Routledge Medieval Encyclopedia Online

Computational Stylistics and Medieval Texts

--Manuscript Draft--

Manuscript Number:	RMEO-D-22-00103R1
Full Title:	Computational Stylistics and Medieval Texts
Article Type:	Revised
Section/Category:	Samu Niskanen & James Willoughby - (WRITING HISTORY, METHODOLOGY)
Keywords:	medieval text; practitioners of stylometry; stylometry; computational stylistics; authorship; digital scholarship; digital humanities; principal components analysis; state of the art; digital text analysis; digital philology; authorship attribution
Corresponding Author:	Jeroen De Gussem Ghent University Ghent, BELGIUM
First Author:	Jeroen De Gussem
Order of Authors:	Jeroen De Gussem
Abstract:	This entry gives a survey of the application of computational stylistics to medieval texts. Computational stylistics, alternatively stylometry, is an umbrella term for a set of computer algorithms that quantify and statistically harvest a document's stylistic features in order to make statistically informed deductions concerning authorship, dating, influence, provenance and/or stylistic-literary characteristics. The entry firstly traces the history of the method from proto-stylometric ancestors such as the sixteenth-century humanist Lorenzo Valla to modern twentieth-century founding fathers Frederick Mosteller and David L. Wallace. Secondly, it traces the broad diversity of research aims that scholars of the Middle Ages have when making use of this method, e.g. authorship attribution, the impact of time on style (stylochronometry), the influence of copyists and scribes, assessing dependencies between manuscripts, tracing reliances on same source texts or school models, literary collaboration, scriptological and dialectological questions, translingual influence, the impact of genre, etc. Thirdly, the entry surveys some of the most common techniques applied to medieval texts (preprocessing, feature extraction and selection, vectorization and scaling, data analysis and visualization). In conclusion, an attempt at an adequate summary is made of the future(s) which practitioners of stylometry for medieval texts envision for this relatively young field.
Additional Information:	
Question	**Response**

