apparent in his physical description in the "General Prologue," where he is described as having hair as yellow as wax, hanging as smoothly as a hank of flax, and a voice as small as a goat. He is also said to be unable to grow a beard, which was a significant marker of masculinity in the Middle Ages. These descriptions suggest a certain **androgyny or effeminacy**, which is **further emphasized by the fact that he carries a bag of relics, an occupation typically associated with women in the Middle Ages.**

Sexuality is also a significant theme in the Pardoner's Prologue and Tale. The Pardoner's sexuality is hinted at in the text, with some critics interpreting him as a eunuch or a homosexual, based on his effeminate appearance, his **apparent lack of sexual desire for women**, and his close relationship with the Summoner. The Pardoner's Tale itself **also deals with sexual sins**, as the three rioters in the tale are guilty of gluttony, gambling, and lechery.

Moreover, the Pardoner's **hypocritical preaching** against the very sins he himself indulges in can be seen as a commentary on the corruption of the Church and its officials, who often violated the vows of chastity, poverty, and obedience they were supposed to uphold.

In conclusion, the themes of gender and sexuality in the Pardoner's Prologue and Tale are **intricately linked with the broader themes of hypocrisy, corruption, and moral decay in the Church and society.** The Pardoner's ambiguous gender and sexual identity not only challenge the rigid gender norms of the time, but also serve as a critique of the **moral failings of religious officials.**

ISSUES WITH LLM RESPONSES FOR CHAUCER/MEDIEVAL TEXTS

Input (ChatGPT, version 4o, 18 November 2024): You are a **scholarly editor of Geoffrey Chaucer's text**, The Pardoner's Prologue and Tale from The Canterbury Tales, with an interest in pedagogical approaches to the text. Your task is to identify and annotate words from the tale from an **intersectional feminist and queer lens**. Ensure the output is suitable for **undergraduate university students**.

Output (Smaller extracts taken for analysis):

"I trowe he were a geldyng or a mare" (General Prologue, line 691, describing the Pardoner)

Annotation: This line has sparked debates about the Pardoner's gender and sexual identity. The ambiguous reference to the Pardoner as possibly a "gelding" (castrated male) or "mare" (**feminized term**) destabilizes fixed notions of gender and sexuality. Feminist and queer interpretations highlight this ambiguity as a space for **questioning medieval norms of masculinity and the rigid binary of male/female**.

Discussion Points for Pedagogy:

Queering the Pardoner: How does the Pardoner's ambiguous gender presentation and sexual identity invite discussions about medieval perceptions of queerness? How do these perceptions align with or differ from contemporary understandings of **gender fluidity and nonconformity**?

Gendered Morality: How are sins like greed, gluttony, or avarice **gendered** in the text, and what might this suggest about medieval attitudes toward morality and gender roles?

Institutional Hypocrisy: What parallels can be drawn between the Pardoner's critique of sin and his embodiment of it, particularly regarding **feminist critiques of institutional power**?

Greed and Colonial Narratives

The "Three Rioters" and Stereotypes of Disorder

Textual Reference: "**Sworn brotherhood in filthee rage and pride**" (Tale, line 709).

Analysis: The three rioters embody moral and social disorder, engaging in drunkenness and violence. Medieval European narratives often associated social chaos with groups perceived as "other," such as non-Christian or non-European peoples. **The alignment of moral decay with "outgroups" can be interrogated for its proto-racial undertones.**

A GENAI EDITION: WHAT ARE THE PROBLEMS FOR DIGITAL EDITING & PEDAGOGY?

