Hongfang Liu

List of Publications by Year
in descending order

Source: https://exaly.com/author-pdf/8411794/publications.pdf

Version: 2024-02-01

256	6,637	87888	102487
papers	citations	38 h-index	66 g-index

294	294	294	7694
all docs	docs citations	times ranked	citing authors

#	Article	IF	Citations
1	Clinical information extraction applications: A literature review. Journal of Biomedical Informatics, 2018, 77, 34-49.	4.3	502
2	A clinical text classification paradigm using weak supervision and deep representation. BMC Medical Informatics and Decision Making, 2019, 19, 1.	3.0	348
3	A comparison of word embeddings for the biomedical natural language processing. Journal of Biomedical Informatics, 2018, 87, 12-20.	4.3	259
4	CLAMP â€“ a toolkit for efficiently building customized clinical natural language processing pipelines. Journal of the American Medical Informatics Association: JAMIA, 2018, 25, 331-336.	4.4	257
5	Validating drug repurposing signals using electronic health records: a case study of metformin associated with reduced cancer mortality. Journal of the American Medical Informatics Association: JAMIA, 2015, 22, 179-191.	4.4	178
6	Deep learning and alternative learning strategies for retrospective real-world clinical data. Npj Digital Medicine, 2019, 2, 43.	10.9	145
7	Artificial intelligence approaches using natural language processing to advance EHR-based clinical research. Journal of Allergy and Clinical Immunology, 2020, 145, 463-469.	2.9	142
8	Characterizing Long COVID: Deep Phenotype of a Complex Condition. EBioMedicine, 2021, 74, 103722.	6.1	127
9	Using machine learning for concept extraction on clinical documents from multiple data sources. Journal of the American Medical Informatics Association: JAMIA, 2011, 18, 580-587.	4.4	112
10	BioThesaurus: a web-based thesaurus of protein and gene names. Bioinformatics, 2006, 22, 103-105.	4.1	106
11	Deep Phenotyping on Electronic Health Records Facilitates Genetic Diagnosis by Clinical Exomes. American Journal of Human Genetics, 2018, 103, 58-73.	6.2	99
12	Rapid identification of familial hypercholesterolemia from electronic health records: The SEARCH study. Journal of Clinical Lipidology, 2016, 10, 1230-1239.	1.5	98
13	MedXN: an open source medication extraction and normalization tool for clinical text. Journal of the American Medical Informatics Association: JAMIA, 2014, 21, 858-865.	4.4	88
14	Systematic Analysis of Adverse Event Reports for Sex Differences in Adverse Drug Events. Scientific Reports, 2016, 6, 24955.	3.3	88
15	Clinical concept extraction: A methodology review. Journal of Biomedical Informatics, 2020, 109, 103526.	4.3	86
16	MedSTS: a resource for clinical semantic textual similarity. Language Resources and Evaluation, 2020, 54, 57-72.	2.7	81
17	DEEPEN: A negation detection system for clinical text incorporating dependency relation into NegEx. Journal of Biomedical Informatics, 2015, 54, 213-219.	4.3	79
18	Natural language processing of clinical notes for identification of critical limb ischemia. International Journal of Medical Informatics, 2018, 111, 83-89.	3.3	77