Figure 1 Click here to access/download;Figure;figure-1.png

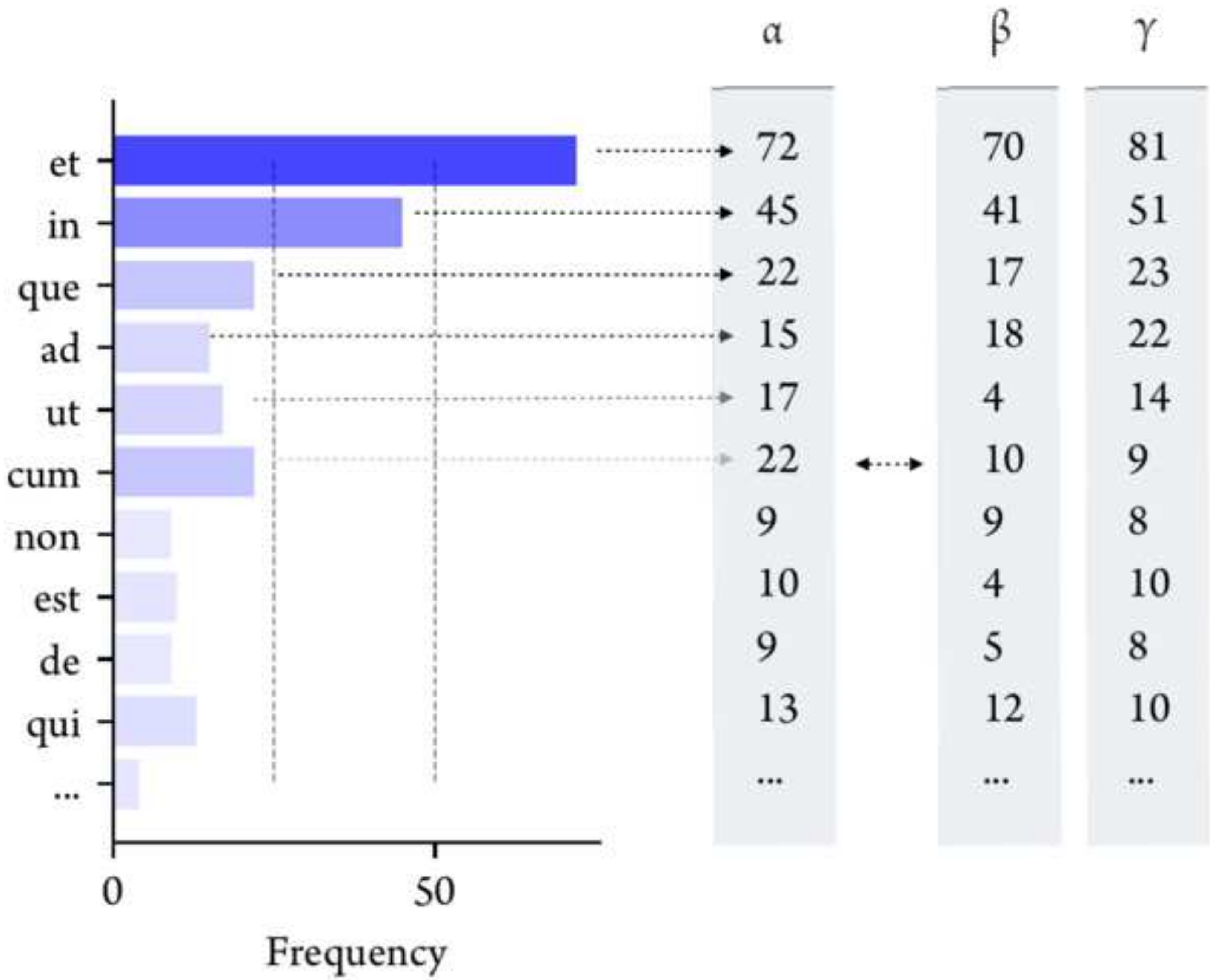


Figure 2

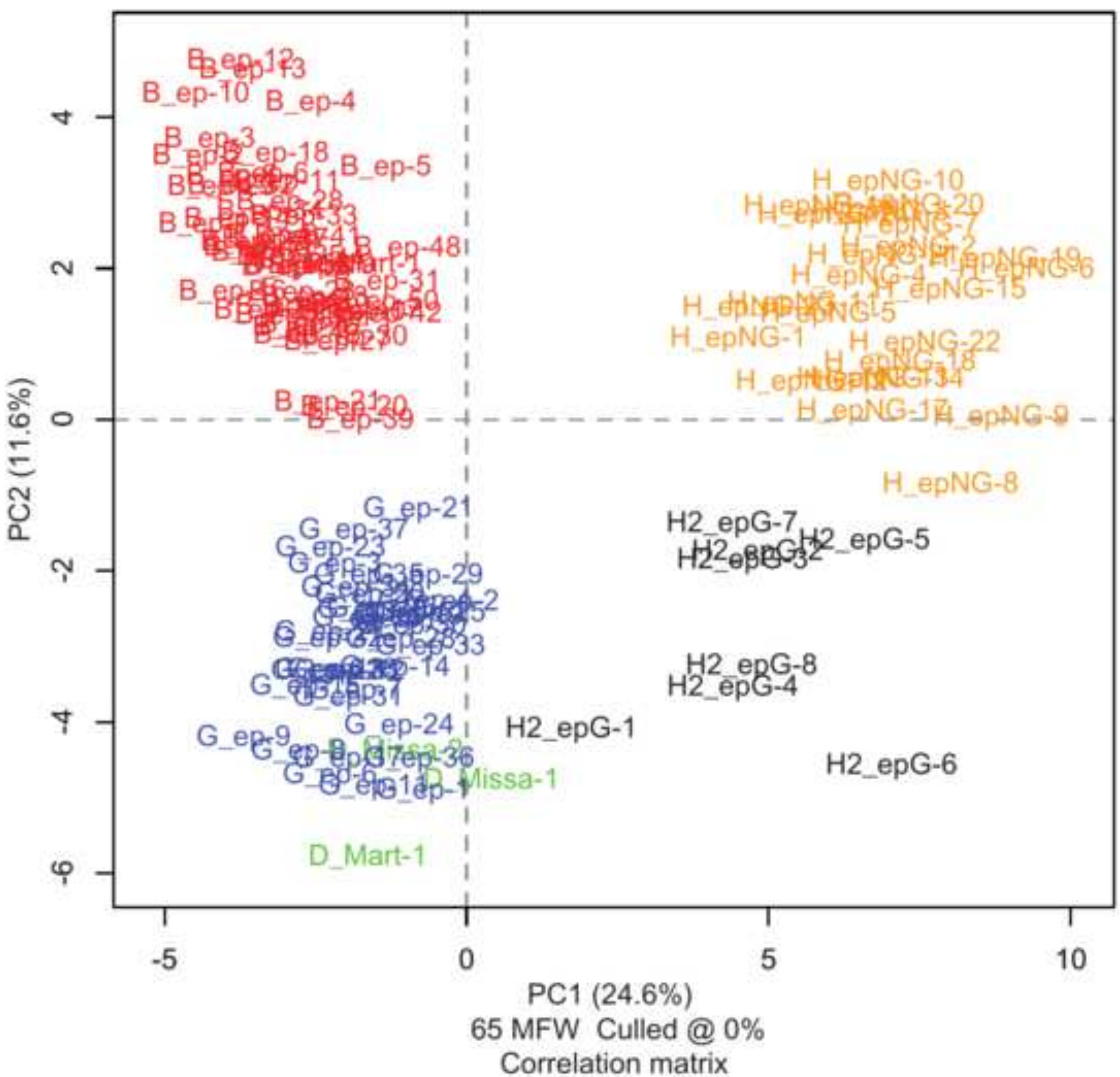
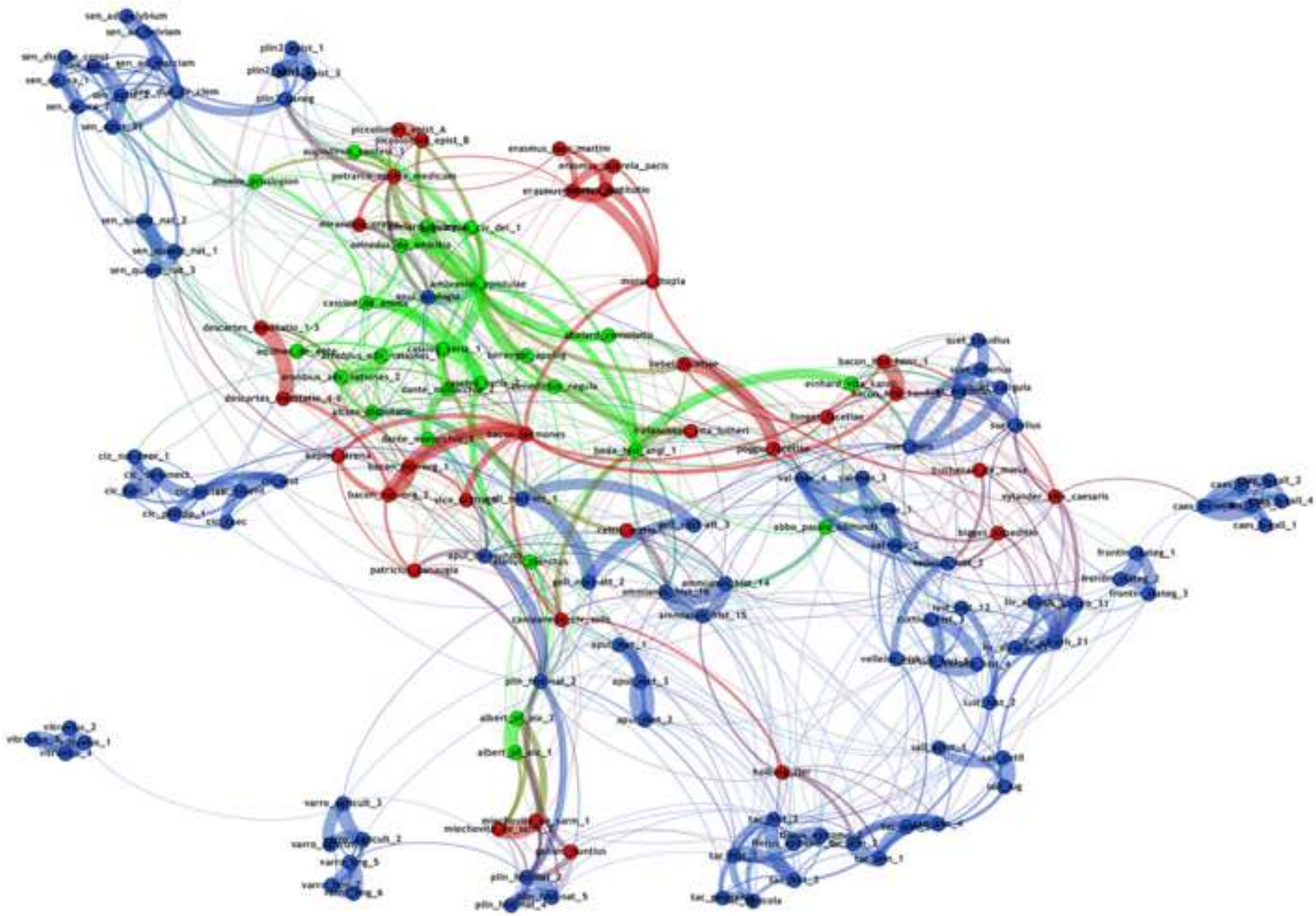

