

# Sophia Ananiadou

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/4357915/publications.pdf>

Version: 2024-02-01

161  
papers

6,886  
citations

66343

42  
h-index

79698

73  
g-index

171  
all docs

171  
docs citations

171  
times ranked

5595  
citing authors

#	ARTICLE	IF	CITATIONS
1	Automatic recognition of multi-word terms: the C-value/NC-value method. International Journal on Digital Libraries, 2000, 3, 115-130.	1.5	542
2	Using text mining for study identification in systematic reviews: a systematic review of current approaches. Systematic Reviews, 2015, 4, 5.	5.3	345
3	Text mining and its potential applications in systems biology. Trends in Biotechnology, 2006, 24, 571-579.	9.3	281
4	Text mining and ontologies in biomedicine: Making sense of raw text. Briefings in Bioinformatics, 2005, 6, 239-251.	6.5	245
5	Event extraction for systems biology by text mining the literature. Trends in Biotechnology, 2010, 28, 381-390.	9.3	160
6	FACTA: a text search engine for finding associated biomedical concepts. Bioinformatics, 2008, 24, 2559-2560.	4.1	152
7	Applications of text mining within systematic reviews. Research Synthesis Methods, 2011, 2, 1-14.	8.7	146
8	A Neural Layered Model for Nested Named Entity Recognition. , 2018, , .		145
9	Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts. Journal of Biomedical Informatics, 2016, 62, 148-158.	4.3	140
10	The C-value/NC-value domain-independent method for multi-word term extraction. Journal of Natural Language Processing, 1999, 6, 145-179.	0.2	135
11	Inter-sentence Relation Extraction with Document-level Graph Convolutional Neural Network. , 2019, , .		134
12	Reducing systematic review workload through certainty-based screening. Journal of Biomedical Informatics, 2014, 51, 242-253.	4.3	116
13	Event extraction across multiple levels of biological organization. Bioinformatics, 2012, 28, i575-i581.	4.1	107
14	Connecting the Dots: Document-level Neural Relation Extraction with Edge-oriented Graphs. , 2019, , .		107
15	Europe PMC: a full-text literature database for the life sciences and platform for innovation. Nucleic Acids Research, 2015, 43, D1042-D1048.	14.5	104
16	Discovering and visualizing indirect associations between biomedical concepts. Bioinformatics, 2011, 27, i111-i119.	4.1	101
17	Boosting automatic event extraction from the literature using domain adaptation and coreference resolution. Bioinformatics, 2012, 28, 1759-1765.	4.1	95
18	Supporting Systematic Reviews Using Text Mining. Social Science Computer Review, 2009, 27, 509-523.	4.2	90

