

Chapter 8

FUTURE TRENDS IN AUTHORSHIP ATTRIBUTION

Patrick Juola

Abstract Authorship attribution, the science of inferring characteristics of an author from the characteristics of documents written by that author, is a problem with a long history and a wide range of application. This paper surveys the history and present state of the discipline – essentially a collection of *ad hoc* methods with little formal data available to select among them. It also makes some predictions about the needs of the discipline and discusses how these needs might be met.

Keywords: Authorship attribution, stylometrics, text forensics

1. Introduction

Judges 12:5–6 describes a harsh, but linguistically insightful, solution to the problem of identifying potential security threats:

5 And the Gileadites took the passages of Jordan before the Ephraimites: and it was so, that when those Ephraimites which were escaped said, Let me go over; that the men of Gilead said unto him, Art thou an Ephraimite? If he said, Nay;
6 Then said they unto him, Say now Shibboleth: and he said Sibboleth: for he could not frame to pronounce it right. Then they took him, and slew him at the passages of Jordan: and there fell at that time of the Ephraimites forty and two thousand.

This barbaric method would not stand up to modern standards regarding rules of evidence – or for that matter, civil rights such as fair trials, appeals or access to counsel. But it illustrates a problem that modern society also has to grapple with. In a world full of bad guys, how can one sort the bad guys from the good ones? If you are looking for a specific bad guy, how can you tell when you have found him?

Please use the following format when citing this chapter:

Juola, P., 2007, in IFIP International Federation for Information Processing, Volume 242, Advances in Digital Forensics III; eds. P. Craiger and S Shenoii;(Boston: Springer), pp. 119-132.

Much of the field of forensics addresses this problem. If you know something about who you are looking for, you look for identifying features specific to that person. These features might be his fingerprints, his DNA, his shoeprints – or his language.

2. Problem Definition

“Authorship attribution,” is broadly defined as the task of inferring characteristics of a document’s author, including but not limited to identity, from the textual characteristics of the document itself. This task, of course, has been the bread-and-butter of handwriting examiners, who are recognized as experts at spotting the idiosyncratic loop of an ‘e’ or slant of an ‘l’ that reliably characterize the writer. Similar experts can testify to the off-line character that identifies a particular typewriter. But software-only documents – this paper qualifies – do not have these quirks; one flat-ASCII ‘A’ looks identical to any other. Traditional network forensics may be able to trace a piece of email to a specific computer at a specific time. But who was actually sitting at that computer (perhaps at a public-access terminal in a public library) and typing the document?

Chaski [10, 11] has published several case studies where the authorship of specific digital documents has been a crucial factor in the case. In one case, for example, an employee was dismissed on the basis of emails admittedly written at her desk and on her computer. But in an open-plan office, anyone can wander into any cubicle and use any computer. Did she really write the relevant emails, or was she wrongfully dismissed? In another case, a software “diary” provided crucial exculpatory evidence against the claims of its author. But were the entries genuine, or had they been planted? In a third case, an investigation of a death turned up a suicide note “written” on a computer. Was this note genuinely written by the decedent, or had it been written by the murderer to cover his tracks?

In each of these cases, the computer that created the document was not in question, but the authorship of the document was. And in each case, traditional handwriting analysis was out of the question because the documents were purely electronic. As electronic documents continue to flourish (and possibly displace paper), one can expect that questions and cases like these will only become more common, more divisive and more important.

Of course, the problem of inferring authorship is not limited to the present nor to issues of litigation and criminal prosecution. Generations of scholars have discussed at length the question of whether or not

William Shakespeare was the actual author of the works commonly attributed to him [14, 30]. The 17th-century tradition of anonymous political pamphlets (e.g., *Common Sense*, attributed by scholars to Thomas Paine) has provided work for thousands of historians. This tradition, of course, continues to the present day, for example, in the anonymous publication of *Primary Colors*, attributed to Joe Klein, and *Imperial Hubris*, attributed to Michael Scheuer.