B_ep-12
B_ep-10
B_ep-4
B_ep-3
B_ep-18
B_ep-5
B_ep-48
B_ep-31
B_ep-42
B_ep-39
H_epNG-10
H_epNG-7
H_epNG-4
H_epNG-15
H_epNG-1
H_epNG-22
H_epNG-18
H_epNG-9
H_epNG-8
G_ep-21
G_ep-37
G_ep-23
G_ep-29
G_ep-33
G_ep-14
G_ep-31
G_ep-24
G_ep-9
G_ep-36
D_Missa-1
D_Mart-1
H2_epG-7
H2_epG-5
H2_epG-3
H2_epG-8
H2_epG-4
H2_epG-1
H2_epG-6
PC2 (11.6%)
4
2
0
-2
-4
-6
-5
0
5
10
PC1 (24.6%)
65 MFW Culled @ 0%
Correlation matrix

Figure 3



Computational Stylistics and Medieval Texts

JEROEN DE GUSSEM
Ghent University

1. Terminology and History

Computational stylistics, alternatively stylometry, is an umbrella term for a set of computer algorithms that quantify and statistically harvest a document's stylistic features in order to make statistically informed deductions concerning authorship, dating, influence, provenance and/or stylistic-literary characteristics. Stylometry is closely associated with computational linguistics and natural language processing (NLP), and it is popularly applied in the area of research which identifies as Digital Humanities (DH).

A coherent narrative of stylometry's history is lacking (Binongo and Smith 1996; Grzybek 2014). Practitioners tend to adhere to national traditions (Delcourt 2002:980) or focus on a single branch or conventional version. Some studies refer to 'proto-stylometric' approaches avant la lettre, the most popular of which is Lorenzo Valla's (1407–1457) famous unmasking on the basis of stylistic arguments of a controversial donation charter ascribed to Constantine as an early medieval Carolingian forgery (Eder 2016:61). Despite such ancestors, all historical surveys clearly indicate that the cradle of stylometry is the mid- to late-nineteenth century, when different scholars across the globe began making initial (simultaneous but not seldom independent) explorations into quantitative approaches to writing style. Thomas C. Mendenhall (1841–1924) and Wincenty Lutosławski (1863–1954) are key figures in this context. Mendenhall made the first publication on stylometry applied to Charles Dickens in *Science* (1887) and applied a "simple counting machine" (Mendenhall 1901:102), perhaps the first ever applied to literary questions (Bailey 1969:218), to investigate if Sir Francis Bacon (1561–1626) was responsible for authoring plays under signature of William Shakespeare (1564–1616). Attempting to establish the chronology of Plato's writings (1897 [1905]), the Polish philosopher and philologist Lutosławski appears to have coined 'stylometry' as a term, first using it for a paper in Oxford titled 'On Stylometry' (1897). Around that same time, Andrej A. Markov (1856–1922), George K. Zipf (1902–1950) and G. Udny Yule (1871–1951) were also making significant contributions to stylometric techniques.

