Efstathios Stamatatos

List of Publications by Citations

73 papers	2,185 citations	21 h-index	45 g-index
74 ext. papers	2,618 ext. citations	1.6 avg, IF	5.75 L-index

#	Paper	IF	Citations
73	A survey of modern authorship attribution methods. *Journal of the Association for Information Science and Technology*, **2009**, 60, 538-556		669
72	Automatic Text Categorization in Terms of Genre and Author. *Computational Linguistics*, **2000**, 26, 471-495	2.8	187
71	Syntactic N-grams as machine learning features for natural language processing. *Expert Systems With Applications*, **2014**, 41, 853-860	7.8	139
70	Authorship Attribution for Social Media Forensics. *IEEE Transactions on Information Forensics and Security*, **2017**, 12, 5-33	8	92
69	Computer-Based Authorship Attribution Without Lexical Measures. *Computers and the Humanities*, **2001**, 35, 193-214		87
68	Author identification: Using text sampling to handle the class imbalance problem. *Information Processing and Management*, **2008**, 44, 790-799	6.3	76
67	Plagiarism detection using stopword n-grams. *Journal of the Association for Information Science and Technology*, **2011**, 62, 2512-2527		63
66	N-Gram Feature Selection for Authorship Identification. *Lecture Notes in Computer Science*, **2006**, 77-86	0.9	62
65	WORDS VERSUS CHARACTER N-GRAMS FOR ANTI-SPAM FILTERING. *International Journal on Artificial Intelligence Tools*, **2007**, 16, 1047-1067	0.9	59
64	Effective identification of source code authors using byte-level information **2006**,		47
63	AUTHORSHIP ATTRIBUTION BASED ON FEATURE SET SUBSPACING ENSEMBLES. *International Journal on Artificial Intelligence Tools*, **2006**, 15, 823-838	0.9	37
62	Authorship Attribution Using Text Distortion **2017**,		34
61	Automatic identification of music performers with learning ensembles. *Artificial Intelligence*, **2005**, 165, 37-56	3.6	30
60	Discriminative subprofile-specific representations for author profiling in social media. *Knowledge-Based Systems*, **2015**, 89, 134-147	7.3	29
59	Examining the significance of high-level programming features in source code author classification. *Journal of Systems and Software*, **2008**, 81, 447-460	3.3	28
58	Author Identification Using Imbalanced and Limited Training Texts **2007**,		27
57	Syntactic Dependency-Based N-grams as Classification Features. *Lecture Notes in Computer Science*, **2013**, 1-11	0.9	26

56	Source Code Author Identification Based on N-gram Author Profiles **2006**, 508-515		26
55	Improving the quality of degraded document images		25
54	Learning to recognize webpage genres. *Information Processing and Management*, **2009**, 45, 499-512	6.3	24
53	Overview of the PAN/CLEF 2015 Evaluation Lab. *Lecture Notes in Computer Science*, **2015**, 518-538	0.9	23
52	Overview of PAN17. *Lecture Notes in Computer Science*, **2017**, 275-290	0.9	20
51	Overview of PAN 2019: Bots and Gender Profiling, Celebrity Profiling, Cross-Domain Authorship Attribution and Style Change Detection. *Lecture Notes in Computer Science*, **2019**, 402-416	0.9	20
50	A Profile-Based Method for Authorship Verification. *Lecture Notes in Computer Science*, **2014**, 313-326	0.9	20
49	Plagiarism analysis, authorship identification, and near-duplicate detection PAN'07. *ACM SIGIR Forum*, **2007**, 41, 68-71	0.9	18
48	Improving the Reproducibility of PAN's Shared Tasks:. *Lecture Notes in Computer Science*, **2014**, 268-299	0.9	17
47	Who Wrote the Web? Revisiting Influential Author Identification Research Applicable to Information Retrieval. *Lecture Notes in Computer Science*, **2016**, 393-407	0.9	17
46	Improving author verification based on topic modeling. *Journal of the Association for Information Science and Technology*, **2019**, 70, 1074-1088	2.7	16
45	Plagiarism and authorship analysis: introduction to the special issue. *Language Resources and Evaluation*, **2011**, 45, 1-4	1.8	15
44	Webpage Genre Identification Using Variable-Length Character n-Grams **2007**,		14
43	Overview of PAN16. *Lecture Notes in Computer Science*, **2016**, 332-350	0.9	12
42	Spam Detection Using Character N-Grams. *Lecture Notes in Computer Science*, **2006**, 95-104	0.9	12
41	Recent Trends in Digital Text Forensics and Its Evaluation. *Lecture Notes in Computer Science*, **2013**, 282-302	0.9	12
40	Masking topic-related information to enhance authorship attribution. *Journal of the Association for Information Science and Technology*, **2018**, 69, 461-473	2.7	12
39	Open-Set Classification for Automated Genre Identification. *Lecture Notes in Computer Science*, **2013**, 207-217	0.9	11