Annotation: nuance added by the editor

The Pardoner's Prologue and Tale


The Ellesmere Manuscript Background to the Text Guides to the Text

The Pardoner's Prologue and Tale

by Geoffrey Chaucer

Show Annotations

Original Text Modern Translation Annotations



138r

123 Al were it yeuen of the pouereste **page**

124 Or of the pouereste wydwe in a village

125 Al sholde hir children sterue for famyne

126 Nay I wol drynke **licour** of the vyne

127 And haue a **joly wenche** in euery toun

128 But herkneth lordynges in conclusioun

129 Your likyng is that I shal telle a tale

130 Now haue I dronke a draughte of corny Ale

131 By god I hope I shal yow telle a thyng

132 That shal by resoū been at youre likyng

133 For though my self be a **ful vicious man**

118: ioly wenche

In Middle English, a *ioly wenche* referred to a pretty young woman, yet frequently with a negative connotation. Here it is used in reference to a woman who might be kept by the Pardoner as his mistress. The word *ioly* has its roots in Old French, where it was used to refer to something that is cheerful or merry. *wenche* comes from Old English *wenche*, and in the [OED](#) its initial meaning is "a child (of either sex); also, a servant or slave; also, a common woman." In the 17th century, the word came to be racialised and used against enslaved Black women (e.g. see research by Carissa Harris).

Transcription Lines: [127](#)

Translation Lines: [127](#)

Tags: Medieval words and phrases

[View this Annotation in the Annotation List.](#)

Chaucer Edition prototype (C21 Editions project), landing page

A GENAI EDITION: WHAT ARE THE PROBLEMS FOR DIGITAL EDITING & PEDAGOGY?

The Pardoner's Prologue and Tale

The Ellesmere Manuscript Background to the Text Guides to the Text Activities Source Book About the Project

The Pardoner's Prologue and Tale
by
Geoffrey Chaucer

Show Annotations Original Text Modern Translation Annotation List Topic List



140v

360 By seinte Marie seyde this Tauerner
361 The child seith sooth for he hath slayn this yeer
362 Henne ouer a Mile with Inne a greet village
363 Bothe man and wōman child & hylde & page
364 I trowe his habitaciōū be there
365 To been auysed greet wysdom it were
366 Er that he dide a man a dishonour
367 Ye goddes Armes quod this Riȝour
"By God's arms!" said this rioter,
368 Is it swich peril with hym for to meete
369 I shal hym seke by wey and eek by Strete
370 I make auow to goddes digne bones

Chaucer Edition prototype (C21 Editions project), landing page

Topic summaries: few in-text examples generated, and linked to a topic guide (also enriched by the editor)

Notions of Masculinity

The tale which the Pardoner tells is devoid of any women characters, focusing instead on the masculine world of the three rioters. Their aggressive masculinity is displayed through their violent actions and their quest to conquer Death, which ultimately leads to their own downfall. Their hyper-masculinity is contrasted with the Pardoner's fluid gender identity, further highlighting the intricacies of the exploration of gender and sexuality in the tale.

Topic: [2. Gender & Sexuality](#)

[View this Topic in the Guide to the Text](#)

Chaucer Edition prototype (C21 Editions project), Notions of Masculinity

Notions of Masculinity

The Pardoner's 'transgression' is evident from a number of specific moments in the prologue and tale, which could be seen as disruptive of some of the norms of the time.



The Canterbury pilgrims depicted in a mural by Ezra Winter (1939). It is based in the North Reading Room, west wall, Library of Congress John Adams Building, in Washington D.C. (Photographed by Carol Highsmith in 2007 who stipulates that her photographs are in the public domain, [Wikimedia Commons](#)).

The relics that the Pardoner holds are said to symbolise many different aspects of his identity. Dinshaw (1989: 164) first considered the objects as euphemistic for the Pardoner's 'lacking' physicality, yet, more recently, scholars have suggested the objects harbour the Pardoner's sexual power, which he uses against other pilgrims in the tale.

The Pardoner's focus on his relics to alleviate others of their sin might carry connotations of sexual potency, especially the way in which he uses them to humiliate or outwit others. Zeikowitz (2002: 56) states that "the Pardoner's gender and sexuality are a problem precisely because we can find neither his 'gender core' nor his 'sexuality core'", and instead, the Pardoner "challenges the idea that masculine normalcy has an exclusive claim to power." Despite the way in which the Pardoner's actions and behaviour are stigmatised, he is able to reconstruct a new perception of himself, which encompasses his sharp wit, preaching abilities, and enticing nature.

