



Accessus:

A Journal of Premodern Literature and New Media

Accessus

Volume 5
Issue 2 *Gower Shorts*

Article 8

2019

Gower as Data: Exploring the Application of Machine Learning to Gower's Middle English Corpus

Kara L. McShane

Ursinus College, kmcshane@ursinus.edu

Alvin Grissom II

Ursinus College, agrissom@ursinus.edu

Follow this and additional works at: <https://scholarworks.wmich.edu/accessus>



Part of the English Language and Literature Commons, European Languages and Societies Commons, Film and Media Studies Commons, and the Medieval Studies Commons

Recommended Citation

McShane, Kara L. and Grissom II, Alvin (2019) "Gower as Data: Exploring the Application of Machine Learning to Gower's Middle English Corpus," *Accessus*: Vol. 5 : Iss. 2 , Article 8.

Available at: <https://scholarworks.wmich.edu/accessus/vol5/iss2/8>

This Article has been accepted for inclusion in Accessus by the editorial board of the journal and administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



Gower as Data: Exploring the Application of Machine Learning to Gower's Middle English Corpus

This paper tracks two investigations in applying machine learning to Middle English literature with a particular focus on the English works of John Gower. Using machine learning techniques from computational linguistics, we open new doors of inquiry that potentially break with accepted assumptions. We aim to start a conversation that addresses the question of what, if anything, the value of machine learning for partial, disruptive, speculative scholarship in medieval literature is. It is within such a framework that we undertake this experiment, one that tries to use the small (and for many scholars, familiar) corpus to our advantage. Machine learning has been used in literary study primarily for large-scale projects that handle an amount of text that a single reader might not be able to consider; further, these texts are to some extent regularized, with consistency across versions (as in the printed editions of, for example, eighteenth-century novels). Middle English literature, on the other hand, is a comparatively small corpus, here made smaller as we focus on a single author, and it is marked by variation across manuscripts. While we acknowledge the problem of scale, here we build upon similar small projects to explore the possibilities of machine reading Middle English literature. We suggest that machine learning methods can be productive for scholars of medieval literature precisely due to the idiosyncrasies of their application: as a mechanism for hypothesis generation and a space of exploration. Machine learning offers a valuable mechanism for play and for distancing a critic from a familiar text, and thus these tools can suggest new links and new readings of old works.

Gower's corpus is particularly suited to such study. To begin, his English writings, by the standards of the field, are substantial yet fairly uniform, consisting primarily of a single long

poem written in regular meter. Moreover, the critical tradition in Gower studies has continued to have strong influence on contemporary readings of his work.¹ For Gowerians, this method helps us reexamine established ways of reading Gower's work. Through new associations, suggestions, and disruptive possibilities, we as critics may gain fresh perspectives and develop new potential readings. Beyond its usefulness for scholars of Gower's English works, however, these two case studies offer an example of distant reading with a smaller data set than those typically modeled with such methods, perhaps opening possibilities for digital work on other similarly small corpora. If digital humanities (DH) as a field is largely about how computing gives us new ways to answer old questions and helps us ask new questions, this project attempts to bridge those two outcomes. In doing so, we hope to disturb expectations in both digital humanities and Gower criticism and invite practitioners in both fields to ask further questions about the scale of digital studies, its usefulness to texts existing primarily in manuscript, and what these methods might suggest about the formal and thematic features accepted as central to Gower's English work.

Machine reading methods present challenges for practitioners working with Middle English texts, and medievalists have detailed some of these challenges. Manuscripts in particular, and Middle English-language texts more generally, resist machine reading due to particulars of their construction. Orthography is the largest concern: Middle English spelling is not consistent, and each variant of a word is read by a machine as its own distinct word. Beyond this, the accessibility of texts for analysis is a problem. Though medievalists and computer scientists have begun collaborating to make manuscripts machine readable, the technology is by no means

¹ For a succinct yet thorough overview of this critical history, see Siân Echard, "Introduction: Gower's Reputation" in *A Companion to Gower*, ed. Siân Echard (Cambridge, UK: D. S. Brewer, 2004), 1-22.