#	ARTICLE	IF	CITATIONS
19	Machine learning algorithms for systematic review: reducing workload in a preclinical review of animal studies and reducing human screening error. <i>Systematic Reviews</i> , 2019, 8, 23.	5.3	90
20	Large-Scale Event Extraction from Literature with Multi-Level Gene Normalization. <i>PLoS ONE</i> , 2013, 8, e55814.	2.5	83
21	Learning string similarity measures for gene/protein name dictionary look-up using logistic regression. <i>Bioinformatics</i> , 2007, 23, 2768-2774.	4.1	81
22	U-Compare: share and compare text mining tools with UIMA. <i>Bioinformatics</i> , 2009, 25, 1997-1998.	4.1	81
23	Natural language processing applied to mental illness detection: a narrative review. <i>Npj Digital Medicine</i> , 2022, 5, 46.	10.9	78
24	Generating Natural Language specifications from UML class diagrams. <i>Requirements Engineering</i> , 2008, 13, 1-18.	3.1	77
25	Prioritising references for systematic reviews with RobotAnalyst: A user study. <i>Research Synthesis Methods</i> , 2018, 9, 470-488.	8.7	77
26	Construction of an annotated corpus to support biomedical information extraction. <i>BMC Bioinformatics</i> , 2009, 10, 349.	2.6	73
27	Adverse drug events and medication relation extraction in electronic health records with ensemble deep learning methods. <i>Journal of the American Medical Informatics Association: JAMIA</i> , 2020, 27, 39-46.	4.4	72
28	Building an abbreviation dictionary using a term recognition approach. <i>Bioinformatics</i> , 2006, 22, 3089-3095.	4.1	71
29	Topic detection using paragraph vectors to support active learning in systematic reviews. <i>Journal of Biomedical Informatics</i> , 2016, 62, 59-65.	4.3	67
30	A Walk-based Model on Entity Graphs for Relation Extraction. , 2018, , .		60
31	BioCause: Annotating and analysing causality in the biomedical domain. <i>BMC Bioinformatics</i> , 2013, 14, 2.	2.6	58
32	Event-based text mining for biology and functional genomics. <i>Briefings in Functional Genomics</i> , 2015, 14, 213-230.	2.7	58
33	Enriching a biomedical event corpus with meta-knowledge annotation. <i>BMC Bioinformatics</i> , 2011, 12, 393.	2.6	57
34	Disambiguating the species of biomedical named entities using natural language parsers. <i>Bioinformatics</i> , 2010, 26, 661-667.	4.1	55
35	Argo: an integrative, interactive, text mining-based workbench supporting curation. <i>Database: the Journal of Biological Databases and Curation</i> , 2012, 2012, bas010-bas010.	3.0	55
36	Extracting semantically enriched events from biomedical literature. <i>BMC Bioinformatics</i> , 2012, 13, 108.	2.6	52

#	ARTICLE	IF	CITATIONS
37	Negated bio-events: analysis and identification. BMC Bioinformatics, 2013, 14, 14.	2.6	52
38	Identification of research hypotheses and new knowledge from scientific literature. BMC Medical Informatics and Decision Making, 2018, 18, 46.	3.0	52
39	Overview of the ID, EPI and REL tasks of BioNLP Shared Task 2011. BMC Bioinformatics, 2012, 13, S2.	2.6	51
40	Anatomical entity mention recognition at literature scale. Bioinformatics, 2014, 30, 868-875.	4.1	49
41	PathText: a text mining integrator for biological pathway visualizations. Bioinformatics, 2010, 26, i374-i381.	4.1	47
42	Text Mining the History of Medicine. PLoS ONE, 2016, 11, e0144717.	2.5	47
43	Overview of the Cancer Genetics and Pathway Curation tasks of BioNLP Shared Task 2013. BMC Bioinformatics, 2015, 16, S2.	2.6	44
44	Text mining resources for the life sciences. Database: the Journal of Biological Databases and Curation, 2016, 2016, .	3.0	44
45	Thalia: semantic search engine for biomedical abstracts. Bioinformatics, 2019, 35, 1799-1801.	4.1	43
46	Building a high-quality sense inventory for improved abbreviation disambiguation. Bioinformatics, 2010, 26, 1246-1253.	4.1	41
47	The BioLexicon: a large-scale terminological resource for biomedical text mining. BMC Bioinformatics, 2011, 12, 397.	2.6	41
48	An ensemble of neural models for nested adverse drug events and medication extraction with subwords. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 22-30.	4.4	41
49	UKPMC: a full text article resource for the life sciences. Nucleic Acids Research, 2011, 39, D58-D65.	14.5	40
50	Supporting systematic reviews using LDA-based document representations. Systematic Reviews, 2015, 4, 172.	5.3	40
51	Boosting drug named entity recognition using an aggregate classifier. Artificial Intelligence in Medicine, 2015, 65, 145-153.	6.5	40
52	Mining metabolites: extracting the yeast metabolome from the literature. Metabolomics, 2011, 7, 94-101.	3.0	37
53	Normalizing biomedical terms by minimizing ambiguity and variability. BMC Bioinformatics, 2008, 9, S2.	2.6	36
54	Overview of the interactive task in BioCreative V. Database: the Journal of Biological Databases and Curation, 2016, 2016, baw119.	3.0	36

