A Decade of Shared Tasks in Digital Text Forensics at PAN

Martin Potthast[1] Paolo Rosso[2] Efstathios Stamatatos[3] Benno Stein[4]

[1]Department of Computer Science, University of Leipzig, Germany
[2]PRHLT Research Center, Universitat Politècnica de València, Spain
[3]Dept. of Information & Communication Systems Eng., University of the Aegean, Greece
[4]Web Technology & Information Systems, Bauhaus-Universität Weimar, Germany

pan@webis.de http://pan.webis.de

Abstract Digital text forensics aims at examining the originality and credibility of information in electronic documents and, in this regard, to extract and analyze information about the authors of these documents. The research field has been substantially developed during the last decade. PAN is a series of shared tasks that started in 2009 and significantly contributed to attract the attention of the research community in well-defined digital text forensics tasks. Several benchmark datasets have been developed to assess the state-of-the-art performance in a wide range of tasks. In this paper, we present the evolution of both the examined tasks and the developed datasets during the last decade. We also briefly introduce the upcoming PAN 2019 shared tasks.

1 Introduction

Digital Text Forensics is a text mining field examining authenticity and credibility issues of information included in electronic documents. It is closely related with text reuse and deception detection applications. But its main focus is on authorship analysis, aiming to reveal information about the author(s) of electronic documents. This is crucial in applications of cybersecurity, digital humanities, and social media analytics. Writing style, rather than topic information, is the primary factor in text forensics tasks [11].

PAN[1] is a series of shared tasks in digital text forensics, started in 2009, and held in conjunction with CLEF evaluation labs since 2010 [38,35]. During the last decade, PAN explored several text forensics tasks and attracted the attention of the international research community. A significant number of new evaluation datasets covering multiple languages and genres have been developed and quickly established as reference benchmarks in this area. Since 2013, only software submissions are allowed in PAN tasks and all submitted software is evaluated on the specifically developed TIRA experimentation platform [26]. Apart from enabling reproducibility of results, the collected software can easily be tested on alternative datasets. In this paper, we present the evolution of main tasks organized by PAN during the last decade depicted in Figure 1. In addition, we describe the datasets introduced by PAN to estimate the effectiveness and weaknesses of state-of-the-art methods.

[1] The acronym originates from the title of the first PAN workshop held at SIGIR-2007: Plagiarism analysis, Authorship identification, and Near-duplicate detection [36].

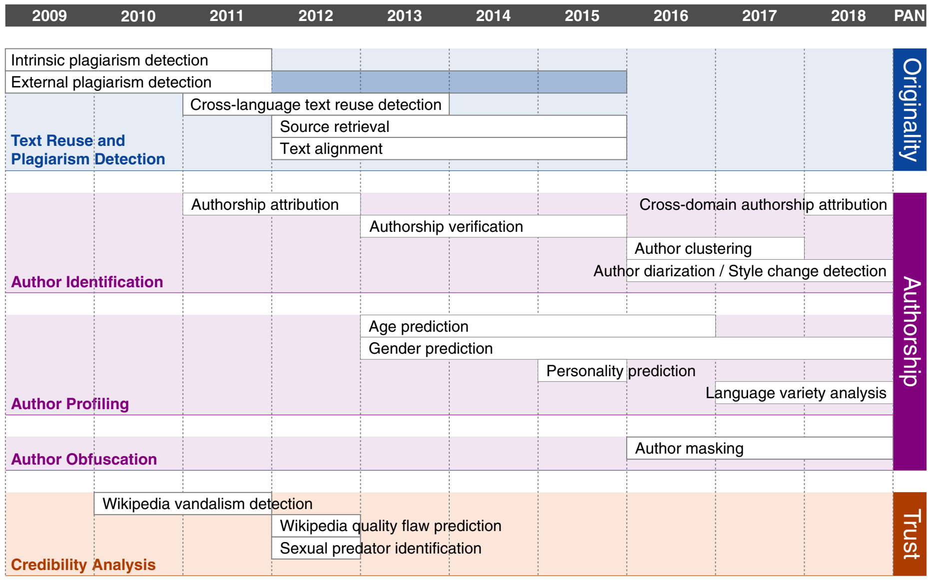

Figure 1. Development of the most important digital text forensics tasks at PAN, starting at 2009. The tasks address three aspects: originality (top), authorship (middle), and trust (bottom). For each aspect various tasks have been suggested, varied, and further specialized.