perfect.² Beginning with edited texts is not necessarily a perfect solution, since, as Michael Widner observes, many texts available online, including those of the TEAMS Middle English Texts Series that we use here, are not encoded in a way that makes them easily available for download and manipulation.³ Even when one circumvents these challenges, using data sets this small can raise problems, as Scott Kleinman observes: “It’s not that machine learning tools don’t work on them; it’s just that the insights you get provide little added value When machine learning techniques do detect patterns that surprise us, we are forced to doubt the statistical significance of these patterns because of the small size of the data.”⁴ Our question in part, then, builds from Kleinman’s observation: does machine analysis have to tell us something statistically significant to be useful in some way? Jennifer Boyle and Helen Burgess consider that manuscripts and medieval poetry do not necessarily lend themselves easily to DH work; they offer that this resistance in the materials that medievalists study creates “an opportunity for weird, partial, and disruptive scholarship,” and Kleinman further suggests “we may need to

² See Brandon Hawk, Antonia Karaisl, and Nick White, “Modelling Medieval Hands: Practical OCR for Caroline Minuscule,” *Digital Humanities Quarterly* 13.1 (2019) <http://www.digitalhumanities.org/dhq/vol/13/1/000412/000412.html>, n.p.a. Work on machine reading Middle English is ongoing, as recent presentations by David Mimno, Laure Thompson, and Anna Fore Waymack of Cornell University and Tobias Hodel and Maria Kallio, both of the READ (Recognition and Enrichment of Archival Documents, <https://read.transkribus.eu/>) project, demonstrate. These talks, given at ICMS Kalamazoo in 2019, address issues of using MALLETT for topic modeling (for the team at Cornell) and using READ to search medieval documents (Hodel & Kallio). A package for standardizing Middle English within Python is available through the Classical Language Toolkit (Kyle P. Johnson, et al. (2014-2019). CLTK: The Classical Language Toolkit. DOI 10.5281/zenodo.593336, available at <https://github.com/cltk>)

³ Widner, “Toward Text-Mining the Middle Ages: Digital Scriptoria and Networks of Labor,” in *The Routledge Research Companion to Digital Medieval Literature*, ed. Jennifer E. Boyle and Helen J. Burgess (New York: Routledge, 2018), 131-44. Long-term, this will change: as Pamela M. Yee, assistant editor of METS, indicated in her recent (2019) presentation at ICMS Kalamazoo, METS staff have begun what will be a slow and painstaking process of encoding texts in the series using TEI. For this project, we constructed a program that transfers the text on the website into a .txt file.

⁴ Scott Kleinman, “Digital Humanities Projects with Small and Unusual Data: Some Experiences from the Trenches,” <http://scottkleinman.net/blog/2016/03/15/digital-humanities-projects-with-small-and-unusual-data/>. Posted 15 March 2016, n.p.a.

cultivate (and reward) a sort of playful hermeneutics in working with this data.”⁵ Our interventions, then, seek to engage in precisely this exploratory, creative interpretation.

Phase 1: Word Embeddings

Our first small project explores what word embeddings⁶ might suggest about John Gower’s use of emotion, focusing on an expressive trait often seen at odds with his plain style.⁷ Compared specifically to his contemporary Chaucer, Gower is known for what R. F. Yeager and Brian Gastle call “his economical depiction of emotion.”⁸ Rather than emotional, Gower’s sphere has long been understood as political and moral; it is Gower, not Chaucer, who engages most explicitly in the political issues of his day.⁹ In examining the poet’s use of emotive terms, then, we hope to find surprises: patterns, associations, and links between his depiction of emotion and his moral and political concerns. While Gower’s straightforward poetic lines have often been considered poetically inferior by literary scholars,¹⁰ simple lines take a great deal of care to craft; as Maura Nolan has recently suggested, they may be particularly well-suited to emotive expression because their parsimony draws attention to specific moments of emotion.¹¹

⁵ Jennifer E. Boyle and Helen J. Burgess, “Introduction: Resistance in the Materials,” in *The Routledge Research Companion to Digital Medieval Literature*, ed. Jennifer E. Boyle and Helen J. Burgess (New York: Routledge, 2018), 4; Kleinman, “Digital Humanities Projects,” n.p.a.