A GENAI EDITION: WHAT ARE THE PROBLEMS FOR DIGITAL EDITING & PEDAGOGY?

I. Editor required to write guidance for the discussion of potentially sensitive subjects

Guidance for Discussions of the Pardoner

As we have discussed in the previous pages, many views have been provided about the Pardoner's character and body, which have real-world consequences for individuals existing and living in society today. It is therefore important in class (whether you are a student or a teacher) to approach the topic with sensitivity and care for others. As recent literature has shown, it is helpful to view the Pardoner's character as limitless, not from an 'either/or' perspective. Further suggestions about the Pardoner's physical being, as well his sexuality – particularly those arising from a normative, patriarchal sphere – are potentially harmful for individuals in the past, present, and future.

We encourage you to approach the Pardoner's character with an open mind, focusing on the range of ways the Pardoner presents himself to his audiences, such as the sharp wit, foolery and playfulness that he exudes. We have included further open access reading below that you may want to use in the classroom, for assignments, or teaching. Note that these openly accessible resources are unfortunately limited, which is why we have summarised some recent literature in the previous sections. We aimed to show here that a queer reading of the Pardoner sheds light on a figure who challenges traditional, patriarchal, and strict binary modes of thinking, and highlights the endless forms of 'being' in both medieval and modern society.

2. How do we cite 'GenAI editions'?

AI Tools

OpenAI. 2024. 'Gender and sexuality of the Pardoner and in Chaucer's "Pardoner's Prologue and Tale", and the history of homosexuality in medieval Europe' prompt. ChatGPT-4 (Playground) [Large language model], generated August-December 2023, <https://platform.openai.com/playground/chat>. In *The Pardoner's Prologue and Tale: A Generative AI Edition for Teaching*, edited by Sophie Whittle, Matthew Groves and Michael Pidd. Digital Humanities Institute. Available at <https://www.dhi.ac.uk/c21-chaucer-interface-prototype/topic/gender-sexuality>

AI & Wikipedia Content

The content on this page was generated using AI alongside an editor. Data from the following Wikipedia articles was also rewritten by generative AI where applicable: [The Pardoner's Tale](#), [Homosexuality in medieval Europe](#). Please treat the text on this page with caution and refer to our [source book](#) for further information about how AI was prompted to generate specific sections of this page.

Chaucer Edition prototype (C21 Editions project),
Gender and sexuality of the Pardoner

COLLABORATIVE SOLUTIONS

Critical, flexible and collaborative translation:

- Different translation possibilities
- Annotations
- Initial suggestions for discussion points
- Follow-up discussion boards
- Similar to 'Beowulf By All': a community translation and workbook providing opportunities for contribution from multiple lived experiences

BEOWULF BY ALL

Heyla!¹ We have a story about the Spear-Danes, from the old days when they were big and their kings showed their strength. There was one king, Shield Schefing, who stole many mead-benches from other tribes and terrified their leaders. At first, he was found weak and wandering, but was taken in and then grew under the comfort of the skies. He consumed honors until each of the other surrounding tribes over the whale's road were forced to obey him and pay tribute. They say, that was a good king.

After all this (when he was old), Shield had a son—a young one in the courtyard—who had been sent by God as a comfort to the people because He had seen how they were distressed, left without a strong leader for a long while.

Beowulf By All (Abbott, Treharne & Fafinski 2021: 18)

Ioly Wenche

The Pardoner discusses what he would rather do with his riches in his prologue:

126 Nay I wol drynke licour of the vyne

127 And haue a *ioly.wenche* in euery toun

Transcription: And haue a **ioly wenche** in euery toun

Translation 1: **And have a jolly wench** in every town.

Translation 2: And have a **pretty wench** in every town.