2 Plagiarism Detection

Plagiarism, the unacknowledged use of another author's original work, is considered a problem in publishing, science, and education. Texts and other works of art have been plagiarized throughout history, but with the advent of the World Wide Web, text reuse and plagiarism have been observed at large scale. Looking for theory, concepts, and algorithms to detect text reuse, computer-based plagiarism detection breaks down this task into manageable parts: *"Given a text d and a reference collection D, does d contain a section s for which one can find a document $d' \in D$ that contains a section s' such that under some retrieval model the similarity between s and s' is above a threshold?"*

The above definition presumes a closed world where a reference collection D is given, which is why this kind of analysis is called *external* plagiarism detection. Since D can be extremely large—possibly the entire indexed part of the World Wide Web—the respective research covers near-similarity search, near-duplicate detection, similarity hashing techniques, and indexes tailored to these problems In addition, situations where one would like to identify sections of plagiarized text if no reference collection is given can be imagined, a setting that is called *intrinsic* plagiarism detection. This problem is closely related to authorship verification: the goal of the former is to identify potential plagiarism by analyzing a document with respect to undeclared changes in writing style. In this regard, intrinsic plagiarism analysis can be understood as a more general form of the authorship verification problem: only a single document is given, and, one is faced with the problem of finding the suspicious sections. Both intrinsic plagiarism detection and authorship verification are one-class classification problems [37].

Against the above background the development of plagiarism detection tasks as shown in Figure 1 (top $\sim$ "originality") becomes plausible: starting 2009, both intrinsic and external plagiarism detection were considered; over three years, the evaluation datasets have been improved and extended [21,23,30]. This experience and the improved problem understanding is also reflected in development of tailored detection measures such as "pladget", which combines precision, recall, and detection accuracy for plagiarized passages. While cross-language text reuse detection lost its importance with gaining popularity of machine translation and the Wikipedia-Based Multilingual Retrieval Model [29], it became clear that research for external plagiarism detection requires a two-fold strategy, adopted in the ensuing three years [24,25,27]: (1) finding promising candidates on the Web (the source retrieval task), and, (2) developing effective algorithms for fuzzy text matching (the text alignment task). Meanwhile, spin-off tasks at FIRE [6,8], also in the form of source code reuse detection [9,10], used the original tasks' setup to develop resources for other languages

3 Author Identification

Author identification focuses on the personal style of the author(s) of electronic documents. The main assumption is that every author has her own stylistic "fingerprint" and that it is possible to identify the author(s) of a disputed document based on them [33]. There are several variations of this problem and PAN has explored many of them as shown in Figure 1. In more detail, in *closed-set authorship attribution*, a well-defined list of suspects and samples of texts they authored are provided. The task is to identify the most likely author of a questioned document among them. In *open-set attribution*, the true author may not be included in the list of suspects. The first editions of PAN related to author identification focused on tasks already popular in the research community [33]. In the 2011 edition, a dataset using emails (extracted from the Enron corpus) and relatively large sets of candidate authors was developed [4]. In 2012, emphasis was put on smaller candidate sets and fiction in English [15]. Another important task is *author verification* where there is only one candidate author. This is an especially challenging task considered fundamental in authorship attribution [17]. PAN has spurred widespread interest in this task among the research community, obtaining rather high participation figures in verification tasks from 2013 to 2015 [11,26,34]. The developed datasets for these tasks cover four languages (Dutch, English, Greek, Spanish) and a variety of genres (e.g., newspaper articles, student essays, reviews, novels, textbooks).

PAN also explored tasks where no labeled (known authorship) documents are provided. One such task is *author clustering* where the goal is to group documents written by the same author given a document collection. Two editions of PAN in 2016 and 2017 introduced an evaluation framework that also considers a retrieval task (ranking document pairs by likelihood of common authorship) [31,28]. Three languages (English, Greek, and Spanish) and two genres (reviews and newspaper articles) are included in the developed datasets focusing on either full texts (2016 edition) or fragments (paragraphs) of texts (2017 edition). Another unsupervised task is *author diarization*, where the assumption that each document is written by a single author does not hold. The task aims to determine how many authors wrote the document and extract the authorial components. A few variations of this task have been included in recent PAN editions, moving from complicated ones (e.g., detection of the exact number of co-authors and their exact contribution) [31,28], which proved to be extremely difficult at present, to more basic ones (e.g. *style change detection*: distinguishing between single-author and multi-author documents) [35], which is more feasible with current technology. The datasets to support these tasks include synthetic multi-author documents in English (essays or Q&As) where topic is controlled [31,28,35].