⁶ Thomas Mikolov, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean, “Distributed representations of words and phrases and their compositionality.” *Advances in Neural Information Processing Systems* 26 (2013), 3111-19.

⁷ For more on Gower’s style and its appropriateness to conveying emotion, see especially Maura Nolan, ““Sensation and the Plain Style in Gower’s *Confessio Amantis*,” in *John Gower: Others and the Self*, ed. Russell A. Peck and R. F. Yeager (Cambridge, UK: D. S. Brewer, 2017), 111-40.

⁸ See “Introduction: Gower in Context,” in *Approaches to Teaching the Poetry of John Gower*, ed. R. F. Yeager and Brian W. Gastle (New York: Modern Language Association, 2011), 1.

⁹ Yeager and Gastle, “Gower in Context,” 2.

¹⁰ Even praise of Gower’s poetry, such as that by early editor G. C. Macaulay, is often tepid: while Macaulay praises Gower’s writing, he claims that Chaucer was “an artist of an altogether superior kind” (qtd. in Echard, *A Companion to Gower*, 13).

¹¹ Nolan, “Sensation and the Plain Style,” 113.

By using word embeddings, we hope to see what kinds of terms are associated with Gower's "emotive" words. If he is economical, then his discussions of emotion become all the more important. Our hope in using word embeddings is that the technology might serve to create distance, to discourage us from seeking known patterns or reinforcing critically accepted readings.

Word embeddings have seen widespread usage in recent computational linguistics and natural language processing (NLP) research. Given some text—in our case, a single work by Gower—an algorithm learns a semantic representation of words. Unlike in traditional logical semantics, the meaning of each word is defined in terms of its relationship to other words. By themselves, they have been used for crafting analogies (for example, man/woman is to king/queen)¹² for automatically finding synonyms, for diachronically studying word change over time,¹³ for studying bias in text, and for many other tasks. In addition, word embeddings are often used in applications when working with text data, as this representation of the data is often more useful for AI than raw text.

Word embeddings are based on the principle that "You shall know a word by the company it keeps,"¹⁴ a principle that guides many other NLP approaches to characterizing text, such as topic modeling. While the details are complex, the idea behind word embeddings is

¹² Thomas Mikolov, Wen-tau Yih, and Geoffrey Zweig, "Linguistic regularities in continuous space word representations," *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (2013), 746-51. <https://www.aclweb.org/anthology/N13-1090>

¹³ William L. Hamilton, Jure Leskovec, and Dan Jurafsky, "Diachronic word embeddings reveal statistical laws of semantic change," *Proceedings of the 2016 Conference of the Association for Computational Linguistics* (2016), 1489-1501. <https://www.aclweb.org/anthology/P16-1141>

¹⁴ John R. Firth. "A synopsis of linguistic theory, 1930-1955," *Studies in Linguistic Analysis* (Oxford: Blackwell, 1957), 1-32.

simple: given a corpus, the algorithm looks at all short spans within a “sliding window.” Inside this window, we focus on one word at a time and try to predict the surrounding words. The algorithm goes through the text several times, and each time, it gets better at predicting the surrounding words. Based on the assumption that similar words tend to occur together, the algorithm develops a semantic model wherein each word is closer to or farther away from every other word along many different dimensions. This representation of the words can be thought of geometrically: words that are closer along several dimensions are closer to each other in general and vice versa in this vector space, but words may be close to each other in some dimensions and far away in others. We then have a model that clusters similar words together in this space. So, for example, in a standard, general purpose corpus, such as a newspaper or data collected from the web, we would expect that the word “love” would be closer to the word “affection” than the word “blasphemy.” Once the model is built—a process that is nearly instantaneous on a modern computer with a corpus of our size—we can query the model for single words or sets of words to see which words the model has clustered together.