#	ARTICLE	IF	CITATIONS
55	Wide coverage biomedical event extraction using multiple partially overlapping corpora. BMC Bioinformatics, 2013, 14, 175.	2.6	35
56	DeepEventMine: end-to-end neural nested event extraction from biomedical texts. Bioinformatics, 2020, 36, 4910-4917.	4.1	35
57	A method for integrating and ranking the evidence for biochemical pathways by mining reactions from text. Bioinformatics, 2013, 29, i44-i52.	4.1	34
58	Enriching news events with meta-knowledge information. Language Resources and Evaluation, 2017, 51, 409-438.	2.7	33
59	Terminology-driven literature mining and knowledge acquisition in biomedicine. International Journal of Medical Informatics, 2002, 67, 33-48.	3.3	32
60	Enhancing Search: Events and Their Discourse Context. Lecture Notes in Computer Science, 2013, , 318-334.	1.3	32
61	A semi-supervised approach using label propagation to support citation screening. Journal of Biomedical Informatics, 2017, 72, 67-76.	4.3	31
62	biochem4j: Integrated and extensible biochemical knowledge through graph databases. PLoS ONE, 2017, 12, e0179130.	2.5	31
63	Detecting experimental techniques and selecting relevant documents for protein-protein interactions from biomedical literature. BMC Bioinformatics, 2011, 12, S11.	2.6	30
64	How to make the most of NE dictionaries in statistical NER. BMC Bioinformatics, 2008, 9, S5.	2.6	29
65	A mental state Knowledge-aware and Contrastive Network for early stress and depression detection on social media. Information Processing and Management, 2022, 59, 102961.	8.6	29
66	Proximity-Based Frameworks for Generating Embeddings from Multi-Output Data. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34, 2216-2232.	13.9	28
67	BioCreative V BioC track overview: collaborative biocurator assistant task for BioGRID. Database: the Journal of Biological Databases and Curation, 2016, 2016, baw121.	3.0	28
68	Using text mining techniques to extract phenotypic information from the PhenoCHF corpus. BMC Medical Informatics and Decision Making, 2015, 15, S3.	3.0	26
69	Risk of bias reporting in the recent animal focal cerebral ischaemia literature. Clinical Science, 2017, 131, 2525-2532.	4.3	26
70	TRUCKS: A Model for Automatic Multi-Word Term Recognition.. Journal of Natural Language Processing, 2001, 8, 101-125.	0.2	25
71	Optimising chemical named entity recognition with pre-processing analytics, knowledge-rich features and heuristics. Journal of Cheminformatics, 2015, 7, S6.	6.1	24
72	Using uncertainty to link and rank evidence from biomedical literature for model curation. Bioinformatics, 2017, 33, 3784-3792.	4.1	23