#	Article	IF	Citations
19	A Multi-aspect Comparison Study of Supervised Word Sense Disambiguation. Journal of the American Medical Informatics Association: JAMIA, 2004, 11, 320-331.	4.4	76
20	An information extraction framework for cohort identification using electronic health records. AMIA Summits on Translational Science Proceedings, 2013, 2013, 149-53.	0.4	76
21	Mining peripheral arterial disease cases from narrative clinical notes using natural language processing. Journal of Vascular Surgery, 2017, 65, 1753-1761.	1.1	75
22	Desiderata for delivering NLP to accelerate healthcare AI advancement and a Mayo Clinic NLP-as-a-service implementation. Npj Digital Medicine, 2019, 2, 130.	10.9	70
23	Application of a Natural Language Processing Algorithm to Asthma Ascertainment. An Automated Chart Review. American Journal of Respiratory and Critical Care Medicine, 2017, 196, 430-437.	5.6	67
24	AffyProbeMiner: a web resource for computing or retrieving accurately redefined Affymetrix probe sets. Bioinformatics, 2007, 23, 2385-2390.	4.1	66
25	Automated chart review for asthma cohort identification using natural language processing: an exploratory study. Annals of Allergy, Asthma and Immunology, 2013, 111, 364-369.	1.0	63
26	Unified Medical Language System term occurrences in clinical notes: a large-scale corpus analysis. Journal of the American Medical Informatics Association: JAMIA, 2012, 19, e149-e156.	4.4	60
27	Toward a Learning Health-care System â€“ Knowledge Delivery at the Point of Care Empowered by Big Data and NLP. Biomedical Informatics Insights, 2016, 8s1, BII.S37977.	4.6	56
28	Unsupervised machine learning for the discovery of latent disease clusters and patient subgroups using electronic health records. Journal of Biomedical Informatics, 2020, 102, 103364.	4.3	56
29	BioTagger-GM: A Gene/Protein Name Recognition System. Journal of the American Medical Informatics Association: JAMIA, 2009, 16, 247-255.	4.4	54
30	Clinical documentation variations and NLP system portability: a case study in asthma birth cohorts across institutions. Journal of the American Medical Informatics Association: JAMIA, 2018, 25, 353-359.	4.4	52
31	Correlations Between COVID-19 Cases and Google Trends Data in the United States: A State-by-State Analysis. Mayo Clinic Proceedings, 2020, 95, 2370-2381.	3.0	52
32	Automated chart review utilizing natural language processing algorithm for asthma predictive index. BMC Pulmonary Medicine, 2018, 18, 34.	2.0	51
33	Developing a FHIR-based EHR phenotyping framework: A case study for identification of patients with obesity and multiple comorbidities from discharge summaries. Journal of Biomedical Informatics, 2019, 99, 103310.	4.3	48
34	A common type system for clinical natural language processing. Journal of Biomedical Semantics, 2013, 4, 1.	1.6	47
35	Recommendations for the safe, effective use of adaptive CDS in the US healthcare system: an AMIA position paper. Journal of the American Medical Informatics Association: JAMIA, 2021, 28, 677-684.	4.4	46
36	Microarray probes and probe sets. Frontiers in Bioscience - Elite, 2010, E2, 325-338.	1.8	45

#	Article	IF	Citations
37	Overview of the gene ontology task at BioCreative IV. Database: the Journal of Biological Databases and Curation, 2014, 2014, bau086-bau086.	3.0	45
38	A Robust e-Epidemiology Tool in Phenotyping Heart Failure with Differentiation for Preserved and Reduced Ejection Fraction: the Electronic Medical Records and Genomics (eMERGE) Network. Journal of Cardiovascular Translational Research, 2015, 8, 475-483.	2.4	44
39	Detection of clinically important colorectal surgical site infection using Bayesian network. Journal of Surgical Research, 2017, 209, 168-173.	1.6	42
40	iProLINK: an integrated protein resource for literature mining. Computational Biology and Chemistry, 2004, 28, 409-416.	2.3	40
41	Natural Language Processing for Asthma Ascertainment in Different Practice Settings. Journal of Allergy and Clinical Immunology: in Practice, 2018, 6, 126-131.	3.8	40
42	Natural Language Processing for the Identification of Silent Brain Infarcts From Neuroimaging Reports. JMIR Medical Informatics, 2019, 7, e12109.	2.6	40
43	Modeling asynchronous event sequences with RNNs. Journal of Biomedical Informatics, 2018, 83, 167-177.	4.3	39
44	Peripheral Neutrophil to Lymphocyte Ratio Improves Prognostication in Colon Cancer. Clinical Colorectal Cancer, 2017, 16, 115-123.e3.	2.3	38
45	BELMiner: adapting a rule-based relation extraction system to extract biological expression language statements from bio-medical literature evidence sentences. Database: the Journal of Biological Databases and Curation, 2017, 2017, .	3.0	38
46	SpliceCenter: A suite of web-based bioinformatic applications for evaluating the impact of alternative splicing on RT-PCR, RNAi, microarray, and peptide-based studies. BMC Bioinformatics, 2008, 9, 313.	2.6	36
47	A study of transportability of an existing smoking status detection module across institutions. AMIA ... Annual Symposium proceedings, 2012, 2012, 577-86.	0.2	36
48	DynGO: a tool for visualizing and mining of Gene Ontology and its associations. BMC Bioinformatics, 2005, 6, 201.	2.6	35
49	Developing a scalable FHIR-based clinical data normalization pipeline for standardizing and integrating unstructured and structured electronic health record data. JAMIA Open, 2019, 2, 570-579.	2.0	35
50	Valx: A System for Extracting and Structuring Numeric Lab Test Comparison Statements from Text. Methods of Information in Medicine, 2016, 55, 266-275.	1.2	34
51	Proton Pump Inhibitors and the Risk for Fracture at Specific Sites: Data Mining of the FDA Adverse Event Reporting System. Scientific Reports, 2017, 7, 5527.	3.3	34
52	BioCreative/OHNLP Challenge 2018. , 2018, , .		34
53	Cohort Profile: The Right Drug, Right Dose, Right Time: Using Genomic Data to Individualize Treatment Protocol (RIGHT Protocol). International Journal of Epidemiology, 2020, 49, 23-24k.	1.9	34
54	CancerBERT: a cancer domain-specific language model for extracting breast cancer phenotypes from electronic health records. Journal of the American Medical Informatics Association: JAMIA, 2022, 29, 1208-1216.	4.4	33