Help [collapse]

Carissa Harris states that, through time, 'wench' came to be known as a 'structural position', rather than a 'gender identity'. She states that "the label of wench is capacious enough to include many, yet binding enough to compel to coerce and to constrain" (Harris 2023). Its roots are the product of colonialism and slavery, and the modern-day use of the term is embedded in racialised misogyny arising from the 17th century.

The medieval use of the term is also encompassed by negative connotations of women in society. In the OED, the term 'wench' is derived from Old English 'wencel', a now obsolete meaning to refer to "a child (of either sex)' or 'a servant or slave; also, a common woman" (OED Online, 'wencel (n.)' 2024). Harris (2023) explains that 'wench' is used in Chaucer for the purposes of control and coercion, domination, and to denote masculine possession, and is often used to justify violence against those with a degraded social position.

The same could be said of the use of 'wench' in the Pardoner's prologue. In this case, the use of *ioly* as a premodifier for 'wench' objectifies the woman's body as central to the Pardoner's pleasure. In Translation 2, *ioly* is translated as 'pretty', which may relate to the physical sense in the OED of 'having the freshness and lively spirits of youth or good health; fresh, lively, sprightly, spirited' (OED Online, 'jolly (adj. & adv.)' 2023), or in the MED where *joli* is presented as "pleasant, beautiful, pretty, elegant, excellent". While the Pardoner might be simply expressing his sexual desires and sin, regardless of his 'responsibilities' within the church, Chaucer's use of the term in this context appears to be embedded in a societal expectation that women be controlled and objectified.

Chaucer Edition prototype (C21 Editions project), Student Translation Task I

COLLABORATIVE SOLUTIONS

Chaucer Edition prototype (C2I Editions project), translation discussion board

The Pardoner's Prologue and Tale (PPT): Translation Activity discussion board

What do you make of the translations used in either this edition, or elsewhere? Would you make any changes? Comment in the discussion board below; we would like to hear your thoughts! (You may post with your name or anonymously)

Nouns and terms of address | **The Pardoner's character** | **ChatGPT translation outputs**

Example post 1
Middle English, line 3: LOrdynge's quod He in chirches when I preche.
ChatGPT + editor translation: "Lords," he said, "when I preach in churches.
Harvard translation: "Gentlemen," he said, "in churches when I preach.

Example post 2
Middle English, line 4: I peyne me to han an hauteyn speche
ChatGPT + editor translation: I strive to have a haughty speech,
Harvard translation: I take pains to have a loud voice, when I preach.

Example post 3
Do you have any other comments related to ChatGPT's translation of the tale?

Discussion boards

Programming Historian, tutorials for humanists

Enter *Programming Historian*
(The Initial English version)

[PH]

111 lessons | ISSN: 2397-2068

We publish novice-friendly, peer-reviewed tutorials that help humanists learn a wide range of digital tools, techniques, and workflows to facilitate research and teaching.

Network building

READ-COOP to improve access to historical documents via Transkribus and a European Cooperative Society

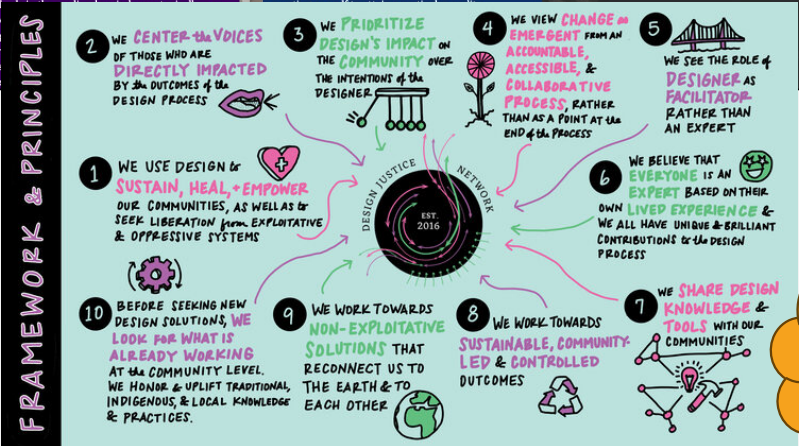
READ COOP | Transkribus | ScanTent | read&search | About | Resources | Plans & Pricing | Sign In | App

We revolutionise Access to Historical Documents

Our European Cooperative Society has more than 100 Members globally.