Traditional approaches to this kind of analysis involve close reading by scholarly experts. The potential problems with this are apparent. Although Shakespeare experts may be readily available, how many experts are there on the writing style of Michael Scheuer? And how can the accuracy of such experts be measured, especially to the demanding standards of a Daubert hearing? For this reason, there has been increasing attention paid in recent decades to the development of testable, objective, “non-traditional” methods of authorship attribution, methods that rely not simply on expert judgment, but on the automated detection and statistical analysis of linguistic features. This emerging discipline has been variously called “stylometry,” “non-traditional authorship analysis,” or simply “authorship attribution.”

Within this general framework, we can also identify several different categories of problems. For example, the Judges quotation above is not about identifying individual persons, but identifying members of a specific class (the Ephraimites). This is analogous to the difference between analyzing handwriting using “document analysis” and “graphoanalysis” – where a document analyst may feel comfortable identifying a specific person as the author of a document, she may be quite out of her depth at describing characteristics of that person, such as age, sex, race and personality profile. The discipline of “graphoanalysis” claims roughly the opposite, to be able to analyze handwriting “to provide insight into personality traits and evaluation of a writer’s personality” [20], but not to identify individuals. In contrast, stylometrists have been exploring a wide variety of related questions and research topics:

- **Did this person write that document?** This, of course, is probably the oldest, clearest and best-defined version of the authorship attribution question. It has garnered considerable research attention. Holmes [15] provides a good history of the research in this field dating back to the late 1800s.
- **Which of these people wrote that document?** This slightly harder problem describes a typical authorship investigation [35, 43]. The Mosteller-Wallace investigation [35] described later is almost a prototypical example, in that the set of candidate authors

was well-defined by previous investigation, and the analysts were merely called upon to referee among them. One can, of course, distinguish two sub-categories of this problem, depending upon whether or not “none of the above” is an acceptable answer.

- **Were all these documents written by the same person?** In many cases, this could serve as a starting point for further investigation [5, 7], but this question may be evidence in and of itself. For example, if a purportedly single-author “diary” or “journal” could be shown to have multiple authors, that in and of itself could show tampering, without needing to find and name alternate authors.
- **When was this document written?** This type of analysis could apply either to the development of a person’s writing style [9, 24] or to the general *Zeitgeist* typical of a certain period [23]. In either case, the evidence provided by document dating could prove to be crucial in settling issues of timing.
- **What was the sex of the author?** This question could obviously be generalized from sex [32] to many other aspects of group identity, such as education level [4, 28].
- **What was the mental capacity of the author?** To the extent that language capacity can be used as a diagnostic instrument [6], it can also be used as evidence of someone’s mental capacity or to identify the severity and type of insanity.

Given the wide variety of questions that can be posed within this framework, it is not surprising that an equally wide variety of people may be interested in the answers. Beyond the obvious forensic applications for law enforcement and the legal profession, and the equally obvious applications for literature scholars, interested groups might include historians, sociologists and other humanities scholars, educators in general (especially those concerned about plagiarism and academic integrity), psychologists, intelligence analysts and journalists. The ability to infer past the words to the author is a key aspect of many types of humanistic inquiry.

3. Theory of Stylometry

Although the problem of stylometry has been around since antiquity, the specific application of statistics to this problem is about 150 years old. Holmes [15] cites Mendenhall’s 1887 study [34] as the first modern example of statistical stylometry, and traces a flurry of research during the 20th century. A key historical development was the detailed

Mosteller-Wallace [35] study of *The Federalist Papers*, as described below.