#	Article	IF	Citations
55	Ease of adoption of clinical natural language processing software: An evaluation of five systems. Journal of Biomedical Informatics, 2015, 58, S189-S196.	4.3	32
56	Standardizing adverse drug event reporting data. Journal of Biomedical Semantics, 2014, 5, 36.	1.6	31
57	Using large clinical corpora for query expansion in text-based cohort identification. Journal of Biomedical Informatics, 2014, 49, 275-281.	4.3	31
58	Comparison of Three Information Sources for Smoking Information in Electronic Health Records. Cancer Informatics, 2016, 15, CIN.S40604.	1.9	30
59	A Part-Of-Speech term weighting scheme for biomedical information retrieval. Journal of Biomedical Informatics, 2016, 63, 379-389.	4.3	30
60	Postoperative bleeding risk prediction for patients undergoing colorectal surgery. Surgery, 2018, 164, 1209-1216.	1.9	30
61	Rare disease knowledge enrichment through a data-driven approach. BMC Medical Informatics and Decision Making, 2019, 19, 32.	3.0	30
62	The 2019 n2c2/OHNLP Track on Clinical Semantic Textual Similarity: Overview. JMIR Medical Informatics, 2020, 8, e23375.	2.6	30
63	Developing an FHIR-Based Computational Pipeline for Automatic Population of Case Report Forms for Colorectal Cancer Clinical Trials Using Electronic Health Records. JCO Clinical Cancer Informatics, 2020, 4, 201-209.	2.1	28
64	Implementation of preemptive DNA sequenceâ€“based pharmacogenomics testing across a large academic medical center: The Mayo-Baylor RIGHT 10K Study. Genetics in Medicine, 2022, 24, 1062-1072.	2.4	28
65	Extracting chemicalâ€“protein relations using attention-based neural networks. Database: the Journal of Biological Databases and Curation, 2018, 2018, .	3.0	27
66	Assessing Unmet Information Needs of Breast Cancer Survivors: Exploratory Study of Online Health Forums Using Text Classification and Retrieval. JMIR Cancer, 2018, 4, e10.	2.4	27
67	HPO2Vec+: Leveraging heterogeneous knowledge resources to enrich node embeddings for the Human Phenotype Ontology. Journal of Biomedical Informatics, 2019, 96, 103246.	4.3	26
68	Natural language processing of radiology reports for identification of skeletal site-specific fractures. BMC Medical Informatics and Decision Making, 2019, 19, 73.	3.0	26
69	Assessment of the impact of EHR heterogeneity for clinical research through a case study of silent brain infarction. BMC Medical Informatics and Decision Making, 2020, 20, 60.	3.0	26
70	Utilization of Electronic Medical Records and Biomedical Literature to Support the Diagnosis of Rare Diseases Using Data Fusion and Collaborative Filtering Approaches. JMIR Medical Informatics, 2018, 6, e11301.	2.6	26
71	Association of Silent Cerebrovascular Disease Identified Using Natural Language Processing and Future Ischemic Stroke. Neurology, 2021, 97, e1313-e1321.	1.1	25
72	Artificial intelligence-assisted clinical decision support for childhood asthma management: A randomized clinical trial. PLoS ONE, 2021, 16, e0255261.	2.5	25