Modern computational stylometry broke through with the 1964 foundational scholarly work of Frederick Mosteller (1916–2006) and David L. Wallace (1928–2017). In their analysis of the Federalist papers, the American scholars showed that function words, inconspicuous and very frequent syntactic words unsusceptible to stylistic imitation and forgery, were extremely efficient for distinguishing works of different authorship, and their book became the seminal scholarly work of non-traditional authorship attribution. One of Mosteller and Wallace's relatively recent and impactful successors to develop the function word method was the Australian scholar John F. Burrows (1928–2019), who introduced multivariate analysis of style with principal components analysis (PCA) and the distance metric 'Delta' (2002). There is insufficient space here to give credit to the works of current household names in the stylometric state of the art such as Hugh Craig (a close associate of Burrows in the Centre for Literary and Linguistic Computing at the University of Newcastle), David I. Holmes, David L. Hoover (1949–), Stanford Literary Lab founders Franco Moretti (1950–) and Matthew L. Jockers, Patrick Juola (1966–), Efstathios Stamatatos, Moshe Koppel (1956–), Shlomo Engelson Argamon (1967–), Adam Kilgariff (1960–2015), James Pennebaker (1950–) and many more, but I hope the reader will find compensation for their absence in already extant, more general stylometric surveys (Stamatatos 2009, Tempestt 2017). Specifically for stylometry applied to medieval texts, no scholar today has done more than Mike Kestemont (1985–; e.g. Kestemont 2012, but cited *passim* below), stationed at the University of Antwerp, and a close associate member of the Computational Stylistics Group at Kraków's Institute of Polish Language, which houses important contemporary names within the field of stylometry such as Maciej Eder and Jan Rybicki.

2. Research Questions and Applications

Writing style is argued to be inferable from 'stylistic features,' also see §3(a) below, that recur with a predictable regularity in a writer's language use. Thanks to frequency patterns, computers recognize a writing style and can predict with a high accuracy if two segments of text are mathematically 'different' or 'distant' in function of these predetermined formal-linguistic 'features.' Stylometry's most successful (yet far from exclusive) application is non-traditional authorship attribution, where on the basis of feature frequencies a disputed anonymous text is ascribed to a candidate author. In the context of attribution, practitioners of stylometry often argue that writers can be shown to have a stylistic fingerprint, or a stylistic DNA, which makes their writing style predictable and distinguishable from that of others. In this context it deserves mentioning that stylometry and related methods also have a practical use in society, e.g. to extract profile information with commercial purposes, or to identify ill-intended written communication, e.g. cyber criminals, fake news spreaders or other kinds of impersonators. This asset has resulted in some popularizing press coverage for the method (Cafiero and Camps 2022). J.K. Rowling (1965–), for instance, was compelled to admit to stylometric evidence that not Robert Galbraith (a pseudonym) but she herself was the true author of *The Cuckoo's Calling* (2013). Similar stories have been published in popular media, where stylometric experimentation demonstrated that Harper Lee's classic *To Kill a Mocking Bird* (1960) exhibited substantial interventions from her editor, or where the pseudonymous Elena Ferrante was argued to be the Neapolitan writer Domenico Starnone (1943–) (possibly in collaboration with his spouse, translator Anita Raja), an ascription which Starnone, however, denies in public.

In its application to medieval texts, computational stylistics can be understood as a unique extension of the medievalists' text-critical toolset (e.g. close reading, palaeography, codicology, diplomatics etc.) Stylometry applied to medieval texts comes with domain-specific challenges which is brought about by a combination of factors that uniquely characterize the period's literary production. The Middle Ages boast a varied corpus of texts in different forms (cf. prose and verse), registers (cf. *probationes pennae* vs. highly stylized literature) and genres (cf. theological treatises vs. courtly romances). Moreover, the texts are written in various languages, sometimes with a degree of orthographical flux. Medieval writing products are fairly more complex and layered than those of modern times, with authorship being not seldom collective, rooted in an oral-aural interaction, reliant on a written and oral tradition, prone to imitation of normative common stock examples, and with a material transmission of manuscripts through the hands of copyists and scribes.

Many different medieval text types and genres have been subjected to stylometric analyses, with varying research aims. Sometimes the hope is to provide more objective angles on ongoing attribution disputes (which is at least as old as Yule's 1939 attribution of *De imitatione Christi* to Thomas of Kempen). This cannot only provide a breakthrough in the reconstruction of a historical context or an author's biography, it can also challenge established literary viewpoints and stir up ideological discussions. Since its early days with the figure of Lutosławski, stylometry has also been applied to assess how time impacts an author's stylistic development, a subfield called stylochronometry (Stamou 2007), a potential that however has been largely underexplored for medieval texts. Another subtask to which stylometry has been applied is to gauge the impact of scribes and copyists (van Dalen-Oskam 2012), or to investigate manuscript dependencies (MacPherson and Tirosh 2020), which can be a helpful extension to palaeography or stemmatology. Stylometry has been applied to investigate medieval writers' preoccupation with identical source texts, or their adherence to similar schools by means of imitation, influence and apprenticeship (Battles 2019). It can cast new light on the interactions of authors, and the dynamics at play between texts' multiple contributors during the composition process (Kestemont *et al.* 2015, De Gussem 2017). Especially for the less conservative vernaculars, scriptological and dialectological questions on the impact of diachronic and diatopic linguistic phenomena (spelling and dialectical variations within written and spoken languages through time and across regions) on individual and collective styles play a vital role (Camps 2018). Although direct cross-lingual comparisons between medieval texts written in different languages have not been attempted thus far, questions of language contacts, translingual influence, or translatorship in the Middle Ages occasionally enter the discussion in order to explain intrinsic stylistic deviations (Edlich-Muth and Edlich-Muth 2019). Finally, although studies rarely have an explicit ambition to 'detect' the genre of a medieval text, many studies reflect on the way in which genre affects style, or at the extent by which polygeneric or pluriform (prose and verse) oeuvres of authors may jeopardize their unitary stylistic signal (Richards *et al*. 2016). Some stylometric scholarship has investigated the (dis)continuities within well-demarcated genres by pinpointing its generic features, which naturally works particularly well for rather conservative writing traditions (e.g. Latin epistolography in Stella 2014).