More recently, PAN focused on a challenging, but quite realistic problem: *cross-domain authorship attribution*. In this task, the labeled and unlabeled documents differ with respect to topic, genre, or even language. Fanfiction, a large part of contemporary fiction written by non-professionals following a canon (e.g., a well-known novel or TV series), has been adopted to allow for controlling the domain of documents. Thus the target domain (fandom) is excluded from the training documents in a closed-set attribution framework. The datasets built for this task include five languages (English, French, Italian, Polish, and Spanish) [35].

4 Author Profiling

Author profiling aims at identifying personal traits of an author on the basis of her writing. Traits, such as gender, age, language variety, or personality, are of high interest for areas such as forensics, security, and also marketing. From a forensic linguistics perspective, one would like to be able to know the linguistic profile of the author of a harassing text message (language used by a certain type of people). From a security perspective, these technologies may allow to profile and identify criminals. From the marketing perspective, being able to identify personal traits from comments to blogs or reviews may provide advertisers with the possibility of better segmenting their audience, which is an important competitive advantage. Traditional investigations in computational linguistics [5] and social psychology [20] have been carried out mainly for English. Furthermore, pioneering research from Argamon et al. [5] and Holmes et al. [13] focused on formal and well-written texts. With the rise of social media, however, the focus has shifted to more informal usage found in blogs and forums [16,32].

Starting in 2013, PAN has been organizing author profiling-related tasks with several objectives as depicted in Figure 1. We have covered different profiling aspects (age, gender, native language, language variety, personality), languages (Arabic, Dutch, English, Italian, Portuguese, Russian, Bengali, Hindi, Kannada, Malayo, Tamil, and Telugu), and genres (blogs, reviews, social media, and Twitter). The first edition was organized with the aim of investigating *age and gender identification* in a social media realistic scenario [11]. We collected thousands of social media posts in English and Spanish with a high variety of topics. With respect to age, we considered three classes following previous work by Schler et al. [32]: 10s (13-17), 20s (23-27) and 30s (33-47). Furthermore, we wanted to test the robustness of the systems when dealing with fake age profiles such as those induced by sexual predators. Therefore, we included texts from the previous year's shared task on sexual predator identification [14]. In the second edition [26], we extended the task to other genres besides social media focusing on Twitter, blogs, and hotel reviews, in English and Spanish. We realized the difficulty of obtaining high-quality labeled data and proposed a methodology to annotate age and gender. In 2014, we opted for modeling age classes without gaps: 18-24; 25-34; 35-49; 50-64; 65+. Finally, the Twitter sub-corpus was constructed in cooperation with RepLab [3] in order to address also the reputational perspective (e.g., profiling social media influencers, journalists, professionals, celebrities, among others).

In 2015 [34], besides age and gender identification, we introduced the task of *personality recognition* in Twitter. We maintained the age ranges defined in 2014 (except "50-64" and "65+" that were merged to "50-XX") and, besides English and Spanish, we included also Dutch and Italian (yet, only for gender and personality recognition). The objective of the shared task organized in 2016 [31] was to investigate the robustness of the systems in a cross-genre scenario. That is, training the systems in one genre and testing their performance in other genres. In particular, we provided Twitter data for training in English, Spanish, and Dutch. The approaches were then tested on blogs and social media genres in English and Spanish, and essays and reviews in Dutch. In 2017 [28], we introduced two novelties: *language variety identification* (together with gender), and Arabic and Portuguese languages (besides English and Spanish). This marked the first time a task has been organized covering gender and language variety identification combined. Language variety was addressed from a fine-grained and coarse-grained perspective, where varieties that are close, geographically, were grouped together (e.g., Canada and United States, Great Britain and Ireland, or New Zealand and Australia). Finally, in 2018 [35], gender identification on Twitter was approached from a multimodal perspective. Three languages have been considered: English, Spanish, and Arabic. Further spin-off profiling tasks were organized at FIRE [18,19].