#	Article	IF	Citations
73	Using Social Media Data to Identify Potential Candidates for Drug Repurposing: A Feasibility Study. JMIR Research Protocols, 2016, 5, e121.	1.0	25
74	Network-based analysis reveals distinct association patterns in a semantic MEDLINE-based drug-disease-gene network. Journal of Biomedical Semantics, 2014, 5, 33.	1.6	24
75	A Hybrid Approach to Sentiment Sentence Classification in Suicide Notes. Biomedical Informatics Insights, 2012, 5s1, BII.S8961.	4.6	23
76	Text mining facilitates database curation - extraction of mutation-disease associations from Bio-medical literature. BMC Bioinformatics, 2015, 16, 185.	2.6	23
77	Clinical element models in the SHARPn consortium. Journal of the American Medical Informatics Association: JAMIA, 2016, 23, 248-256.	4.4	23
78	Early genetic aberrations in patients with sporadic colorectal cancer. Molecular Carcinogenesis, 2018, 57, 114-124.	2.7	23
79	Detecting Pharmacovigilance Signals Combining Electronic Medical Records With Spontaneous Reports: A Case Study of Conventional Disease-Modifying Antirheumatic Drugs for Rheumatoid Arthritis. Frontiers in Pharmacology, 2018, 9, 875.	3.5	23
80	A new method for prioritizing drug repositioning candidates extracted by literature-based discovery. , 2015, , .		22
81	Testing the theory of relative defect proneness for closed-source software. Empirical Software Engineering, 2010, 15, 577-598.	3.9	21
82	Association of Ankle-Brachial Indices With Limb Revascularization or Amputation in Patients With Peripheral Artery Disease. JAMA Network Open, 2018, 1, e185547.	5.9	21
83	Automated extraction of sudden cardiac death risk factors in hypertrophic cardiomyopathy patients by natural language processing. International Journal of Medical Informatics, 2019, 128, 32-38.	3.3	21
84	MayoNLP at SemEval-2016 Task 1: Semantic Textual Similarity based on Lexical Semantic Net and Deep Learning Semantic Model. , 2016, , .		21
85	Expert artificial intelligence-based natural language processing characterises childhood asthma. BMJ Open Respiratory Research, 2020, 7, e000524.	3.0	20
86	Identification of Patients with Family History of Pancreatic Cancer--Investigation of an NLP System Portability. Studies in Health Technology and Informatics, 2015, 216, 604-8.	0.3	20
87	Ascertainment of asthma prognosis using natural language processing from electronic medical records. Journal of Allergy and Clinical Immunology, 2018, 141, 2292-2294.e3.	2.9	19
88	Ensembles of natural language processing systems for portable phenotyping solutions. Journal of Biomedical Informatics, 2019, 100, 103318.	4.3	19
89	A corpus-driven standardization framework for encoding clinical problems with HL7 FHIR. Journal of Biomedical Informatics, 2020, 110, 103541.	4.3	19
90	Mining severe drug-drug interaction adverse events using Semantic Web technologies: a case study. BioData Mining, 2015, 8, 12.	4.0	18