3. State-of-the-art Techniques

Far from pretending to be an exhaustive state of the art, this four-tiered section covers common techniques, workflows and setups for stylometry's application to medieval sources.

(1) *Preprocessing*

The preprocessing of electronic text files or transcriptions / editions of medieval texts is an elementary yet fundamental step in ensuring valid and reliable results.

It entails the removal and editing of all irrelevant characters in the text (punctuation, numerals, optical character recognition errors, case-folding, titles or annotations, etc …) and 'tokenize' the text to meaningful units, often word tokens. If so desired, the preprocessing phase takes care also to normalize (or align) variant orthographical and editorial conventions between text versions, especially with the vernaculars. This is especially relevant when considering that many current editions of medieval texts tend to comply with the orthography of base manuscripts instead of standardizing spelling to classical models. In some cases, normalization comes paired with disambiguation after recognition, for instance to semantically distinguish homographs. More sophisticated automated annotation techniques have been proposed as part of preprocessing. Stemming, for instance, recovers the basis stem or morphological root of word tokens by removing affixes and pruning it down to a base root or unit for analysis. Lemmatization transforms word tokens to a standard dictionary form. Morphological taggers and syntactic parsers can identify a token's part of speech and syntactic function. One can use automated scansion tools for prosodic units of analysis. At this moment, fairly accurate software circulates to supply rich annotations as mentioned above, but they are predominantly attuned to modern-day natural languages, and despite ongoing progress (e.g. Manjavacas *et al*. 2019, Camps *et al.* 2021) there still remains much work to be done to increase the performance for pre-orthographic, historical languages.

Finally, preprocessing also entails the ‘sampling’ or ‘segmentation’ of a longer text into shorter segments. Establishing a secure minimal length to guarantee accuracy remains a tenacious problem for medieval texts, whose characterization by sparseness and fragmentation are not seldom responsible for the problems of authorship, dating, transmission, etc. which stylometry is regularly expected to solve in the first place. Exhaustive studies on sample length have shown its significance for achieving robust results (Eder 2013), but with some exceptions (Eder 2015) too few studies have attempted to make a systematic assessment of what minimal requirements of sample length apply to medieval texts (in different languages). With omission of permutations, sampling methods generally fall within four categories. (1) Discrete sampling divides a longer text in segments according to a predefined fixed sample size. (2) Rolling sampling divides the text in non-identical, partially overlapping windows instead of discrete chunks of text. (3) Random sampling selects sentences from a certain author’s or text’s entire oeuvre until a predefined sample size limit is reached (e.g. 1,000 words) in order to come to an almost inexhaustible number of new, real-world representations of the author’s lexical distribution. (4) Generative sampling, hardly explored for historical texts, uses text generation methods to expand an extant corpus by generating new samples (Manjavacas *et al*. 2017).

(2) *Feature Extraction and Selection*

Stylometric studies focus on patterns in the language which are argued to relate to or have an impact on style. With feature extraction, practitioners of stylometry identify an initial set of linguistic items which they believe is relevant to their problem. Early practitioners of stylometry focussed on *sentence length* (Yule 1939:372–77), but that feature type is rarely revisited nowadays. The majority of stylometric studies of medieval texts have focussed on lexical features, especially *most frequent words* (MFW) and *function words*. Some studies automatically extract a set of words through a fixed list (e.g. retain the *hapax legomena*) or by integrating restrictions during extraction (e.g. only compound words, rhyme words, etc.). *Lexical richness* measures the diversity within an author’s vocabulary. Another popular type of feature is the *n*-gram, where sequences of *n* (variable number of) characters / words / parts-of-speech from a given sample of text or speech are taken as the feature input (e.g. character bigrams of *uenerabili* are [‘ue’, ‘ne’, ‘ra’, ‘bi’, ‘li’]). Some feature types can only be extracted after more sophisticated preprocessing has taken place (see §3(1) above). Examples are *part-of-speech* and *syntactic tags*, or *prosodic features* such as phonological stress / accent or syllable quantity (metre). Recently, even semantic features with word embeddings (Mikolov *et al*. 2013) are gaining some leeway, but its application to medieval texts is near non-existent.

Not to be confused with feature extraction, feature selection techniques filter out features considered redundant. Sometimes the selection is done manually on the basis of qualitative arguments (Kestemont *et al.* 2013:206), others (filter-based and wrapper-based methods) are automated and base themselves on quantitative properties of the features (e.g. ‘culling’ in Hoover 2004). Machine-learning techniques (below §4(b)) can be particularly helpful in evaluating the efficiency and relevance of features.

(3) *Vectorization and Scaling*

Once the stylistic feature set has been identified and selected with care, each text (segment) can be characterized and numerically represented in terms of how frequent each of these features is. This process is called vectorization, and it lies at the core of what makes stylometric analysis so powerful. Vectors guarantee an objective basis for comparability of texts. They allow algorithms to process textual characteristics automatically, simultaneously, fast and on a large scale, often disregarding word order (the bag-of-words approach). A vector α, filled with the feature frequencies of text α, after all, becomes a solid basis for comparison with vectors β, γ, ... which in their own turn represent the stylistic properties of text samples β, γ, ... This common ground allows to approach stylistic difference literally as mathematical difference and geometric distance. Often, the ensuing analysis will take place not on the raw frequencies of features, but on 'normalized' or 'weighted' frequencies (feature scaling).

<FIGURE 1 HERE>

Figure 1. Schematic intuition of vectorization, from unscaled function word frequencies (left) to vectors (columns of numbers on the right containing frequencies).

The length of the vector corresponds to the number of features and is sometimes referred to as the vector dimensionality.

(4) *Data Analysis and Visualization*

Once converted to a numerical format, the ensuing ways to manipulate and analyse medieval texts are numerous. In general, two fundamental approaches can be discerned, based on a classic data-scientific distinction, (a) non-supervised and (b) supervised.