5 Author Obfuscation

Author obfuscation (in particular, author masking as a special case) was launched in 2016 within the PAN task series: as the adversary task to authorship verification, it deals with preventing verification by altering a to-be-verified text. The underlying question is whether the authorial style of a text can be consistently manipulated. Though this task is of public interest and has various applications, only a handful of approaches have been proposed so far, and they achieved limited success only. We hope that this dedicated PAN task will push the research boundaries for both obfuscation and verification, and help to develop theoretical backgrounds and new evaluation frameworks: an obfuscation software is called *safe* if a forensic analysis does not reveal the original author of its obfuscated texts, it is called *sound* if its obfuscated texts are textually entailed with their originals, and it is called *sensible* if its obfuscated texts are inconspicuous.

6 Trust-Related Tasks

The PAN tasks related to trust (see Figure 1 bottom) have foreshadowed today's challenges that the Web and, in particular, social media platforms provide to computer scientists, psycholinguists, and psychologists, among others. Driven by the ideal of social responsibility and the scientific curiosity of the limits of "detectability", different tasks have been devised and operationalized.

Wikipedia vandalism detection (2010-2011) addressed the intentional damage of Wikipedia articles: given a set of edits on Wikipedia articles, the task was to separate ill-intentioned edits from well-intentioned edits. Wikipedia quality flaw prediction (2012) can be considered as a generalization of the vandalism detection task, focusing on the prediction of quality flaws in Wikipedia articles. It was driven by the observation that the majority of quality flaws in Wikipedia is not caused due to malicious intentions but stem from edits by inexperienced authors; examples include poor writing style, unreferenced statements, or missing neutrality. Since, by nature, no representative "negative" training data can be provided (articles that are tagged to not suffer from vandalism, articles that are tagged to not contain a particular flaw), vandalism detection and quality flaw prediction in Wikipedia represent one-class classification problems.

The goal of the sexual predator identification task (2012) was to identify online predators: the participants were given chat logs involving two (or more) people for which they had to determine who is the one trying to convince the other(s) to provide some sexual favor.

7 Discussion

During the last decade, PAN contributed to focus the attention of the research community on specific digital text forensics tasks, built benchmark datasets, and estimated the effectiveness as well as the weaknesses of the state of the art. The developed datasets cover multiple genres and languages while the top-ranked PAN submissions have been used as baselines in subsequent research [12,22]. In addition, the evolution of tasks within PAN made the exploration of new tasks feasible. For example, author obfuscation is based on the results of the author verification tasks. PAN also achieved to highlight the close relationship among certain tasks. For example, an approach to authorship clustering can be based on a verification method [7].

The upcoming edition of PAN will focus on four tasks. Two new tasks are introduced—*bots and gender profiling*, whose aim is to discriminate between human and robot Twitter profiles and in case of humans to profile their gender, and *celebrity profiling*, whose aim is to profile celebrities with regard to how they present themselves in public, be it personally or via public relations staff. In addition, the *cross-domain authorship attribution* task based on fanfiction documents, introduced in 2018, will continue. However, this time the open-set attribution scenario is adopted. Finally, another variant of the *style change detection* task will be included, this time focusing on the exact number of co-authors in a multi-author document.

Acknowledgements

We are indebted to many colleagues and friends who contributed greatly to PAN's tasks: Maik Anderka, Shlomo Argamon, Alberto Barrón-Cedeño, Fabio Celli, Fabio Crestani, Walter Daelemans, Andreas Eiselt, Tim Gollub, Parth Gupta, Matthias Hagen, Teresa Holfeld, Patrick Juola, Giacomo Inches, Mike Kestemont, Moshe Koppel, Manuel Montes-y-Gómez, Aurelio Lopez-Lopez, Francisco Rangel, Miguel Angel Sánchez-Pérez, Günther Specht, Michael Tschuggnall, and Ben Verhoeven. Our special thanks go to PAN's sponsors throughout the years and not least to the hundreds of participants.