#	Article	IF	Citations
91	Automatic extraction and assessment of lifestyle exposures for Alzheimerâ€™s disease using natural language processing. International Journal of Medical Informatics, 2019, 130, 103943.	3.3	18
92	Ascertainment of Delirium Status Using Natural Language Processing From Electronic Health Records. Journals of Gerontology - Series A Biological Sciences and Medical Sciences, 2022, 77, 524-530.	3.6	18
93	Patient-level temporal aggregation for text-based asthma status ascertainment. Journal of the American Medical Informatics Association: JAMIA, 2014, 21, 876-884.	4.4	17
94	Need of informatics in designing interoperable clinical registries. International Journal of Medical Informatics, 2017, 108, 78-84.	3.3	17
95	Electronic Health Record Phenotypes for Precision Medicine: Perspectives and Caveats From Treatment of Breast Cancer at a Single Institution. Clinical and Translational Science, 2018, 11, 85-92.	3.1	17
96	Constructing co-occurrence network embeddings to assist association extraction for COVID-19 and other coronavirus infectious diseases. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 1259-1267.	4.4	17
97	Modeling cancer clinical trials using HL7 FHIR to support downstream applications: A case study with colorectal cancer data. International Journal of Medical Informatics, 2021, 145, 104308.	3.3	17
98	Are synthetic clinical notes useful for real natural language processing tasks: A case study on clinical entity recognition. Journal of the American Medical Informatics Association: JAMIA, 2021, 28, 2193-2201.	4.4	17
99	Detecting Lifestyle Risk Factors for Chronic Kidney Disease With Comorbidities: Association Rule Mining Analysis of Web-Based Survey Data. Journal of Medical Internet Research, 2019, 21, e14204.	4.3	17
100	Identifying Abdominal Aortic Aneurysm Cases and Controls using Natural Language Processing of Radiology Reports. AMIA Summits on Translational Science Proceedings, 2013, 2013, 249-53.	0.4	17
101	Identifying peripheral arterial disease cases using natural language processing of clinical notes. , 2016, 2016, 126-131.		16
102	Leveraging word embeddings and medical entity extraction for biomedical dataset retrieval using unstructured texts. Database: the Journal of Biological Databases and Curation, 2017, 2017, .	3.0	16
103	Natural language processing for populating lung cancer clinical research data. BMC Medical Informatics and Decision Making, 2019, 19, 239.	3.0	16
104	Time event ontology (TEO): to support semantic representation and reasoning of complex temporal relations of clinical events. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 1046-1056.	4.4	16
105	Implementation of a Cohort Retrieval System for Clinical Data Repositories Using the Observational Medical Outcomes Partnership Common Data Model: Proof-of-Concept System Validation. JMIR Medical Informatics, 2020, 8, e17376.	2.6	16
106	Patient Portal Messaging for Asynchronous Virtual Care During the COVID-19 Pandemic: Retrospective Analysis. JMIR Human Factors, 2022, 9, e35187.	2.0	16
107	Computational drug repurposing based on electronic health records: a scoping review. Npj Digital Medicine, 2022, 5, .	10.9	16
108	Ontology-based systematical representation and drug class effect analysis of package insert-reported adverse events associated with cardiovascular drugs used in China. Scientific Reports, 2017, 7, 13819.	3.3	15