Our text is not a general-purpose corpus like one generated from web or newspaper data; rather, it is a corpus by a single author, and not general prose but poetry. We therefore expect that word embeddings will encode the idiosyncrasies of this author’s language use. While we find some synonyms or related terms in our queries, we also find some elements that are distinctive in Gower’s writing. Our corpus, which uses the digital texts available through the METS website, consists of some 33,000 lines of the *Confessio Amantis* in Russell A. Peck’s 2004-2013 editions, with the Latin lines removed, as well as Gower’s 385-line English poem *In Praise of Peace* as edited by Michael Livingston.¹⁵

¹⁵ Texts are available at <https://d.lib.rochester.edu/teams/publication/peck-confessio-amantis-volume-1> and <https://d.lib.rochester.edu/teams/text/yeager-gower-minor-latin-works-livingston-in-praise-of-peace>. The METS

From the perspective of the person querying the model, the process works in the following way: a researcher enters a term to query, and the algorithm subsequently returns a list of the most similar words to the queried term in the corpus, with a similarity score between 0 and 1 for each word.¹⁶ Since each term can be similar across a number of dimensions, it is also possible to visualize them and examine representations of the terms graphically. While the nature of the visualizations depends on a number of tweakable factors and should not be considered “objective” in any meaningful sense, such visualizations afford a natural way of exploring themes and patterns, using the data as a guide. As a process, exploration of Gower’s English corpus through this program is iterative; that is, one result yields new possible associations and new terms to query.

Determining which terms to query, of course, requires human input.¹⁷ It is impossible---and, we would suggest, not desirable---to entirely circumvent the human element of such study. Yet the program’s ability to make connections between terms serves as a kind of hypothesis-generating device, another tool for literary scholars to use in exploring links between particular words in Gower’s English corpus.¹⁸ The data suggest that there might be a connection: the place

website currently reflects the text of Peck’s 2006 edition of Volume 1 (the Prologue, Book 1, and Book 8); his 2013 edition of Volume 2 (Books 2, 3, and 4), and his 2004 edition of Volume 3 (Books 5, 6, and 7). We created text files of these works using Python, through a process known as “scraping.” More information on web scraping is available through Jeri Wieringa’s lesson on the Programming Historian, which introduces learners to a program, BeautifulSoup, designed for web scraping; see <https://programminghistorian.org/en/lessons/intro-to-beautiful-soup>.

¹⁶ The program is also available for download alongside this article. We do not include the numerical scores in our tables here, but list words from most to least similar to the queried word.

¹⁷ As David Birnbaum, Sheila Bonde, and Mike Kestemont’s case study of topic modeling in back issues of *Speculum* reminds us, models “reproduce the cultural biases that are present in the material on which they have been based” (S31). See Birnbaum, Bonde, and Kestemont, “The Digital Middle Ages: An Introduction,” *Speculum* 92:1 (Oct 2017), <https://www.journals.uchicago.edu/doi/full/10.1086/694236?mobileUi=0&..>, S1-S38.

¹⁸ Consider, for example, arguments made about distant reading in Franco Moretti’s *Graphs, Maps, Trees: Abstract Models for Literary History* (London: Verso, 2007). In his analysis, Moretti claims that quantification leads to problems rather than providing answers (26).

of close reading, then, is to pick up this connection and consider its nature more fully. We began our study by querying terms for emotion that we expected to find in Gower's work, such as *love*, *lust*, *rage*, *wraththe*, *wanhope*. Thus, rather than conclusions, we here offer observations and, with them, questions for further consideration.

In our initial queries of emotion terms, we noted the prevalence of location-related words, place terms. *Contre*'s association with rage, considered alongside *worldes*' association with vice, could suggest connections between emotion and location. A glance at the two terms together yields interesting similarities: *partie* and *noble* appear in the search for *rage*, while *worldes* and *pes* appear in the search for *vice*. *Partie* may suggest something about Gower's preoccupation with division---its most common usage is to indicate a part, piece, or portion of something---but it might also suggest a group of people or a social class.¹⁹ Does *partie* suggest something about the effects of rage? Or does Gower tend to use it instead to refer to collectives of people?