References

1. FIRE 2015 Working Notes Papers, 4-6 December, Gandhinagar, India (2015), http://www.uni-weimar.de/medien/webis/events/pan-at-fire-15
2. FIRE 2017 Working Notes Papers, 8-11 December, Bangalore, India (2017)
3. Amigó, E., Carrillo de Albornoz, J., Chugur, I., Corujo, A., Gonzalo, J., Meij, E., de Rijke, M., Spina, D.: Overview of replab 2014: Author profiling and reputation dimensions for online reputation management. In: Information Access Evaluation. Multilinguality, Multimodality, and Interaction - 5th International Conference of the CLEF Initiative, CLEF 2014. Proceedings. pp. 307–322 (2014)
4. Argamon, S., Juola, P.: Overview of the International Authorship Identification Competition at PAN-2011. In: Petras, V., Forner, P., Clough, P. (eds.) Notebook Papers of CLEF 2011 Labs and Workshops, 19-22 September, Amsterdam, Netherlands (Sep 2011), http://www.clef-initiative.eu/publication/working-notes
5. Argamon, S., Koppel, M., Fine, J., Shimoni, A.R.: Gender, genre, and writing style in formal written texts. TEXT 23, 321–346 (2003)
6. Asghari, H., Mohtaj, S., Fatemi, O., Faili, H., Rosso, P., Potthast, M.: Algorithms and Corpora for Persian Plagiarism Detection: Overview of PAN at FIRE 2016. In: FIRE 2016 Working Notes Papers, 7-10 December, Kolkata, India (Dec 2016)
7. Bagnall, D.: Authorship Clustering Using Multi-headed Recurrent Neural Networks—Notebook for PAN at CLEF 2016. In: Balog, K., Cappellato, L., Ferro, N., Macdonald, C. (eds.) CLEF 2016 Evaluation Labs and Workshop – Working Notes Papers, 5-8 September, Évora, Portugal. CEUR Workshop Proceedings, CEUR-WS.org (Sep 2016), http://ceur-ws.org/Vol-1609/
8. Bensalem, I., Boukhalfa, I., Rosso, P., Abouenour, L., Darwish, K., Chikhi, S.: Overview of the AraPlagDet PAN@FIRE2015 Shared Task on Arabic Plagiarism Detection. In: FIRE 2015 Working Notes Papers, 4-6 December, Gandhinagar, India [1]
9. Flores, E., Rosso, P., Moreno, L., Villatoro-Tello, E.: On the Detection of SOurce COde Re-use. In: FIRE 2014 Working Notes Papers, 5-7 December, Bangalore, India. pp. 21–30 (Dec 2014)
10. Flores, E., Rosso, P., Villatoro-Tello, E., Moreno, L., Alcover, R., Chirivella, V.: PAN@FIRE: Overview of CL-SOCO Track on the Detection of Cross-Language SOurce COde Re-use. In: FIRE 2015 Working Notes Papers, 4-6 December, Gandhinagar, India [1], pp. 1–5
11. Gollub, T., Potthast, M., Beyer, A., Busse, M., Rangel, F., Rosso, P., Stamatatos, E., Stein, B.: Recent trends in digital text forensics and its evaluation. In: Forner, P., Müller, H., Paredes, R., Rosso, P., Stein, B. (eds.) Information Access Evaluation. Multilinguality, Multimodality, and Visualization. pp. 282–302. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
12. Halvani, O., Graner, L., Vogel, I.: Authorship verification in the absence of explicit features and thresholds. In: Advances in Information Retrieval - 40th European Conference on IR Research, ECIR 2018, Proceedings. pp. 454–465 (2018)
13. Holmes, J., Meyerhoff, M.: The Handbook of Language and Gender. Blackwell Handbooks in Linguistics, Wiley (2003)
14. Inches, G., Crestani, F.: Overview of the International Sexual Predator Identification Competition at PAN-2012. In: Forner, P., Karlgren, J., Womser-Hacker, C. (eds.) CLEF 2012 Evaluation Labs and Workshop – Working Notes Papers, 17-20 September, Rome, Italy (Sep 2012), http://www.clef-initiative.eu/publication/working-notes
15. Juola, P.: An Overview of the Traditional Authorship Attribution Subtask. In: Forner, P., Karlgren, J., Womser-Hacker, C. (eds.) CLEF 2012 Evaluation Labs and Workshop –