AI-powered Handwritten Text Recognition with Transkribus

Transkribus is a comprehensive solution for the digitisation, AI-powered text recognition, transcription and searching of historical documents.



Iterative and inclusive design processes

AI PEDAGOGY PROJECT

metaLAB (at) Harvard

Harvard AI Pedagogy Project: "We are at an inflection point in the development of AI tools—who will decide what capabilities are developed, by whom, and to what end?"

marinelives-collaboratory

Repository for MarineLives Collaboratory

To accompany Colin Greenstreet's [November 19th 2024 talk](#) at the IHR Digital History seminar on [Machine Learning and Historical Research](#), MarineLives is launching an online collaboratory. We welcome non-coding and very-low-coding historians with open arms.

The collaboratory is targeted at masters and doctoral students working in the field of history, who would like to explore how large language model based machine learning techniques can be incorporated into their research design and research processes. Each session will be built around [specific Historical Research Use Cases](#). Participants are invited to bring their own use cases for discussion.

We will look at a wide range of analytical frameworks, subject matter and periods, directed by the research interests of participants in the collaboratory, and will explore how large language models can assist with the design and conduct of different types of research.

The goals of the collaboratory are to:

1. Build useful research tools for real historical research use cases which can be put into immediate effect
2. Develop and document shared knowledge of large language models applied to historical research use cases
3. Build a community of doctoral students interested in large language model applications to historical research

Collaboratory by Colin Greenstreet for students to gain understanding of machine learning techniques for research in digital history

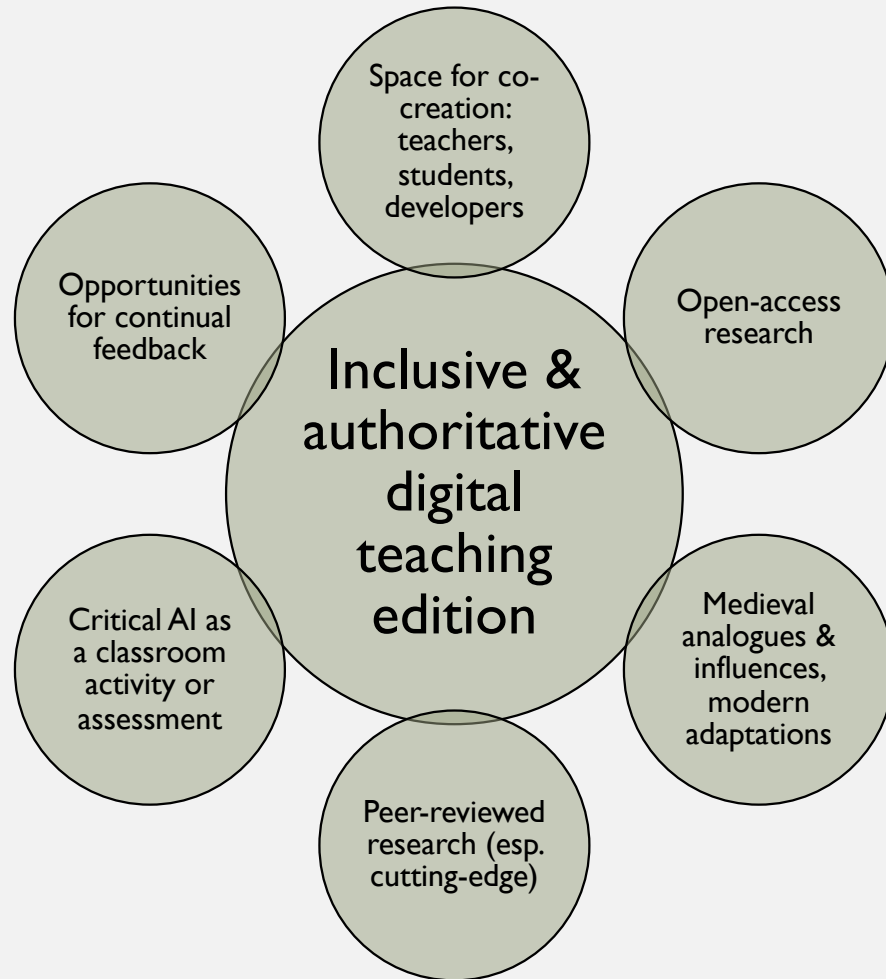
Design Justice Network principles, from the Philanthropy Workshop, 14 April 2021