#	ARTICLE	IF	CITATIONS
73	Improving reference prioritisation with PICO recognition. BMC Medical Informatics and Decision Making, 2019, 19, 256.	3.0	22
74	Facilitating the Analysis of Discourse Phenomena in an Interoperable NLP Platform. Lecture Notes in Computer Science, 2013, , 559-571.	1.3	22
75	Text-mining-assisted biocuration workflows in Argo. Database: the Journal of Biological Databases and Curation, 2014, 2014, .	3.0	21
76	Biosimilar vs originator insulins: Systematic review and meta-analysis. Diabetes, Obesity and Metabolism, 2018, 20, 1787-1792.	4.4	21
77	Annotation and detection of drug effects in text for pharmacovigilance. Journal of Cheminformatics, 2018, 10, 37.	6.1	21
78	SciLite: a platform for displaying text-mined annotations as a means to link research articles with biological data. Wellcome Open Research, 2016, 1, 25.	1.8	21
79	Customised OCR correction for historical medical text. , 2015, , .		20
80	Thesaurus or Logical Ontology, Which One Do We Need for Text Mining?. Computers and the Humanities, 2005, 39, 77-90.	1.4	19
81	Divide-and-Conquer: Post-User Interaction Network for Fake News Detection on Social Media. , 2022, , .		19
82	U-Compare: A modular NLP workflow construction and evaluation system. IBM Journal of Research and Development, 2011, 55, 11:1-11:10.	3.1	18
83	Using Workflows to Explore and Optimise Named Entity Recognition for Chemistry. PLoS ONE, 2011, 6, e20181.	2.5	18
84	ASCOT: a text mining-based web-service for efficient search and assisted creation of clinical trials. BMC Medical Informatics and Decision Making, 2012, 12, S3.	3.0	18
85	Proactive Learning for Named Entity Recognition. , 2017, , .		18
86	Adaptable, high recall, event extraction system with minimal configuration. BMC Bioinformatics, 2015, 16, S7.	2.6	17
87	A Text Mining Pipeline Using Active and Deep Learning Aimed at Curating Information in Computational Neuroscience. Neuroinformatics, 2019, 17, 391-406.	2.8	17
88	Automatic identification of suicide notes with a transformer-based deep learning model. Internet Interventions, 2021, 25, 100422.	2.7	17
89	Named Entity Recognition for Bacterial Type IV Secretion Systems. PLoS ONE, 2011, 6, e14780.	2.5	17
90	Text mining meets workflow: linking U-Compare with Taverna. Bioinformatics, 2010, 26, 2486-2487.	4.1	16

#	ARTICLE	IF	CITATIONS
91	Semi-supervised learning of causal relations in biomedical scientific discourse. <i>BioMedical Engineering OnLine</i> , 2014, 13, S1.	2.7	16
92	Supporting the annotation of chronic obstructive pulmonary disease (COPD) phenotypes with text mining workflows. <i>Journal of Biomedical Semantics</i> , 2015, 6, 8.	1.6	16
93	Descriptive document clustering via discriminant learning in a co-embedded space of multilevel similarities. <i>Journal of the Association for Information Science and Technology</i> , 2016, 67, 106-133.	2.9	16
94	Argo: enabling the development of bespoke workflows and services for disease annotation. <i>Database: the Journal of Biological Databases and Curation</i> , 2016, 2016, baw066.	3.0	15
95	Cited text span identification for scientific summarisation using pre-trained encoders. <i>Scientometrics</i> , 2020, 125, 3109-3137.	3.0	15
96	The pace of modern culture. <i>Nature Human Behaviour</i> , 2020, 4, 352-360.	12.0	14
97	Improving classification of Adverse Drug Reactions through Using Sentiment Analysis and Transfer Learning. , 2019, , .		14
98	COPIOUS: A gold standard corpus of named entities towards extracting species occurrence from biodiversity literature. <i>Biodiversity Data Journal</i> , 2019, 7, e29626.	0.8	14
99	Accelerating the annotation of sparse named entities by dynamic sentence selection. <i>BMC Bioinformatics</i> , 2008, 9, S8.	2.6	13
100	Syntactically-informed word representations from graph neural network. <i>Neurocomputing</i> , 2020, 413, 431-443.	5.9	13
101	Automatic extraction of angiogenesis bioprocess from text. <i>Bioinformatics</i> , 2011, 27, 2730-2737.	4.1	12
102	News search using discourse analytics. , 2013, , .		12
103	Processing biological literature with customizable Web services supporting interoperable formats. <i>Database: the Journal of Biological Databases and Curation</i> , 2014, 2014, bau064-bau064.	3.0	12
104	Constructing a biodiversity terminological inventory. <i>PLoS ONE</i> , 2017, 12, e0175277.	2.5	12
105	Supporting the education evidence portal via text mining. <i>Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences</i> , 2010, 368, 3829-3844.	3.4	11
106	Text Mining Supporting Search for Knowledge Discovery in Diabetes. <i>Current Cardiovascular Risk Reports</i> , 2013, 7, 1-8.	2.0	11
107	Mapping Phenotypic Information in Heterogeneous Textual Sources to a Domain-Specific Terminological Resource. <i>PLoS ONE</i> , 2016, 11, e0162287.	2.5	11
108	U-Compare bio-event meta-service: compatible BioNLP event extraction services. <i>BMC Bioinformatics</i> , 2011, 12, 481.	2.6	10