(a) *Unsupervised*

Unsupervised methods allow to inspect data and enable a detection of trends on an exploratory basis. As opposed to supervised learning (explained below, §4(b)), it needs no labelled training input in order to be operational. Its aim is to describe without prescience, not to classify. By implication, the general disadvantage of unsupervised methods is that there is limited possibility of evaluation (a degree-of-error) for each observation, nor can it give estimations of which textual material is perhaps redundant for certain questions. Therefore, practitioners strongly advise to back up unsupervised analysis with supervised evidence when that proves feasible.

Examples of unsupervised techniques applied to medieval texts are numerous. Popularly used are principal components analysis or PCA (Kestemont *et al.* 2015), *k* nearest neighbours (potentially visualized with network analysis; e.g. Eder 2016:72), t-distributed stochastic neighbor embedding or t-SNE (Leclercq and Kestemont 2021:229) and dendrogram analysis (e.g. Dockray-Miller *et al.* 2021).

<FIGURE 2 HERE>

Figure 2. Example of PCA, where the text corpora of three twelfth-century Latin authors appear as three distinct clusters (Bernard of Clairvaux in red, Hildegard of Bingen in orange, and the latter's secretary Guibert of Gembloux in blue). The text samples annotated with H2_epG represent letters in Hildegard's epistolarium carrying her secretary Guibert's stylistic revisions (approaching his cluster in the bottom left corner), and the green text samples D_Mart and D_Missa are dubious visions (hence, 'D') *Visio de Sancto Martino* and the *Visio ad Guibertum missa*, betraying a strong influence of Guibert's idiom (reproduced from Kestemont *et al.* 2013:257, also reprinted in Kestemont *et al.* 2015:218).

<FIGURE 3 HERE>

Figure 3. Example of consensus network of 150 Latin texts produced with Gephi, coloured according to chronology: ancient texts in blue, medieval texts in green, and early modern ones in red (Eder 2016:72). Networks allow to include many authors and facilitate a comprehensive exploration of the styles of texts from different periods (stylochronometry) and geographical contexts.

(b) *Supervised*

Supervised algorithms (or models) observe the way in which the feature inputs *X* (e.g. function word frequencies of a text) correlate with class outputs *y* (e.g. authorship of the text). A supervised model 'observes' correctly labelled, preclassified *X*-*y* pairs, in order to register meaningful correlations (e.g. between certain function word patterns *X* with certain authors *y*). This process is called 'training,' and the *X*-*y* pairs which the model trains on correspond to what is commonly referred to as training data. Consequently, once this learning process has taken place, the model can be confronted with 'test data,' comprising previously unobserved and unclassified texts. On the basis of what it has observed, the supervised model can make a prediction, and assign the unseen test data to a class, either by a hard decision or by outputting a probability score.

The advantage of supervised machine-learning methods, 'text classification' (Sebastiani 2002), is the possibility of evaluation. By making different combinations of parameters, such as the feature set, the vector length (number of features), sample length, vectorization method, scaling method, etc., and evaluating how well they can be fitted to a class (author), scholars can finetune and optimize these parameters. The essential insights and practices of machine learning, techniques that enable to statistically evaluate the compatibility of feature sets to certain problems, have perhaps not sufficiently found their way into the field of stylometry applied to historical literature. Perhaps two reasons lie behind this slow integration. Firstly, the learning threshold to 'supervised' techniques can be considerably steep as it intersects with mathematics, statistics, and computer science. Secondly, machine-learning techniques rely on prelabelled, securely attributed text as training data, not always feasible for the Middle Ages.

4. **Future of Stylometry for the Middle Ages**

Like many methods advancing from the ongoing digital revolution, stylometry is a field in its infancy. A glimpse at the publication years of the works cited in this entry's bibliography may suffice as demonstration of this. Nevertheless, recent special issues in *Speculum* (Birnbaum *et al.* 2017) and *Interfaces* (Deploige and De Gussem 2021) and entire dedicated journals such as *The Digital Medievalist* and *Digital Philology: A Journal of Medieval Cultures* suggest that the digital has incontrovertibly arrived in medieval studies and is not leaving any time soon.

Three lines of concerns about the future of stylometry may be observed here. (1) The first is methodological. For medieval texts particularly, there is still a dire lack of a systematic, large-scale encompassing studies that test and evaluate the application of stylometric methods across medieval languages and genres taking into account the specific challenges set by medieval text types. The fact that most stylometric algorithms have been optimized with modern-day texts and languages as their training data, preserved under stable conditions and with a relatively secure provenance, is often recognized in isolated publications, but has not been scientifically responded to. The sometimes inaccessible machine-learning evaluations (discussed above in §3(b)) may be key here. The problem is that progress in this area is also strongly hampered by feasibility, and with the still ongoing, partial disclosure of annotated electronic editions and transcriptions of medieval texts. (2) In order to appeal more to medievalists, stylometry will have to keep insinuating itself in debates medievalists are currently engaged in. Although some large authorship questions remain a matter of fascination and intrigue, stylometry should not only be applied to hackneyed philological, authorship-oriented questions (a point made in Birnbaum *et al*. 2017:S31 and Deploige and De Gussem 2021). (3) Despite approaches to make stylometry more openly available with open-source programming code and available software packages (Stylo with R, JGAAP (Java Graphical Authorship Attribution Program), Lexos), stylometry is still struggling to become less of an in-crowd specialty, which is still discouraging conversation which could otherwise be most fruitful.

Further reading

Bailey, R.W. (1969) ‘Statistics and Style: A Historical Survey’, in L. Doležel and R.W. Bailey (eds) *Statistics and Style*, New York: American Elsevier Publishing, 217–36. (Early chapter on early quantitative approaches to style.)