Table 1. Closest words to *rage*, *vice*, *love*, *lust*, and *bok*

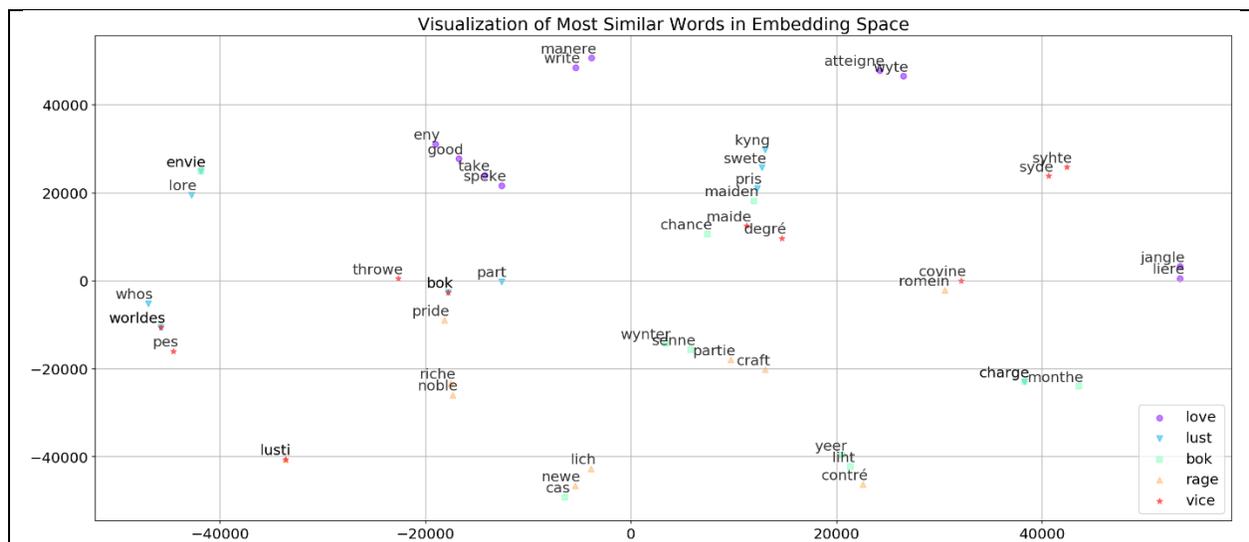
Rage	Vice	Lust	Bok	Love
newe noble lusti partie contré craft riche pride lich romein	lusti maide thief partie knyht throwe degré noble worldes Plit	swete part lore envie kyng whos pris worldes bok charge	chance cas liht yeer senne envie charge maiden monthe wynter	manere wyte good eny jangle liere speke write atteigne take

¹⁹MED partie n.

In Table 1, we can see the results of querying “rage” and “vice,” respectively, and Figure 1 shows visualizations of these queried terms using a technique called t-SNE, which can take the words in a high-dimensional space and plot them on a two-dimensional plane that is to some degree interpretable by humans. Each query yields some terms that one would expect—for example, *love* is seen as a social *good*, while *lust* returns the perhaps obvious *envie*. This makes the perhaps unexpected or surprising associations all the more potentially productive. For example, many emotion terms also seem associated with terms for forms of writing—*love* yields *write* (as well as *speke*), while *lust* returns *lore* and *bok*.²⁰ Of the words associated with *lust*, *bok* is the one more directly linked to a physical written artifact in Middle English; yet *lore* has an educational connotation, linked to teaching as well as to narratives and storytelling. The multivalent nature of the term makes it particularly apt to the structure of the *Confessio* in which storytelling is itself instruction. The association between *lust* and *bok* may reinforce those connections. Yet querying *bok* also yields associations about time—*yeer* and *monthe*, as well as, arguably, *wynter*. This may have implications for our understanding of Gower’s investment in writing as a form of action in the world;²¹ how is emotion linked to Gower’s understanding of writerly craft, of the poet’s social and political role? How is time linked to writerly craft for Gower?

²⁰ While the *Middle English Dictionary*, s.v. *lore*, n (1) refers to losses, *lore*, n(2) refers to teaching and instruction or its results and can refer to a narrative or story. Further analysis would be required to determine how frequently Gower uses the term in which sense, but this second set of meanings is potentially fruitful and makes sense in the context of the *Confessio Amantis* as a teaching text.