#	ARTICLE	IF	CITATIONS
109	Anatomical Entity Recognition with a Hierarchical Framework Augmented by External Resources. PLoS ONE, 2014, 9, e108396.	2.5	10
110	Inferring appropriate eligibility criteria in clinical trial protocols without labeled data. , 2012, , .		9
111	Clinical text classification under the Open and Closed Topic Assumptions. International Journal of Data Mining and Bioinformatics, 2009, 3, 299.	0.1	8
112	Bilingual term alignment from comparable corpora in English discharge summary and Chinese discharge summary. BMC Bioinformatics, 2015, 16, 149.	2.6	8
113	Self-Tuned Descriptive Document Clustering Using a Predictive Network. IEEE Transactions on Knowledge and Data Engineering, 2018, 30, 1929-1942.	5.7	8
114	Data Visualization with Structural Control of Global Cohort and Local Data Neighborhoods. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40, 1323-1337.	13.9	8
115	Combining String and Context Similarity for Bilingual Term Alignment from Comparable Corpora. , 2014, , .		8
116	Neighbour Interaction based Click-Through Rate Prediction via Graph-masked Transformer. , 2022, , .		8
117	Multi-topic Aspects in Clinical Text Classification. , 2007, , .		7
118	Deploying and sharing U-Compare workflows as web services. Journal of Biomedical Semantics, 2013, 4, 7.	1.6	7
119	Womenâ€™s health in<i>The BMJ</i>: a data science history. BMJ Open, 2020, 10, e039759.	1.9	7
120	A Text Mining-Based Framework for Constructing an RDF-Compliant Biodiversity Knowledge Repository. Communications in Computer and Information Science, 2017, , 30-42.	0.5	7
121	THE VALUE OF AN IN-DOMAIN LEXICON IN GENOMICS QA. Journal of Bioinformatics and Computational Biology, 2010, 08, 147-161.	0.8	6
122	A Deep Learning Approach to Refine the Identification of High-Quality Clinical Research Articles From the Biomedical Literature: Protocol for Algorithm Development and Validation. JMIR Research Protocols, 2021, 10, e29398.	1.0	6
123	Paladin: an annotation tool based on active and proactive learning. , 2021, , .		6
124	Terms are not alone: term choice and choice terms. ASLIB Proceedings, 1995, 47, 47-60.	1.2	5
125	RECOGNISING DISCOURSE CAUSALITY TRIGGERS IN THE BIOMEDICAL DOMAIN. Journal of Bioinformatics and Computational Biology, 2013, 11, 1343008.	0.8	5
126	Hypothesis, analysis and synthesis, it's all Greek to me. ELife, 2019, 8, .	6.0	5