3.1 Theoretical Background

From a theoretical perspective, stylometry is no different from many other accepted forensic techniques. Traditional handwriting analysis, for example, assumes that people have specific, individual, persistent and uncontrollable habits of penmanship that can be reliably identified by skilled practitioners. A similar theory underlies DNA analysis, fingerprints, toolmarks, and ballistic markings – one cannot consciously or through an effort of will change the patterns of ridges on one’s fingers or the genetic markers in one’s blood. Authorship attribution assumes similarly that people have specific, individual, persistent and uncontrollable habits of thought and/or phrasing; some researchers [41] call this a “stylome” in deliberate analogy to the DNA “genome.” Although the strict implications of this analogy may be incorrect – in particular, if the “stylome,” like the “genome,” is fixed and unchangeable, how is it possible to do document dating via stylometry? – it is nevertheless fruitful to explore.

Language, like genetics, can be characterized by a very large set of potential features that may or may not show up in any specific sample, and that may or may not have obvious large-scale impact. The Judges passage above is one example; each language (or dialect subgroup) has a characteristic inventory of sounds. Other accessible examples would include lexical items characteristic of a particular dialect, cultural or social group (such as “chesterfield” among Canadians [12] or “mosquito hawk” among southeastern Americans [21]), or individual quirks (such as an idiosyncratic misspelling of “touch” [42]). By identifying the features characteristic of the group or individual of interest, and then finding those features in the studied document, one can support a finding that the document was written by that person or a member of that group. The question, then, becomes what features are “characteristic” and how reliable those features are.

Unfortunately, the current state of affairs is something of an *ad hoc* mess. A 1998 paper [37] stated that more than 1,000 different features have been proposed at various times for authorship attribution. At least several dozen analytic methods have been proposed, and even the methods of document treatment vary widely. As a result, space precludes a detailed or even a cursory examination of all the research over the past century and a half.

3.2 Examining the Proposals

Instead, we will try to analyze the structure of the proposals themselves. The first obvious area of variance is the type of feature to be analyzed. An analysis of “average word length,” for example, is not influenced by an author’s grammar and syntax, but only her vocabulary. An analysis of the distribution of sentence lengths is, by contrast, only influenced by her grammar/syntax and not at all by her vocabulary. We can divide the types of features analyzed into several broad categories:

- **Lexicographic:** Analysis of the letters and sub-lexical units (such as morphemes) in the document [22, 26].
- **Lexical:** Analysis of specific words or their properties such as length and distribution, or an analysis of the general content [3, 18, 19, 43].
- **Syntactic:** Analysis of syntactic patterns, including aspects such as word n-grams, distribution of parts of speech, and punctuation. We can also include function word analysis into this category, since function words tend to illustrate syntactic rather than lexical or semantic aspects of the text [4, 7, 31, 35, 40].
- **Layout:** Use of formatting, spacing, color, fonts and size changes, and similar non-linguistic aspects of information presentation [1, 2].
- **“Unusual:”** Finally, of course, there are examples that fail to fit neatly into any of these categories or that span multiple levels of language [10, 42].

We could make a similar categorization of the proposed analysis methods used – principal components analysis [5, 7, 16], “delta” (*t*-tests) [8, 18, 19], linear discriminant analysis [4, 41], similarity judgments and *k*-nearest neighbors [23, 28], neural networks [40], support vector machines [3], etc. – but would simply end up cataloging the fields of machine learning, classification and pattern recognition.

What has unfortunately not been performed, but is badly needed, is a systematic cross-comparison of these methods. The Daubert criteria for evidence demand as a principle of law that any proposed forensic analysis method be subject to rather stringent study, including published error analysis. Legal requirements aside, any ethical analyst, of course, wishes to use the most appropriate and most accurate methods available, and should be able to make informed and objective choices about what those “best practices” should be. Furthermore, it is possible, perhaps even

likely, that the most effective practices may be a combination of features and techniques from several different proposed methods. What is needed is a stable, objective, and representative test bed to compare proposed methods head-to-head.