Battles, P. (2019) ‘Using N-gram Analysis to Map Intertextual Networks in Old English Verse’, *Digital Philology: A Journal of Medieval Cultures 8*: 155–91. (Original article that extends stylometry’s usefulness beyond attribution to establishing poetic networks.)

Binongo, J.N.G., and Smith, M.W.A. (1996) ‘Stylometry’, *Notes and Queries 43*: 448–52. (Short article on defining ‘stylometry’ from a historical, terminological perspective.)

Birnbaum, D.J., Bonde, S., Kestemont M. (2017) ‘The Digital Middle Ages: An Introduction’, *Speculum 92*: S1–S38. (Introduction to a special supplement of *Speculum* devoted to the possibilities of Digital Humanities in medieval studies, containing a helpful history of DH and introducing the full scope of state-of-the-art possibilities.)

Burrows, J.F. (2002) ‘Delta: A Measure of Stylistic Difference and a Guide to Likely Authorship’, *Literary and Linguistic Computing 17*: 267–87. (Ground-breaking article on the later immensely popular distance metric ‘Delta’.)

Cafiero, F., and Camps, J.-B. (2022) *Affaires de style: du cas Molière à l’affaire Grégory, la stylométrie mène l’enquête*, Paris: Le Robert. (Popularizing book on stylometry, covering many of the best-known cold cases the method has tackled from the Bible to Kurt Cobain.)

Camps, J.-B. (2018) ‘Manuscripts in Time and Space: Experiments in Scriptometrics on an Old French Corpus’, *Corpus-Based Research in the Humanities CRH-2*: 55–64. (Scriptometric article that attempts to isolate the impact of local speech and writing habits in written Old French *chansons de geste*.)

Camps, J.-B., Clérice, T., Duval, F., Ing, L., Kanaoka, N., and Pinche, A. (2019) ‘Corpus and Models for Lemmatisation and POS-tagging of Old French’, *Journal of Data Mining and Digital Humanities*: 1–19. (On the lemmatisation and part-of-speech tagging of Old French, making reference to the still developing language-independent post-correction interface Pyrrha for part-of-speech tagging and lemmatization.)

De Gussem J. (2017) ‘Bernard of Clairvaux and Nicholas of Montiéramey: Tracing the Secretarial Trail with Computational Stylistics’, *Speculum 92*: S190–S225. (Stylometric analysis of the collaboration between Abbot Bernard of Clairvaux and his secretary Nicholas of Montiéramey on Bernard’s letter corpus.)

Deploige J., and De Gussem J. (2021) ‘Medieval Authorship and Canonicity in the Digital Age – an Introduction’, *Interfaces 8*: 113–24. (Introduction to a cluster of stylometric articles in *Interfaces*, containing reflections on the stylometry’s potential to breach medievalists’ restraints caused by canonicity in questions of medieval authorship.)

Delcourt, C. (2002) ‘Stylometry’, *Revue belge de Philologie et d'Histoire 80*: 979–1002. (Critical state of the art with attention to stylometry’s history and eye for the use and misuse of stylometry.)

Dockray-Miller D., Drout M. D.C., with Kinkade, S., and Valerio J. (2021) ‘The Author and the Authors of the *Vita Ædwardi Regis*: Women’s Literary Culture and Digital Humanities’, *Interfaces 8*: 160–213. (Recent article on the collaborative authorship of the *Vita Ædwardi Regis* that applies dendrogram analysis.)

Eder, M. (2016) ‘A Bird’s Eye View of Early Modern Latin: Distant Reading, Network Analysis and Style Variation’, in L. Estill, D.K. Jakacki, and M. Ullyot (eds) *Early Modern Studies after the Digital Turn*, 61–88. (Chapter advocating the use of network analysis for large-scale stylometry, and moreover containing a good introduction to stylometry’s history.)

—. (2013) ‘Does Size Matter? Authorship Attribution, Small Samples, Big Problem’, *Digital Scholarship in the Humanities 30*: 167–82. (Important article with large-scale benchmarking experiments for different languages and genres that demonstrates the significant impact of sample length for reliable attribution.)

—. (2015) ‘Taking Stylometry to the Limits: Benchmark Study on 5,281 Texts from *Patrologia Latina*’ *Digital Humanities 2015: Conference Abstracts*, Sydney: 1919–24. (Enormous benchmarking experiment on the entire *Patrologia Latina* to assess the impact of corpus size on stylometric accuracy.)

Edlich-Muth, C., and Edlich-Muth, M. (2019) ‘A Computational Approach to Source Adaptation in Thomas Malory’s Morte Darthur’, *Digital Medievalist 12*: 1–26. (Stylometric analysis of Thomas Malory’s *Morte D’Arthur* which identifies sections that stylistically differ due to differing source text languages.)

Grzybek, P. (2014) ‘The Emergence of Stylometry: Prolegomena to the History of Term and Concept’, in K. Kroó and P. Torop (eds) *Text within Text - Culture within Culture*, Budapest, Tartu: L’Harmattan, 60–77. (Interesting chapter on the history of stylometry, which poignantly shows the many varieties in the historical narrative of the method’s conception.)

Hoover, D.L. (2004) ‘Testing Burrows’s Delta’, *Literary and Linguistic Computing 19*: 453–75. (An article approving, replicating and improving Burrows’s Delta, which also shows the benefice of ‘culling’.)

Kestemont, M. (2012) ‘Stylometry for Medieval Authorship Studies: An Application to Rhyme words’, *Digital Philology: A Journal of Medieval Cultures 1*: 42–72. (This article, aside from its exploration of the advantage of rhyme words for Middle Dutch, also contains interesting reflections of a more general nature on the application of stylometry to medieval texts.)