38	Improving Cross-Topic Authorship Attribution: The Role of Pre-Processing. *Lecture Notes in Computer Science*, **2018**, 289-302	0.9	11
37	Syntactic Dependency-Based N-grams: More Evidence of Usefulness in Classification. *Lecture Notes in Computer Science*, **2013**, 13-24	0.9	11
36	Intrinsic Author Verification Using Topic Modeling **2018**,		10
35	Paraphrase plagiarism identification with character-level features. *Pattern Analysis and Applications*, **2019**, 22, 669-681	2.3	10
34	Masking domain-specific information for cross-domain deception detection. *Pattern Recognition Letters*, **2020**, 135, 122-130	4.7	9
33	Cross-Domain Authorship Attribution Using Pre-trained Language Models. *IFIP Advances in Information and Communication Technology*, **2020**, 255-266	0.5	9
32	An Improved Impostors Method for Authorship Verification. *Lecture Notes in Computer Science*, **2017**, 138-144	0.9	9
31	Tensor Space Models for Authorship Identification. *Lecture Notes in Computer Science*, **2008**, 239-249	0.9	9
30	Author Identification Using Imbalanced and Limited Training Texts		8
29	Authorship Verification: A Review of Recent Advances. *Research in Computing Science*, **2016**, 123, 9-25	1.2	7
28	Distinguishing the Popularity between Topics: A System for Up-to-Date Opinion Retrieval and Mining in the Web. *Lecture Notes in Computer Science*, **2013**, 197-209	0.9	7
27	An Agent-Based Focused Crawling Framework for Topic- and Genre-Related Web Document Discovery **2012**,		6
26	Fourth international workshop on uncovering plagiarism, authorship, and social software misuse. *ACM SIGIR Forum*, **2011**, 45, 45-48	0.9	6
25	Source Code Authorship Analysis For Supporting the Cybercrime Investigation Process. *Advances in Digital Crime, Forensics, and Cyber Terrorism*, **2010**, 470-495	0.2	6
24	Improved algorithms for extrinsic author verification. *Knowledge and Information Systems*, **2020**, 62, 1903-19216	2.4	6
23	Overview of PAN 2018. *Lecture Notes in Computer Science*, **2018**, 267-285	0.9	6
22	Plagiarism detection based on structural information **2011**,		5
21	Extracting informative textual parts from web pages containing user-generated content **2012**,		5

20	Overview of PAN 2021: Authorship Verification, Profiling Hate Speech Spreaders on Twitter, and Style Change Detection. *Lecture Notes in Computer Science*, **2021**, 419-431	0.9	5
19	The Impact of Noise in Web Genre Identification. *Lecture Notes in Computer Science*, **2015**, 268-273	0.9	3
18	Author Identification in Imbalanced Sets of Source Code Samples **2012**,		3
17	Supporting multilinguality in library automation systems using ai tools. *Applied Artificial Intelligence* , **1999**, 13, 679-703	2.3	3
16	Shared Tasks on Authorship Analysis at PAN 2020. *Lecture Notes in Computer Science*, **2020**, 508-516	0.9	3
15	Identification of Plagiarism Using Syntactic and Semantic Filters. *Lecture Notes in Computer Science*, **2014**, 495-506	0.9	3
14	Open set evaluation of web genre identification. *Language Resources and Evaluation*, **2018**, 52, 949-968	1.8	3
13	Dynamic Ensemble Selection for Author Verification. *Lecture Notes in Computer Science*, **2019**, 102-115	0.9	2
12	A Decade of Shared Tasks in Digital Text Forensics at PAN. *Lecture Notes in Computer Science*, **2019**, 291-300	0.9	2
11	An image processing self-training system for ruling line removal algorithms **2013**,		2
10	Learning How to Propagate Using Random Probing. *Lecture Notes in Computer Science*, **2009**, 263-278	0.9	2
9	Overview of PAN 2021: Authorship Verification, Profiling Hate Speech Spreaders on Twitter, and Style Change Detection. *Lecture Notes in Computer Science*, **2021**, 567-573	0.9	2
8	Devising Rhesus: A strange collaboration between Aeschylus and Euripides. *Digital Scholarship in the Humanities*, **2018**, 33, 347-361	0.6	2
7	Open-Set Web Genre Identification Using Distributional Features and Nearest Neighbors Distance Ratio. *Lecture Notes in Computer Science*, **2019**, 3-11	0.9	1
6	Evolution of the PAN Lab on Digital Text Forensics. *The Kluwer International Series on Information Retrieval*, **2019**, 461-485	0.7	1
5	Supporting the Cybercrime Investigation Process: Effective Discrimination of Source Code Authors Based on Byte-Level Information. *Communications in Computer and Information Science*, **2007**, 163-173	0.3	1
4	Universality of Stylistic Traits in Texts. *Lecture Notes in Morphogenesis*, **2016**, 143-155	0	1
3	A transfer learning approach to cross-domain authorship attribution. *Evolving Systems*, **2021**, 12, 625-643	2.1	0

2	Overview of PAN 2022: Authorship Verification, Profiling Irony and Stereotype Spreaders, Style Change Detection, and Trigger Detection. *Lecture Notes in Computer Science*, **2022**, 331-338	0.9	0
1	Music Performer Verification Based on Learning Ensembles. *Lecture Notes in Computer Science*, **2004**, 122-131	0.9	