#	ARTICLE	IF	CITATIONS
127	A Narrative Literature Review of Natural Language Processing Applied to the Occupational Exposome. International Journal of Environmental Research and Public Health, 2022, 19, 8544.	2.6	5
128	Highly scalable Text Mining - parallel tagging application. , 2009, , .		4
129	A Cross-Lingual Similarity Measure for Detecting Biomedical Term Translations. PLoS ONE, 2015, 10, e0126196.	2.5	4
130	LitPathExplorer: a confidence-based visual text analytics tool for exploring literature-enriched pathway models. Bioinformatics, 2018, 34, 1389-1397.	4.1	4
131	Annotating and detecting phenotypic information for chronic obstructive pulmonary disease. JAMIA Open, 2019, 2, 261-271.	2.0	4
132	Improving Textual Emotion Recognition Based on Intra- and Inter-Class Variations. IEEE Transactions on Affective Computing, 2023, 14, 1297-1307.	8.3	4
133	The GENIA Corpus: Annotation Levels and Applications. , 2017, , 1395-1432.		4
134	APLenty: annotation tool for creating high-quality datasets using active and proactive learning. , 2018, , .		4
135	A method for discovering and inferring appropriate eligibility criteria in clinical trial protocols without labeled data. BMC Medical Informatics and Decision Making, 2013, 13, S6.	3.0	3
136	Dealing with Data Sparsity in Drug Named Entity Recognition. , 2013, , .		3
137	A Hybrid Approach to Compiling Bilingual Dictionaries of Medical Terms from Parallel Corpora. Lecture Notes in Computer Science, 2014, , 57-69.	1.3	3
138	The strategic impact of META-NET on the regional, national and international level. Language Resources and Evaluation, 2016, 50, 351-374.	2.7	3
139	Quantifying risk factors in medical reports with a context-aware linear model. Journal of the American Medical Informatics Association: JAMIA, 2019, 26, 537-546.	4.4	3
140	On revolutions. Palgrave Communications, 2020, 6, .	4.7	3
141	NERO: a biomedical named-entity (recognition) ontology with a large, annotated corpus reveals meaningful associations through text embedding. Npj Systems Biology and Applications, 2021, 7, 38.	3.0	3
142	A Text Mining Framework for Accelerating the Semantic Curation of Literature. Lecture Notes in Computer Science, 2016, , 459-462.	1.3	3
143	Generalising semantic category disambiguation with large lexical resources for fun and profit. Journal of Biomedical Semantics, 2014, 5, 26.	1.6	2
144	Semantically enhanced search system for historical medical archives. , 2015, , .		2

#	ARTICLE	IF	CITATIONS
145	Supporting Biological Pathway Curation Through Text Mining. Communications in Computer and Information Science, 2017, , 59-73.	0.5	2
146	Mapping anatomical related entities to human body parts based on wikipedia in discharge summaries. BMC Bioinformatics, 2019, 20, 430.	2.6	2
147	Developing Multilingual Text Mining Workflows in UIMA and U-Compare. Lecture Notes in Computer Science, 2012, , 82-93.	1.3	2
148	Comparing neural models for nested and overlapping biomedical event detection. BMC Bioinformatics, 2022, 23, .	2.6	2
149	EXTRACTING SECONDARY BIO-EVENT ARGUMENTS WITH EXTRACTION CONSTRAINTS. Computational Intelligence, 2011, 27, 702-721.	3.2	1
150	"Mining events from the literature for bioinformatics applications" by S. Ananiadou, P. Thompson, and R. Nawaz; with Martin Vesely as coordinator. SIGWEB Newsletter: the Newsletter of ACM's Special Interest Group on Hypertext and Hypermedia, 2013, , 1-12.	0.6	1
151	Adding text mining workflows as web services to the BioCatalogue. , 2012, , .		1
152	Argo as a platform for integrating distinct biodiversity analytics tools into workflows for building graph databases. Biodiversity Information Science and Standards, 0, 1, e20067.	0.0	1
153	A term-based and citation network-based search system for COVID-19. JAMIA Open, 2021, 4, ooab104.	2.0	1
154	AGRA: analysis of gene ranking algorithms. Bioinformatics, 2011, 27, 1185-1186.	4.1	0
155	ASCOT. , 2012, , .		0
156	A hybrid approach to recognising discourse causality in the biomedical domain. , 2013, , .		0
157	DISCOVERING ROBUST EMBEDDINGS IN (DIS)SIMILARITY SPACE FOR HIGH-DIMENSIONAL LINGUISTIC FEATURES. Computational Intelligence, 2014, 30, 285-315.	3.2	0
158	Role of in silico tools and text mining in the safety assessment of selected plant coumarins. Toxicology Letters, 2017, 280, S96.	0.8	0
159	HSEarch: Semantic Search System for Workplace Accident Reports. Lecture Notes in Computer Science, 2021, , 514-519.	1.3	0
160	Towards a Linguistic Treatment of Compounds in a Machine Translation Environment. Journal of Natural Language Processing, 1996, 3, 45-66.	0.2	0
161	On Revolutions. SSRN Electronic Journal, 0, , .	0.4	0