#	Article	IF	Citations
109	Dependency and AMR Embeddings for Drug-Drug Interaction Extraction from Biomedical Literature. , 2017, , .		15
110	Systematic identification of latent disease-gene associations from PubMed articles. PLoS ONE, 2018, 13, e0191568.	2.5	15
111	Innovative Informatics Approaches for Peripheral Artery Disease: Current State and Provider Survey of Strategies for Improving Guideline-Based Care. Mayo Clinic Proceedings Innovations, Quality & Outcomes, 2018, 2, 129-136.	2.4	14
112	Family History Extraction From Synthetic Clinical Narratives Using Natural Language Processing: Overview and Evaluation of a Challenge Data Set and Solutions for the 2019 National NLP Clinical Challenges (n2c2)/Open Health Natural Language Processing (OHNLP) Competition. JMIR Medical Informatics, 2021, 9, e24008.	2.6	14
113	Natural Language Processing and Machine Learning for Identifying Incident Stroke From Electronic Health Records: Algorithm Development and Validation. Journal of Medical Internet Research, 2021, 23, e22951.	4.3	14
114	Longitudinal cohorts for harnessing the electronic health record for disease prediction in a US population. BMJ Open, 2021, 11, e044353.	1.9	14
115	Phenotypic Analysis of Clinical Narratives Using Human Phenotype Ontology. Studies in Health Technology and Informatics, 2017, 245, 581-585.	0.3	14
116	Building a knowledge base of severe adverse drug events based on AERS reporting data using semantic web technologies. Studies in Health Technology and Informatics, 2013, 192, 496-500.	0.3	14
117	Software Engineering Education for Bioinformatics. , 2009, , .		13
118	Ten-Year Trends in Antiemetic Prescribing in Patients Receiving Highly Emetogenic Chemotherapy. Journal of the National Comprehensive Cancer Network: JNCCN, 2018, 16, 294-299.	4.9	13
119	A privacy-preserving distributed filtering framework for NLP artifacts. BMC Medical Informatics and Decision Making, 2019, 19, 183.	3.0	13
120	Predicate Oriented Pattern Analysis for Biomedical Knowledge Discovery. Intelligent Information Management, 2016, 08, 66-85.	0.5	13
121	BELTracker: evidence sentence retrieval for BEL statements. Database: the Journal of Biological Databases and Curation, 2016, 2016, baw079.	3.0	12
122	Burden of hospitalization in clinically diagnosed peripheral artery disease: A community-based study. Vascular Medicine, 2018, 23, 23-31.	1.5	12
123	Selected articles from the BioCreative/OHNLP challenge 2018. BMC Medical Informatics and Decision Making, 2019, 19, 262.	3.0	12
124	COVID-19 TestNorm: A tool to normalize COVID-19 testing names to LOINC codes. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 1437-1442.	4.4	12
125	An aberration detection-based approach for sentinel syndromic surveillance of COVID-19 and other novel influenza-like illnesses. Journal of Biomedical Informatics, 2021, 113, 103660.	4.3	12
126	Privacy-Preserving Predictive Modeling: Harmonization of Contextual Embeddings From Different Sources. JMIR Medical Informatics, 2018, 6, e33.	2.6	12

#	Article	IF	Citations
127	Systematic Analysis of Free-Text Family History in Electronic Health Record. AMIA Summits on Translational Science Proceedings, 2017, 2017, 104-113.	0.4	12
128	Deep Learning Solutions for Classifying Patients on Opioid Use. AMIA ... Annual Symposium proceedings, 2017, 2017, 525-534.	0.2	12
129	Use of echocardiography in outpatients with chest pain and normal resting electrocardiograms referred to Mayo Clinic Rochester. American Heart Journal, 2018, 196, 49-55.	2.7	11
130	Leveraging Collaborative Filtering to Accelerate Rare Disease Diagnosis. AMIA ... Annual Symposium proceedings, 2017, 2017, 1554-1563.	0.2	11
131	Analysis of Cross-Institutional Medication Description Patterns in Clinical Narratives. Biomedical Informatics Insights, 2013, 6s1, BII.S11634.	4.6	10
132	Prioritizing Adverse Drug Reaction and Drug Repositioning Candidates Generated by Literature-Based Discovery. , 2016, , .		10
133	Time Lapse to Colorectal Cancer: Telomere Dynamics Define the Malignant Potential of Polyps. Clinical and Translational Gastroenterology, 2016, 7, e188.	2.5	10
134	Test collections for electronic health record-based clinical information retrieval. JAMIA Open, 2019, 2, 360-368.	2.0	10
135	Agreement between neuroimages and reports for natural language processing-based detection of silent brain infarcts and white matter disease. BMC Neurology, 2021, 21, 189.	1.8	10
136	Using ensemble of ensemble machine learning methods to predict outcomes of cardiac resynchronization. Journal of Cardiovascular Electrophysiology, 2021, 32, 2504-2514.	1.7	10
137	Towards a semantic lexicon for clinical natural language processing. AMIA ... Annual Symposium proceedings, 2012, 2012, 568-76.	0.2	10
138	Integrating Structured and Unstructured EHR Data Using an FHIR-based Type System: A Case Study with Medication Data. AMIA Summits on Translational Science Proceedings, 2018, 2017, 74-83.	0.4	10
139	A hybrid model to identify fall occurrence from electronic health records. International Journal of Medical Informatics, 2022, 162, 104736.	3.3	10
140	An integrative computational approach to identify disease-specific networks from PubMed literature information. , 2013, , .		9
141	Colorectal Cancer with Residual Polyp of Origin: A Model of Malignant Transformation. Translational Oncology, 2016, 9, 280-286.	3.7	9
142	Impact of Patient Reminders on Papanicolaou Test Completion for High-Risk Patients Identified by a Clinical Decision Support System. Journal of Women's Health, 2018, 27, 569-574.	3.3	9
143	Using data-driven sublanguage pattern mining to induce knowledge models: application in medical image reports knowledge representation. BMC Medical Informatics and Decision Making, 2018, 18, 61.	3.0	9
144	Deployment of an Interdisciplinary Predictive Analytics Task Force to Inform Hospital Operational Decision-Making During the COVID-19 Pandemic. Mayo Clinic Proceedings, 2021, 96, 690-698.	3.0	9