As an example of the current muddle, we look in more detail at the Mosteller-Wallace study [35] and its aftermath. *The Federalist Papers* are a set of newspaper essays published between 1787 and 1788 by an anonymous author named “Publius,” in favor of the ratification of the newly-proposed Constitution of the United States. It has since become known that “Publius” was a pseudonym for a group of three authors: John Jay, Alexander Hamilton and James Madison. It has also become known that of the eighty-odd essays, Jay wrote five, Madison wrote fourteen, and Hamilton wrote 51, with three more essays written jointly by Madison and Hamilton. The other twelve essays, the famous “disputed essays,” are attributed to both Madison and Hamilton.

Modern scholarship is almost unanimous in assigning authorship of the disputed essays to Madison on the basis of traditional historical methods. Mosteller and Wallace were able to make this determination purely on the basis of statistically-inferred probabilities and Bayesian analysis.

In particular, we note that an author has almost complete freedom to choose between the words “big” and “large” or similar synonym pairs; neither the structure of English grammar nor the meanings of the words place any constraints. By observing that one author consistently makes one choice and another the opposite, one has a noticeable, topic-free, and consistent way to differentiate between the authors.

Mosteller and Wallace [35] attempted to apply this technique to *The Federalist Papers*, but found that there were not enough synonym pairs to make this practical. Instead, they focused on so-called “function words,” words like conjunctions, prepositions and articles that carry little meaning by themselves (think about what “of” means), but that define relationships of syntactic or semantic functions between other (“content”) words in the sentence. These words are, therefore, largely topic-independent and may serve as useful indicators of an author’s preferred way to express broad concepts such as “ownership.” Mosteller and Wallace, therefore, analyzed the distribution of 30 function words extracted from the text of *The Federalist Papers*.

Because of the circumstances of this problem, *The Federalist Papers* are almost a perfect test bed for new methods of authorship attribution. First, the documents themselves are widely available (albeit with many potential corruptions), including over the Internet through sources such as Project Gutenberg. Second, the candidate set for authorship is

well-defined; the author of the disputed papers is known to be either Hamilton or Madison. Third, the undisputed papers provide excellent samples of undisputed text written by the same authors, at the same time, on the same topic, in the same genre, for publication via the same media. A more representative training set would be hard to imagine. For this reason, it has become almost traditional to test a new method on this problem, see, e.g., [17, 22, 33, 36, 38, 40]. However, even in this limited situation, it is not possible to compare results directly between analyses, since the corpus itself is not consistently defined. As Rudman [37, 38] has pointed out, different versions of the corpus contain numerous textual flaws, including differences in source versions, wrong letters and misspelling, inconsistency in decisions about what to include or to exclude (such as titles), inclusions of foreign language phrases and quotations, and so on. Even when two researchers ostensibly analyze the same documents, they may not be analyzing the same data!

Some attempts have been made to standardize test corpora for such purposes; examples include the Forsyth corpus [13], the Baayen corpus [4] and the Juola corpus [25]. For the most part, however, papers report on an analysis of samples of convenience that are not necessarily representative or even widely available. The accuracy rates reported usually hover in the 90% range, but that means little when one considers that 90% of ten documents between three authors is a far cry from 90% of a thousand documents between 250 authors.

The generalization question is also unaddressed. Do we have reason to believe that a method that performs brilliantly on Dutch will perform equally well on English? (Or vice versa?) Do successful methods generalize across genres? How much data is needed for a given method to work? Current studies [27] suggest that there is a strong correlation of method performance across different environments, but that may amount to little more than the observation that a method with only mediocre performance in one environment is not likely to miraculously improve in a new and untested one.

4. The Future

The field of authorship attribution is, therefore, in need of clearly-defined and well-documented standards of practice. Juola has defined a theoretical framework [26] to help establish these standards through the development of modular software [29] specifically to encourage this sort of validation, testing and possible cross-development. The details of this have been described elsewhere [27], and so will be summarized here. In short, the proposed system uses a three-phase structure abstraction of

canonization, event determination and inferential statistics, any of which can be defined in several different technical implementations.