²¹ On this, see Brian Gastle, “‘The Lucre of Marchandie’: Poet, Patron, and Payment in Gower’s *Confessio Amantis*,” in *John Gower: Others and the Self*, ed. Russell A. Peck and R. F. Yeager (Cambridge, UK: D. S. Brewer, 2017), 283-94 and Gabrielle Parkin, “Hidden Matter in John Gower’s *Confessio Amantis*,” in *John Gower: Others and the Self*, ed. Russell A. Peck and R. F. Yeager (Cambridge, UK: D. S. Brewer, 2017), 295-305. Foundational work on the topic is by Russell A. Peck, *Kingship and Common Profit in Gower’s Confessio Amantis* (Carbondale: Southern Illinois University Press, 1978) and R. F. Yeager, *John Gower’s Poetic: The Search for a New Arion* (Cambridge, UK: D. S. Brewer, 1990).

Figure 1: 2D Visualization (t-SNE) of the ten most similar words to *rage*, *vice*, *love*, *lust*, and *bok*.

Based on the results that the algorithm generated for emotion terms, we then queried the model again, seeking associations with terms related to governance and social order, the sort of terms one might expect to be central to Gower's concerns---words such as *parlement*, *governance*, *kyng*, *lawe*. *Lawe* returns both *vice* and *pride*---fascinating links for someone who was well versed in law, as Conrad Van Dijk, Candace Barrington, and others have demonstrated.²² How does Gower imagine law as being linked to vice? To pride? Along similar lines, *lawe* also returns *knyht* as a closely associated term. Is this link a matter of social and political standing? Do Gower's knights often find themselves in conflict with the law? The program does not suggest how terms are associated: it identifies them simply as semantically proximate. Close reading, then, may pick up where these connections leave off.

Phase 2: Generating Middle English

²² See Conrad Van Dijk, *John Gower and the Limits of the Law* (Cambridge, UK: D. S. Brewer, 2013) for an innovative reading and useful bibliography.

Another direction we've taken this work is a slightly more creative one: we can train a program to produce "Gowerian-inspired" lines. While we do not expect the program to preserve meter, the goal is for it to produce grammatical if not necessarily sensical Middle English. In computational linguistics, this is known as language modeling; the particular model we use is called a Long Short-term Memory neural network, or LSTM. While word embeddings give us associations based on word proximity, this model is different. LSTMs can be used to create sequence-based *language models*. These models can be used to generate new sentences based on the prior context, by producing the most likely words to follow the previous ones, sometimes capturing turns of phrase and stylistic nuance. Such models implicitly create content connections and syntactic models, producing sentences that syntactically resemble those produced by the text on which the model is based. These models, like word embeddings, offer opportunities for unexpected connections to be produced; they both care about proximity, and so presumably, one should find some kind of consistency between the two methods.

When used for mobile devices, the longstanding assumption is that predictive text is for "efficiency"---that is, if a phone can successfully predict the next word in a sequence, it saves the user from having to type out whole words. The predictive ability of these programs, as many users know, is somewhat variable. The program we have applied to Gower's work uses the same fundamental algorithm, but efficiency is not the goal; rather, we hope that the program might suggest interesting connections and offer perspectives on Gower's frequently used phrases and expressions. So, for example, if one provides the program with "middel," "weie" would be a likely next word for the program to choose, given Gower's extensive discussion of the "middel weie" his poetry seeks to walk. As with word embeddings, the point is partly to confirm

expectations set by the critical tradition, but also partly to shatter those expectations, to generate surprising, evocative suggestions.

Table 2: Two examples of Gower-like generated Middle English, with initial seed phrases in boldface.