#	Article	IF	Citations
145	Quantifying the Importance of COVID-19 Vaccination to Our Future Outlook. Mayo Clinic Proceedings, 2021, 96, 1890-1895.	3.0	9
146	Recommending Education Materials for Diabetic Questions Using Information Retrieval Approaches. Journal of Medical Internet Research, 2017, 19, e342.	4.3	9
147	Using SNOMED-CT to encode summary level data - a corpus analysis. AMIA Summits on Translational Science Proceedings, 2012, 2012, 30-7.	0.4	9
148	Discovering adverse drug events combining spontaneous reports with electronic medical records: a case study of conventional DMARDs and biologics for rheumatoid arthritis. AMIA Summits on Translational Science Proceedings, 2017, 2017, 95-103.	0.4	9
149	CliniViewer: a tool for viewing electronic medical records based on natural language processing and XML. Studies in Health Technology and Informatics, 2004, 107, 639-43.	0.3	9
150	Next generation informatics for big data in precision medicine era. BioData Mining, 2015, 8, 34.	4.0	8
151	Intrainstitutional EHR collections for patientâ€level information retrieval. Journal of the Association for Information Science and Technology, 2017, 68, 2636-2648.	2.9	8
152	Early Identification of Childhood Asthma: The Role of Informatics in an Era of Electronic Health Records. Frontiers in Pediatrics, 2019, 7, 113.	1.9	8
153	Detection of Surgical Site Infection Utilizing Automated Feature Generation in Clinical Notes. Journal of Healthcare Informatics Research, 2019, 3, 267-282.	7.6	8
154	Feasibility of pooling annotated corpora for clinical concept extraction. AMIA Summits on Translational Science Proceedings, 2012, 2012, 38.	0.4	8
155	Coverage Evaluation of CTCAE for Capturing the Immune-related Adverse Events Leveraging Text Mining Technologies. AMIA Summits on Translational Science Proceedings, 2019, 2019, 771-778.	0.4	8
156	MedTator: a serverless annotation tool for corpus development. Bioinformatics, 2022, 38, 1776-1778.	4.1	8
157	Semi-supervised Learning of Text Classification on Bacterial Protein-Protein Interaction Documents. , 2009, , .		7
158	A frequency-filtering strategy of obtaining PHI-free sentences from clinical data repository. , 2015, , .		7
159	BmQGen: Biomedical query generator for knowledge discovery. , 2015, , .		7
160	Preliminary exploration of survival analysis using the OHDSI common data model: a case study of intrahepatic cholangiocarcinoma. BMC Medical Informatics and Decision Making, 2018, 18, 116.	3.0	7
161	Salience of Medical Concepts of Inside Clinical Texts and Outside Medical Records for Referred Cardiovascular Patients. Journal of Healthcare Informatics Research, 2019, 3, 200-219.	7.6	7
162	Evaluation of patient-level retrieval from electronic health record data for a cohort discovery task. JAMIA Open, 2020, 3, 395-404.	2.0	7