Projects like this are, one hopes, only the tip of the iceberg in terms of what the future of authorship attribution will bring. There are a number of crucial issues that need to be addressed to make stylometry a fully-fledged and standard forensic discipline. Fortunately, the seeds of most of the issues and developments have already been planted.

Better test data, for example, is something of a *sine qua non*. Some test corpora have already been developed, and others are on the way. A key aspect to be addressed is the development of specific corpora representing the specific needs of specific communities. For example, researchers such as NYU's David Hoover have been collecting large sets of literary text such as novels, to better aid in the literary analysis of major authors. Such corpora can easily be deployed to answer questions of literary style, such as whether or not a given (anonymous) political pamphlet was actually written by an author of recognized merit, and as such reflects his/her political and social views, to the enrichment of scholars. Such a corpus, however, would not be of much use to law enforcement; not only is 18th or 19th century text unrepresentative of the 21st century, but the idea of a 100,000 word ransom note being analyzed with an eye towards criminal prosecution borders on the ludicrous. The needs of law enforcement are much better served by developing corpora of web log (blog) entries, email, etc. – document styles that are used routinely in investigations. So while we can expect to see much greater development of corpora to serve community needs, we can also expect a certain degree of fragmentation as different subcommunities express (and fund) different needs.

We can expect to see the current *ad hoc* mess of methods and algorithms to be straightened out, as testing on the newly developed corpora becomes more commonplace. Programs such as JGAAP [29] will help support the idea of standardized testing of new algorithms on standardized problems, and the software programs themselves can and will be made available in standard (tested) configurations for use by non-experts. Just as digital forensic tools like EnCase and FTK make file carving and undeletion practical, so will the next generation of authorship attribution tools.

At the same time, concerns such as Rudman's about the handling of questioned documents can be expected to crystallize into standardized procedures about treatment of "dubitanda," aspects of documents of questionable relevance to the authorship question. Further testing will formalize issues such as what kind of documents are "comparable" (it is highly unlikely that there will be ransom notes of known authorship

to compare with the ransom note found at the crime scene; are business letters sufficiently “comparable”?) or how much data is “enough” for a confident analysis (it is unlikely that anything useful can be learned from a single-expletive email, but equally unlikely that the suspect will provide investigators with millions of words of text). TREC-style competitions will more than likely provide a source of continuous improvement as well as establish a continuing stream of new “standard” test beds tuned to specific problems.

The new level of computer support will trigger new levels of understanding of the algorithms. Although some efforts (most notably Stein and Argamon [39]) have been made to explain not only that certain methods work, but also why they work, most research to date has been content with finding accurate methods rather than explaining them. The need to explain one’s conclusions to a judge, jury and opposing counsel will no doubt spur research into the fundamental linguistic, psychological, and cognitive underpinnings, possibly shedding more light on the purely mental aspects of authorship.

Finally, as scholarship in these areas improves and provides new resources, the acceptance of non-traditional authorship attribution can be expected to improve. Just as handwriting analysis and ballistics are accepted specialist fields, so will “authorship attribution,” with the corresponding professional tools, credentials and societies.

5. Conclusions

This paper has presented a survey, not only of the present state of non-traditional authorship attribution, but of what may be expected in the near-term future. It should be apparent that authorship attribution at present is at a crossroads or, perhaps, at a threshold. The current state of affairs is that of a professional adhococracy, where many different researchers have proposed many different methods, all of which tend to work. However, in the current muddle, there is no clear direction about which methods work better under what circumstances, about what the expected rates of reliability should be under field conditions, and about why particular methods work as well as they do.

These issues will need to be addressed if authorship attribution is to become an accepted forensic discipline like footprint, toolmark and fingerprint analysis. Fortunately, these issues are being addressed by researchers at the same time as non-specialists in the larger community – whether they be forensic scientists, law enforcement agents or even English professors – are becoming more aware and more accepting of the possibilities, pitfalls and potentials of authorship attribution.