Working Notes Papers, 17-20 September, Rome, Italy (Sep 2012), http://www.clef-initiative.eu/publication/working-notes
16. Koppel, M., Argamon, S., Shimoni, A.R.: Automatically categorizing written texts by author gender (2003)
17. Koppel, M., Schler, J., Argamon, S., Winter, Y.: The "fundamental problem" of authorship attribution. English Studies 93(3), 284–291 (2012)
18. Litvinova, T., Rangel, F., Rosso, P., Seredin, P., Litvinova, O.: Overview of the RusProfiling PAN at FIRE Track on Cross-genre Gender Identification in Russian. In: FIRE 2017 Working Notes Papers, 8-11 December, Bangalore, India [2]
19. M, A.K., HB, B.G., Singh, S., KP, S., Rosso, P.: Overview of the INLI PAN at FIRE-2017 Track on Indian Native Language Identification. In: FIRE 2017 Working Notes Papers, 8-11 December, Bangalore, India [2]
20. Pennebaker, J.W.: The Secret Life of Pronouns: What Our Words Say About Us. Bloomsbury USA (2013)
21. Potthast, M., Barrón-Cedeño, A., Eiselt, A., Stein, B., Rosso, P.: Overview of the 2nd International Competition on Plagiarism Detection. In: Braschler, M., Harman, D., Pianta, E. (eds.) Working Notes Papers of the CLEF 2010 Evaluation Labs (Sep 2010), http://www.clef-initiative.eu/publication/working-notes
22. Potthast, M., Braun, S., Buz, T., Duffhauss, F., Friedrich, F., Gülzow, J.M., Köhler, J., Lötzsch, W., Müller, F., Müller, M.E., Paßmann, R., Reinke, B., Rettenmeier, L., Rometsch, T., Sommer, T., Träger, M., Wilhelm, S., Stein, B., Stamatatos, E., Hagen, M.: Who wrote the web? revisiting influential author identification research applicable to information retrieval. In: Advances in Information Retrieval - 38th European Conference on IR Research, ECIR 2016, Proceedings. pp. 393–407 (2016)
23. Potthast, M., Eiselt, A., Barrón-Cedeño, A., Stein, B., Rosso, P.: Overview of the 3rd International Competition on Plagiarism Detection. In: Notebook Papers of the 5th Evaluation Lab on Uncovering Plagiarism, Authorship and Social Software Misuse (PAN). Amsterdam, The Netherlands (September 2011)
24. Potthast, M., Gollub, T., Hagen, M., Graßegger, J., Kiesel, J., Michel, M., Oberländer, A., Tippmann, M., Barrón-Cedeño, A., Gupta, P., Rosso, P., Stein, B.: Overview of the 4th International Competition on Plagiarism Detection. In: Forner, P., Karlgren, J., Womser-Hacker, C. (eds.) Working Notes Papers of the CLEF 2012 Evaluation Labs (Sep 2012), http://www.clef-initiative.eu/publication/working-notes
25. Potthast, M., Gollub, T., Hagen, M., Tippmann, M., Kiesel, J., Rosso, P., Stamatatos, E., Stein, B.: Overview of the 5th International Competition on Plagiarism Detection. In: Forner, P., Navigli, R., Tufis, D. (eds.) Working Notes Papers of the CLEF 2013 Evaluation Labs (Sep 2013), http://www.clef-initiative.eu/publication/working-notes
26. Potthast, M., Gollub, T., Rangel, F., Rosso, P., Stamatatos, E., Stein, B.: Improving the Reproducibility of PAN's Shared Tasks: Plagiarism Detection, Author Identification, and Author Profiling. In: Kanoulas, E., Lupu, M., Clough, P., Sanderson, M., Hall, M., Hanbury, A., Toms, E. (eds.) Information Access Evaluation meets Multilinguality, Multimodality, and Visualization. 5th International Conference of the CLEF Initiative (CLEF 14). pp. 268–299. Springer, Berlin Heidelberg New York (Sep 2014)
27. Potthast, M., Hagen, M., Beyer, A., Busse, M., Tippmann, M., Rosso, P., Stein, B.: Overview of the 6th International Competition on Plagiarism Detection. In: Cappellato, L., Ferro, N., Halvey, M., Kraaij, W. (eds.) Working Notes Papers of the CLEF 2014 Evaluation Labs. CEUR Workshop Proceedings, CLEF and CEUR-WS.org (Sep 2014), http://www.clef-initiative.eu/publication/working-notes
28. Potthast, M., Rangel, F., Tschuggnall, M., Stamatatos, E., Rosso, P., Stein, B.: Overview of pan'17: Author identification, author profiling, and author obfuscation. In: Jones, G.J.,