PROBING LLMs: WHAT MIGHT IT LOOK LIKE IN A CLASSROOM ENVIRONMENT?



1. [Critical AI Literacies: A Guide for Students](#), Design Justice Labs, Rutgers University
2. [Teaching Critical AI literacies document](#)

DIGITAL PEDAGOGICAL EDITION FRAMEWORKS



Returning back to RQs:

1. What are the issues when using LLMs to respond to medieval texts, from an editorial and pedagogical perspective?
 1. Lack of nuance and specificity
 2. Not always able to produce responses based on specific approaches, e.g. intersectional or queer
 3. Moral and ethical considerations when using the technology
 4. Pedagogical implications if resources on genAI are not discussed or made accessible
2. Who should have their say on the development of LLMs?
 1. Everyone that is affected by the developments of generative AI 'tools'



FUTURE
RESEARCH
& GET IN
TOUCH!

- Hope to build a network for understanding challenges within machine-assisted digital editing and pedagogy, please get in touch if you are interested and would like a chat: sophie.whittle@sheffield.ac.uk
- Publication of edition & source book, further funding for iterative testing processes of the Chaucer edition
- Interaction between TEI and AI?: representing responsibility and highlighting contributions from multiple, collaborating sources
- Returning to smaller language models? (I would recommend a recent IHR digital history seminar by Ashley Sanders, on [‘Identifying Latent Textual Bias: Making a case for traditional NLP tasks in the era of AI’](#))
- More reading?: Code Dependent: Living in the Shadow of AI by Madhumita Murgia and Moral Codes: Designing Alternatives to AI by Alan Blackwell



Chaucer Doth Tweet

@levostregc.bsky.social

Gen AI ys not the future. It ys a cash-grab attempt to freeze tyme and enshrine ideas and wayes of thinkinge that are alreadye stale and dated.