References

- [1] A. Abbasi and H. Chen, Identification and comparison of extremist-group web forum messages using authorship analysis, *IEEE Intelligent Systems*, vol. 20(5), pp. 67–75, 2005.
- [2] A. Abbasi and H. Chen, Visualizing authorship for identification, in *Proceedings of the IEEE International Conference on Intelligence and Security Informatics (LNCS 3975)*, S. Mehrotra, et al. (Eds.), Springer-Verlag, Berlin Heidelberg, Germany, pp. 60–71, 2006.
- [3] S. Argamon and S. Levitan, Measuring the usefulness of function words for authorship attribution, *Proceedings of the Joint International Conference of the Association for Literary and Linguistic Computing and the Association for Computers and the Humanities*, 2005.
- [4] R. Baayen, H. van Halteren, A. Neijt and F. Tweedie, An experiment in authorship attribution, *Proceedings of JADT 2002: Sixth International Conference on Textual Data Statistical Analysis*, pp. 29–37, 2002.
- [5] J. Binongo, Who wrote the 15th Book of Oz? An application of multivariate analysis to authorship attribution, *Chance*, vol. 16(2), pp. 9–17, 2003.
- [6] C. Brown, M. Covington, J. Semple and J. Brown, Reduced idea density in speech as an indicator of schizophrenia and ketamine intoxication, presented at the *International Congress on Schizophrenia Research*, 2005.
- [7] J. Burrows, “an ocean where each kind. . . :” Statistical analysis and some major determinants of literary style, *Computers and the Humanities*, vol. 23(4-5), pp. 309–321, 1989.
- [8] J. Burrows, Questions of authorships: Attribution and beyond, *Computers and the Humanities*, vol. 37(1), pp. 5–32, 2003.
- [9] F. Can and J. Patton, Change of writing style with time, *Computers and the Humanities*, vol. 38(1), pp. 61–82, 2004.
- [10] C. Chaski, Who’s at the keyboard: Authorship attribution in digital evidence investigations, *International Journal of Digital Evidence*, vol. 4(1), 2005.
- [11] C. Chaski, The keyboard dilemma and forensic authorship attribution, in *Advances in Digital Forensics III*, P. Craiger and S. Shenoj (Eds.), Springer, New York, pp. 133–146, 2007.
- [12] G. Easson, The linguistic implications of Shibboleths, presented at the *Annual Meeting of the Canadian Linguistics Association*, 2002.

- [13] R. Forsyth, Towards a text benchmark suite, *Proceedings of the Joint International Conference of the Association for Literary and Linguistic Computing and the Association for Computers and the Humanities*, 1997.
- [14] W. Friedman and E. Friedman., *The Shakespearean Ciphers Examined*, Cambridge University Press, Cambridge, United Kingdom, 1957.
- [15] D. Holmes, Authorship attribution, *Computers and the Humanities*, vol. 28(2), pp. 87–106, 1994.
- [16] D. Holmes, Stylometry and the Civil War: The case of the Pickett Letters, *Chance*, vol. 16(2), pp. 18–26, 2003.
- [17] D. Holmes and R. Forsyth, The Federalist revisited: New directions in authorship attribution, *Literary and Linguistic Computing*, vol. 10(2), pp. 111–127, 1995.
- [18] D. Hoover, Delta prime? *Literary and Linguistic Computing*, vol. 19(4), pp. 477–495, 2004.
- [19] D. Hoover, Testing Burrows' delta, *Literary and Linguistic Computing*, vol. 19(4), pp. 453–475, 2004.
- [20] International Graphoanalysis Society (IGAS), (www.igas.com).
- [21] E. Johnson, *Lexical Change and Variation in the Southeastern United States 1930–1990*, University of Alabama Press, Tuscaloosa, Alabama, 1996.
- [22] P. Juola, What can we do with small corpora? Document categorization via cross-entropy, *Proceedings of the Interdisciplinary Workshop on Similarity and Categorization*, 1997.
- [23] P. Juola, The rate of language change, *Proceedings of the Fourth International Conference on Quantitative Linguistics*, 2000.
- [24] P. Juola, Becoming Jack London, *Proceedings of the Fifth International Conference on Quantitative Linguistics*, 2003.
- [25] P. Juola, Ad-hoc authorship attribution competition, *Proceedings of the Joint International Conference of the Association for Literary and Linguistic Computing and the Association for Computers and the Humanities*, 2004.
- [26] P. Juola, On composership attribution, *Proceedings of the Joint International Conference of the Association for Literary and Linguistic Computing and the Association for Computers and the Humanities*, 2004.
- [27] P. Juola, Authorship attribution for electronic documents, in *Advances in Digital Forensics II*, M. Olivier and S. Shenoj (Eds.), Springer, New York, pp. 119–130, 2006.