Lawless, S., Gonzalo, J., Kelly, L., Goeuriot, L., Mandl, T., Cappellato, L., Ferro, N. (eds.) Experimental IR Meets Multilinguality, Multimodality, and Interaction. pp. 275–290. Springer International Publishing, Cham (2017)
29. Potthast, M., Stein, B., Anderka, M.: A Wikipedia-Based Multilingual Retrieval Model. In: Macdonald, C., Ounis, I., Plachouras, V., Ruthven, I., White, R. (eds.) Advances in Information Retrieval. 30th European Conference on IR Research (ECIR 2008). Lecture Notes in Computer Science, vol. 4956, pp. 522–530. Springer, Berlin Heidelberg New York (2008)
30. Potthast, M., Stein, B., Eiselt, A., Barrón-Cedeño, A., Rosso, P.: Overview of the 1st International Competition on Plagiarism Detection. In: Stein, B., Rosso, P., Stamatatos, E., Koppel, M., Agirre, E. (eds.) SEPLN 2009 Workshop on Uncovering Plagiarism, Authorship, and Social Software Misuse (PAN 2009). pp. 1–9. CEUR-WS.org (Sep 2009), http://ceur-ws.org/Vol-502
31. Rosso, P., Rangel, F., Potthast, M., Stamatatos, E., Tschuggnall, M., Stein, B.: Overview of pan'16: New challenges for authorship analysis: Cross-genre profiling, clustering, diarization, and obfuscation. In: Fuhr, N., Quaresma, P., Gonçalves, T., Larsen, B., Balog, K., Macdonald, C., Cappellato, L., Ferro, N. (eds.) Experimental IR Meets Multilinguality, Multimodality, and Interaction. pp. 332–350. Springer International Publishing, Cham (2016)
32. Schler, J., Koppel, M., Argamon, S., Pennebaker, J.W.: Effects of age and gender on blogging. In: AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs. pp. 199–205. AAAI (2006)
33. Stamatatos, E.: A Survey of Modern Authorship Attribution Methods. Journal of the American Society for Information Science and Technology 60, 538–556 (2009)
34. Stamatatos, E., Potthast, M., Rangel, F., Rosso, P., Stein, B.: Overview of the pan/clef 2015 evaluation lab. In: Mothe, J., Savoy, J., Kamps, J., Pinel-Sauvagnat, K., Jones, G., San Juan, E., Capellato, L., Ferro, N. (eds.) Experimental IR Meets Multilinguality, Multimodality, and Interaction. pp. 518–538. Springer International Publishing, Cham (2015)
35. Stamatatos, E., Rangel, F., Tschuggnall, M., Stein, B., Kestemont, M., Rosso, P., Potthast, M.: Overview of pan 2018: Author identification, author profiling, and author obfuscation. In: Bellot, P., Trabelsi, C., Mothe, J., Murtagh, F., Nie, J.Y., Soulier, L., SanJuan, E., Cappellato, L., Ferro, N. (eds.) Experimental IR Meets Multilinguality, Multimodality, and Interaction. pp. 267–285. Springer International Publishing, Cham (2018)
36. Stein, B., Koppel, M., Stamatatos, E. (eds.): SIGIR 07 Workshop Workshop on Plagiarism Analysis, Authorship Identification, and Near-Duplicate Detection (PAN 07). CEUR-WS.org (2007), http://www.uni-weimar.de/medien/webis/events/pan-07
37. Stein, B., Lipka, N., Prettenhofer, P.: Intrinsic Plagiarism Analysis. Language Resources and Evaluation (LRE) 45(1), 63–82 (Mar 2011)
38. Stein, B., Rosso, P., Stamatatos, E., Koppel, M., Agirre, E. (eds.): SEPLN 2009 Workshop on Uncovering Plagiarism, Authorship, and Social Software Misuse (PAN 09). Universidad Politécnica de Valencia and CEUR-WS.org (2009), http://ceur-ws.org/Vol-502