Trial 1 Output	Trial 2 Output
<p>Of hem that writen ous tofore The bokes duelle and we therefore Ben tawht of that was write tho of myht wordes othre me calle syh only upon the werk aboute and so thei made an herte bere and duelle ferst and make him hide ther is toward a gret delit hou thei creatures the word of hem that crist schal liere telle hou ferst ther most be sore do that he mai give with the sonne so be the goode senne fyred for whan thei setten into this that thei most fulle now th ende than forthi sche was nocht cry he seith</p>	<p>Bot wolde God that now were on An other such as Arion which ridende hadde was here wif under his entente: of that his fader him were sone stod the juise nevere for i which may this bet wolden for it is every man to take bot ate laste ther as he bot over this frensche were the king ne time on rome so the king cam til he noble hed a tale which the goddes yhe that he hath on in such a wise that he for al him on him slowh</p>

Our early results give us phrases we'd expect, as well as some loosely sensical lines. In Trial 1, phrases such as “a gret delit” does appear in Gower’s corpus (in the Tale of Nectanabus), as does “schal liere”; “in such a wise,” which appears in Trial 2, is a frequently used phrase in Gower’s work, pointing to a formal feature of his poetry.²³ Equally interesting is what the program guesses will appear after frequently appearing words, such as definite or indefinite articles, “the” or “a.” These terms reinforce common concerns of Gower’s poetry, as here;

²³ “a gret delit” appears at line 1836 in Book VI; “schal liere” appears only in the Tale of Apollonius (8.1462). For “in such a wise,” see for example Prol.589; Prol.759; 1.266; 1.1400; 8.173; 8.2128.

“king” appears after “the,” which provides some indication both of Gower’s concern about kingship as a concept and of the characters who populate many of his tales. Similarly, “a tale” generated in Trial 2 is an incredibly frequent formulation as Genius narrates tale after tale, drawing attention to the poem’s structural features. While the trials yield largely expected associations rather than especially ground-breaking ones, they function as a valuable check on this method; that is, they demonstrate that the LSTM model can successfully capture some high-level elements of Gower’s Middle English.

Conclusions

While the explorations discussed here are preliminary, we suggest that they do support the viability of both methods for future medieval literary scholarship. Relying on algorithmic techniques entirely for analysis is fraught with some of the same dangers inherent to appealing to any authority. But what we suggest here, rather, is a methodology for creative scholarship by which the computational model engages in a dialogue with the scholar. The model provides feedback that can lead the scholar to examine questions that may not have been asked. Our work serves as a preliminary example of how computationally assisted analysis can serve as a kind of digital partner for medieval scholars, without succumbing to the temptations of treating the results of digital analysis as more “objective.” This machine, in this case, ignites new, creative directions for scholars by offering an alternative, computational methodology to generate material from which to draw.

Bibliography

Boyle, Jennifer E., and Helen J. Burgess, “Introduction: Resistance in the Materials.” In *The*

- Routledge Research Companion to Digital Medieval Literature*. Edited by Jennifer E. Boyle and Helen J. Burgess. New York: Routledge, 2018. 1-5.
- Birnbaum, David, Sheila Bonde, and Mike Kestemont. "The Digital Middle Ages: An Introduction," *Speculum* 92:1 (Oct 2017), <https://www.journals.uchicago.edu/doi/full/10.1086/694236?mobileUi=0&>: S1-S38.
- Echard, Siân. "Introduction: Gower's Reputation." In *A Companion to Gower*. Edited by Siân Echard. Cambridge, UK: D. S. Brewer, 2004. 1-22.
- Firth, John R. "A synopsis of linguistic theory, 1930-1955," *Studies in Linguistic Analysis*. Oxford: Blackwell, 1957. 1-32.
- Gastle, Brian W. "'The Lucre of Marchandie': Poet, Patron, and Payment in Gower's *Confessio Amantis*." In *John Gower: Others and the Self*. Edited by Russell A. Peck and R. F. Yeager. Cambridge, UK: D. S. Brewer, 2017. 283-94.
- Hamilton, William L., Jure Leskovec, and Dan Jurafsky, "Diachronic word embeddings reveal statistical laws of semantic change." *Proceedings of the 2016 Conference of the Association for Computational Linguistics* (2016):1489-1501.
- Hawk, Brandon, Antonia Karaisl, and Nick White. "Modelling Medieval Hands: Practical OCR for Caroline Minuscule." *Digital Humanities Quarterly* 13.1 (2019). <http://www.digitalhumanities.org/dhq/vol/13/1/000412/000412.html> , n.p.a.
- Johnson, Kyle P., et al. *CLTK: The Classical Language Toolkit*. 2014-2019. DOI 10.5281/zenodo.593336. <https://github.com/cltk> , n.p.a
- Kleinman, Scott. "Digital Humanities Projects with Small and Unusual Data: Some Experiences from the Trenches." <http://scottkleinman.net/blog/2016/03/15/digital-humanities-projects-with-small-and-unusual-data/>. Posted 15 March 2016.