- [28] P. Juola and H. Baayen, A controlled-corpus experiment in authorship attribution by cross-entropy, *Literary and Linguistic Computing*, vol. 20, pp. 59–67, 2005.
- [29] P. Juola, J. Sofko and P. Brennan, A prototype for authorship attribution studies, *Literary and Linguistic Computing*, vol. 21(2), pp. 169–178, 2006.
- [30] D. Kahn, *The Codebreakers*, Scribner, New York, 1996.
- [31] V. Kešelj and N. Cercone, CNG method with weighted voting, presented at the *Joint International Conference of the Association for Literary and Linguistic Computing and the Association for Computers and the Humanities*, 2004.
- [32] M. Koppel, S. Argamon and A. Shimoni, Automatically categorizing written texts by author gender, *Literary and Linguistic Computing*, vol. 17(4), pp. 401–412, 2002.
- [33] C. Martindale and D. McKenzie, On the utility of content analysis in authorship attribution: The Federalist Papers, *Computers and the Humanities*, vol. 29(4), pp. 259–270, 1995.
- [34] T. Mendenhall, The characteristic curves of composition, *Science*, vol. IX, pp. 237–249, 1887.
- [35] F. Mosteller and D. Wallace, *Inference and Disputed Authorship: The Federalist*, Addison-Wesley, Reading, Massachusetts, 1964.
- [36] M. Rockeach, R. Homant and L. Penner, A value analysis of the disputed Federalist Papers, *Journal of Personality and Social Psychology*, vol. 16, pp. 245–250, 1970.
- [37] J. Rudman, The state of authorship attribution studies: Some problems and solutions, *Computers and the Humanities*, vol. 31, pp. 351–365, 1998.
- [38] J. Rudman, The non-traditional case for the authorship of the twelve disputed Federalist Papers: A monument built on sand, *Proceedings of the Joint International Conference of the Association for Literary and Linguistic Computing and the Association for Computers and the Humanities*, 2005.
- [39] S. Stein and S. Argamon, A mathematical explanation of Burrows' delta, *Proceedings of the Digital Humanities Conference*, 2006.
- [40] F. Tweedie, S. Singh and D. Holmes, Neural network applications in stylometry: The Federalist Papers, *Computers and the Humanities*, vol. 30(1), pp. 1–10, 1996.

- [41] H. van Halteren, R. Baayen, F. Tweedie, M. Haverkort and A. Neijt, New machine learning methods demonstrate the existence of a human stylome, *Journal of Quantitative Linguistics*, vol. 12(1), pp. 65–77, 2005.
- [42] F. Wellman, *The Art of Cross-Examination*, MacMillan, New York, 1936.
- [43] G. Yule, *The Statistical Study of Literary Vocabulary*, Cambridge University Press, Cambridge, United Kingdom, 1944.