14 January 2025 at 17:36  Everybody can reply

20 reposts 141 likes



- Abbott, Jean, Elaine Treharne, and Mateusz Fafinski. 2021. *Beowulf by All: Translation and Workbook*. Arc Humanities Press.
- MetaLAB (at) Harvard. 2025. "AI Pedagogy Project: Creative and critical engagement with AI in education." Accessed January 15, 2025. <https://aipedagogy.org/>.
- Baechle, Sarah, and Carissa M. Harris. 2021. "The Ethical Challenge of Chaucerian Scholarship in the Twenty-First Century." *The Chaucer Review* 56 (4): 311–21. <https://muse.jhu.edu/pub/2/article/804395>.
- Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🐦." *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23. <https://doi.org/10.1145/3442188.3445922>.
- Blackwell, Alan F. 2024. *Moral Codes: Designing Alternatives to AI*. The MIT Press. <https://doi.org/10.7551/mitpress/14872.001.0001>.
- Boyd, Jason, and Bo Ruberg. 2022. "Queer Digital Humanities." In *The Bloomsbury Handbook to the Digital Humanities*, edited by James O'Sullivan. Bloomsbury Publishing.
- Buolamwini, Joy. 2017. "How I'm fighting bias in algorithms." TED Talk. Video, 8 min., 44 sec. https://www.youtube.com/watch?v=UG_X_7g63rY.
- Conrad, Kathryn. 2024. "A Blueprint for an AI Bill of Rights for Education." *Critical AI* 2 (1) <https://doi.org/10.1215/2834703X-11205245>.
- Critical AI @ Rutgers and Design Justice Labs. 2024. "Teaching Critical AI Literacies: 'Explainer' and Resources for the New Semester." Last modified October 20, 2024. <https://docs.google.com/document/d/1TAXqYGid8sQz8vIngTLDIqZBx2rNKHeKn9mcfWbFzRQ/edit?tab=t.0#heading=h.kgds7i8l6uca>.
- Design Justice Labs. 2024. "Critical AI Literacies: A Guide for Students." Accessed January 15, 2025. <https://rutgers.app.box.com/s/ek4mxim2soxho6cb2m2x5ztv92vfqlr8>.
- Design Justice Network. 2021. "Design Justice Principles in Philanthropy Workshop: April 14th Report Back." Last modified April 14, 2021. <https://designjustice.org/news-1/2021/philanthropy-workshop>.
- Evans, Ruth. 2022. "On Not Being Chaucer." *Studies in the Age of Chaucer* 44 (1): 2–26. <https://doi.org/10.1353/sac.2022.0028>.
- Gairola, Rahul K. 2022. "Race, Otherness, and the Digital Humanities." *The Bloomsbury Handbook to the Digital Humanities*, edited by James O'Sullivan. Bloomsbury Publishing.
- Goodlad, Lauren M. E. 2023. "Editor's Introduction: Humanities in the Loop." *Critical AI* 1 (1-2) <https://doi.org/10.1215/2834703X-10734016>.
- Greenstreet, Colin. 2024. "AI assistants and agents: A New Skill Set for Historians?" Moderated by James Baker. Virtual seminar, November 19. Posted November 20, 2024, by Institute for Historical Research: Digital History Seminar. <https://ihrdighist.blogs.sas.ac.uk/2024/09/tuesday-19-november-2024-colin-greenstreet-ai-agents-for-historical-research-a-new-skill-set-for-historians/>.
- Murgia, Madhumita. 2024. *Code Dependent: Living in the Shadow of AI*. Picador.
- O'Sullivan, James, Michael Pidd, Michael Kurzmeier, Orla Murphy, and Bridgette Wessels. 2023. "Interviews about the future of scholarly digital editions." Data files. <https://www.dhi.ac.uk/data/c21editions>.
- Programming Historian. 2025. "The Programming Historian." Accessed January 15, 2025. <https://programminghistorian.org/en/>.
- READ-COOP. 2023. "A cooperative to unlock our written past." Accessed January 15, 2025. <https://readcoop.org>.
- Sanders, Ashley R. 2024. "Identifying Latent Textual Bias: Making a case for traditional NLP tasks in the era of AI." Moderated by Alexandra Ortolja-Baird and Alice Kinghorn. Virtual seminar, October 29. Posted October 30, 2024, by Institute for Historical Research: Digital History Seminar. <https://ihrdighist.blogs.sas.ac.uk/2024/09/tuesday-29-october-2024-ashley-r-sanders-ucla-identifying-latent-textual-bias-making-a-case-for-traditional-nlp-tasks-in-the-era-of-ai/>.
- Schneider, Britta. 2024. "A Sociolinguist's Look at the 'Language' in Large Language Models." *Critical AI* 2 (1) <https://doi.org/10.1215/2834703X-11205168>.
- Stone, Matthew, Lauren M. E. Goodlad, and Mark Sammons. 2024. "The Origins of Generative AI in Transcription and Machine Translation, and Why That Matters." *Critical AI* 2 (1) <https://doi.org/10.1215/2834703X-11256853>.