Livingston, Michael, ed. "John Gower: In Praise of Peace." In *John Gower: The Minor Latin Works*. Edited by R. F. Yeager. <https://d.lib.rochester.edu/teams/text/yeager-gower-minor-latin-works-livingston-in-praise-of-peace>.

Middle English Dictionary. Edited by Robert E. Lewis, et al. Ann Arbor: University of Michigan Press, 1952-2001. Online edition in Middle English Compendium. Edited by Frances McSparran, et al.. Ann Arbor: University of Michigan Library, 2000-2018. <http://quod.lib.umich.edu/m/middle-english-dictionary/>

Mikolov, Thomas, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean, "Distributed representations of words and phrases and their compositionality." *Advances in Neural Information Processing Systems* 26 (2013). 3111-19.

Mikolov, Thomas, Wen-tau Yih, and Geoffrey Zweig. "Linguistic regularities in continuous space word representations." *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (2013): 746-51.

Moretti, Franco. *Graphs, Maps, Trees: Abstract Models for Literary History*. London: Verso, 2007.

Nolan, Maura. "'Sensation and the Plain Style in Gower's *Confessio Amantis*.'" In *John Gower: Others and the Self*. Edited by Russell A. Peck and R. F. Yeager. Cambridge, UK: D. S. Brewer, 2017. 111-40.

Parkin, Gabrielle. "Hidden Matter in John Gower's *Confessio Amantis*." In *John Gower: Others and the Self*. Edited by Russell A. Peck and R. F. Yeager. Cambridge, UK: D. S. Brewer, 2017. 295-305.

Peck, Russell A., ed. *John Gower: Confessio Amantis*. Vol. 1. Second edition. Kalamazoo, MI:

- Medieval Institute Publications, 2006. <https://d.lib.rochester.edu/teams/publication/peck-confessio-amantis-volume-1> n.p.a.
- Peck, Russell A., ed. *John Gower: Confessio Amantis*. Vol. 2. Second edition. Kalamazoo, MI: Medieval Institute Publications, 2013. <https://d.lib.rochester.edu/teams/publication/peck-confessio-amantis-volume-2> n.p.a.
- Peck, Russell A., ed. *John Gower: Confessio Amantis*. Vol. 3. Kalamazoo, MI: Medieval Institute Publications, 2004. <https://d.lib.rochester.edu/teams/publication/peck-confessio-amantis-volume-3> n.p.a.
- Peck, Russell A. *Kingship and Common Profit in Gower's Confessio Amantis*. Carbondale: Southern Illinois University Press, 1978.
- Recognition and Enrichment of Archival Documents* (READ). <https://read.transkribus.eu/>. Modified 2018. n.p.a.
- Van Dijk, Conrad. *John Gower and the Limits of the Law*. Cambridge, UK: D. S. Brewer, 2013.
- Widner, Michael. "Toward Text-Mining the Middle Ages: Digital Scriptoria and Networks of Labor." In *The Routledge Research Companion to Digital Medieval Literature*. Edited by Jennifer E. Boyle and Helen J. Burgess. New York: Routledge, 2018. 131-44.
- Wieringa, Jeri. "Intro to BeautifulSoup." *The Programming Historian*. <https://programminghistorian.org/en/lessons/intro-to-beautiful-soup>. Modified 24 Sept 2018. n.p.a.
- Yeager, R. F. *John Gower's Poetic: The Search for a New Arion*. Cambridge, UK: D. S. Brewer, 1990.
- Yeager, R. F., and Brian W. Gastle. "Introduction: Gower in Context." In *Approaches to*

Teaching the Poetry of John Gower. Edited by R. F. Yeager and Brian W. Gastle. New York: Modern Language Association, 2011. 1-6.

Kara L. McShane
Ursinus College
kmcshane@ursinus.edu

Alvin Grissom II
Ursinus College
agrissom@ursinus.edu