Kestemont, M., Moens, S., and Deploige, J. (2013) ‘Stylometry and the Complex Authorship in Hildegard of Bingen’s Oeuvre’, *Digital Humanities 2013: Conference Abstracts*, Lincoln, Nebraska: 255–58. (A significant extension of stylometry’s usefulness to explore the dynamics of authorship and authority in medieval writing collaborations.)

—. (2015) ‘Collaborative Authorship in the Twelfth Century: A Stylometric Study of Hildegard of Bingen and Guibert of Gembloux’, *Digital Scholarship in the Humanities 30*: 199–224. (Extension and reprint of conference paper above.)

Leclercq, E., and Kestemont, M. (2021) ‘Advances in Distant Diplomatics: A Stylometric Approach to Medieval Charters’, *Interfaces 8*: 214–44. (‘Distant diplomatics’ applied to a corpus of twelfth-century Latin charters from the Cambrai episcopal chancery by making use of t-SNE.)

Lutosławski, W. (1897 [1905]) *The Origin and Growth of Plato's Logic: with an Account of Plato's Style and of the Chronology of his Writings*, New York: Longmans Green. (A book of one of the first practitioners of stylometry, in which the term as such, however, is never used.)

—. (1897) 'On Stylometry', *Classical Review 11*: 284–6. (Lutosławski's paper in Oxford that apparently coined the term 'stylometry'.)

MacPherson, M., and Tirosh, Y. (2020) 'A Stylometric Analysis of *Ljósvetninga saga*', *Gripla 31*: 7–41. (Stylometric analysis of differing manuscript redactions of the Old Norse *Ljósvetninga* to assess the primacy of one variant over another.)

Manjavacas, E., De Gussem, J., Daelemans, W., and Kestemont, M. (2017) 'Assessing the Stylistic Properties of Neurally Generated Text in Authorship Attribution', in *Proceedings of the Workshop on Stylistic Variation*, Copenhagen: 116–25. (The only extant article to my knowledge that generates 'new' literary Latin by making use of generative NLP models.)

Manjavacas, E., Kádár, Á., and Kestemont, M. (2019) 'Improving Lemmatization of Non-Standard Languages with Joint Learning', in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies 1*, Minneapolis, Minnesota: 1493–1503. (Article that proposes the PIE tagger for improving the task of lemmatization for non-standard historical languages tested on medieval and early modern corpora in Dutch, German, French, Slovene and Latin.)

Mendenhall, T.C. (1887) 'The Characteristic Curves of Composition', *Science 9*: 237–49. (Mendenhall's famous *Science* article, in which he arrives at visual evidence suggesting that texts of same authorship exhibit similar 'wave lengths' or 'characteristic curves of composition'.)

—. (1901) 'A Mechanical Solution of a Literary Problem', *Popular Science Monthly 60*: 97–105. (Another study of Mendenhall in which he mentions his tabulating device that produced reels of paper tape imprinted with word-length data.)

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., and Dean, J. (2013) 'Distributed Representations of Words and Phrases and Their Compositionality', *Advances in Neural Information Processing Systems 26*: 3111–9. (Ground-breaking paper on word embeddings that are argued to capture the semantic properties of words.)

Mosteller, F., and Wallace, D.L. (1984 [1964]) *Applied Bayesian and Classical Inference: The Case of the Federalist Papers*, New York: Springer Verlag. (The foundational book that put computational stylometry on the map.)

Richards, E.J., Wrisley D.J., and Dulac L. (2016) 'The Different Styles of Christine de Pizan. An Initial Stylometric Analysis', *Le moyen français 78–9*: 187–206. (A useful reflection on the stylistic variation that a single author can exhibit.)

Sebastiani, F. (2002) 'Machine Learning in Automated Text Categorization', *ACM Computing Surveys 34*: 1–47. (The classic survey article on the main approaches to text categorization.)

Stamatatos, E. (2009) 'A Survey of Modern Authorship Attribution Methods', *Journal of the American Society for Information Science and Technology 60*: 538–56. (A gradually becoming outdated—yet likely to remain relevant—survey article of non-traditional authorship attribution methods.)

Stamou, C. (2007) 'Stylochronometry: Stylistic Development, Sequence of Composition, and Relative Dating', *Literary and Linguistic Computing 23*: 181–99. (A state of the art of stylochronometric research.)

Stella, F. (2014) 'Generic Constants and Chronological Variations in Statistical Linguistics on Latin Epistolography', in T. Andrews and C. Macé (eds) *Analysis of Ancient and Medieval Texts and Manuscripts: Digital Approaches*, Turnhout: Brepols, 159–79. (Interesting stylometric observations on the generic specificities of Latin epistolography from the ancient republican period to the humanists.)

Tempestt, N., Kalaivani, S., Aneez, F., Yiming, Y., Yingfei X., and Woodard D. (2017) 'Surveying Stylometry Techniques and Applications', *ACM Computing Surveys 50*: 1–36. (A very recent and general survey of stylometric techniques of a more technical nature.)

Van Dalen-Oskam, K. (2012) 'The Secret Life of Scribes. Exploring Fifteen Manuscripts of Jacob van Maerlant's *Scolastica* (1271)', *Literary and Linguistic Computing 27*: 355–72. (Application of stylometry to gauge the stylistic impact of scribes on varying versions of Jacob of Maerlant's *Scolastica.*)

Yule, G.U. (1939) 'On Sentence-Length as a Statistical Characteristic of Style in Prose: With Application to Two Cases of Disputed Authorship', *Biometrika 30*: 363–90. (Some of Yule's early work on sentence length, with a section dedicated to the authorship of *De imitatione Christi* by Thomas of Kempen.)