#	Article	IF	Citations
163	Generating real-world evidence from unstructured clinical notes to examine clinical utility of genetic tests: use case in BRCAness. BMC Medical Informatics and Decision Making, 2021, 21, 3.	3.0	7
164	Using Human Phenotype Ontology for Phenotypic Analysis of Clinical Notes. Studies in Health Technology and Informatics, 2017, 245, 1285.	0.3	7
165	Document Classification for Mining Host Pathogen Protein-Protein Interactions. , 2008, , .		6
166	Use of Echocardiography in Olmsted County Outpatients With Chest Pain and Normal Resting Electrocardiograms Seen at Mayo Clinic Rochester. Mayo Clinic Proceedings, 2015, 90, 1492-1498.	3.0	6
167	A computational framework for converting textual clinical diagnostic criteria into the quality data model. Journal of Biomedical Informatics, 2016, 63, 11-21.	4.3	6
168	Artificial intelligence to organize patient portal messages: a journey from an ensemble deep learning text classification to rule-based named entity recognition. , 2019, , .		6
169	Detecting and Filtering Immune-Related Adverse Events Signal Based on Text Mining and Observational Health Data Sciences and Informatics Common Data Model: Framework Development Study. JMIR Medical Informatics, 2020, 8, e17353.	2.6	6
170	A Topic-modeling Based Framework for Drug-drug Interaction Classification from Biomedical Text. AMIA ... Annual Symposium proceedings, 2016, 2016, 789-798.	0.2	6
171	Analysis of Clinical Variations in Asthma Care Documented in Electronic Health Records Between Staff and Resident Physicians. Studies in Health Technology and Informatics, 2017, 245, 1170-1174.	0.3	6
172	Standardizing Heterogeneous Annotation Corpora Using HL7 FHIR for Facilitating their Reuse and Integration in Clinical NLP. AMIA ... Annual Symposium proceedings, 2018, 2018, 574-583.	0.2	6
173	Artificial Intelligence Assesses Cliniciansâ€™ Adherence to Asthma Guidelines Using Electronic Health Records. Journal of Allergy and Clinical Immunology: in Practice, 2022, 10, 1047-1056.e1.	3.8	6
174	Teaching software engineering to end-users. , 2008, , .		5
175	Clinical Decision Support for Colonoscopy Surveillance Using Natural Language Processing. , 2012, , .		5
176	Automated Recommendation for Cervical Cancer Screening and Surveillance. Cancer Informatics, 2014, 13s3, CIN.S14035.	1.9	5
177	Retrieval of Semantically Similar Healthcare Questions in Healthcare Forums. , 2015, , .		5
178	A bibliometric analysis on tobacco regulation investigators. BioData Mining, 2015, 8, 11.	4.0	5
179	Constructing Node Embeddings for Human Phenotype Ontology to Assist Phenotypic Similarity Measurement. , 2018, , .		5
180	Discovering associations between problem list and practice setting. BMC Medical Informatics and Decision Making, 2019, 19, 69.	3.0	5

#	Article	IF	Citations
181	Impact of Diverse Data Sources on Computational Phenotyping. Frontiers in Genetics, 2020, 11, 556.	2.3	5
182	Digital Pathology-based Study of Cell- and Tissue-level Morphologic Features in Serous Borderline Ovarian Tumor and High-grade Serous Ovarian Cancer. Journal of Pathology Informatics, 2021, 12, 24.	1.7	5
183	TOWARDS PATHWAY CURATION THROUGH LITERATURE MINING â€“ A CASE STUDY USING PHARMGKB. , 2013, , .		5
184	Towards pathway curation through literature mining--a case study using PharmGKB. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing, 2014, , 352-63.	0.7	5
185	Discovering associations among diagnosis groups using topic modeling. AMIA Summits on Translational Science Proceedings, 2014, 2014, 43-9.	0.4	5
186	Exploring Gaps of Family History Documentation in EHR for Precision Medicine -A Case Study of Familial Hypercholesterolemia Ascertainment. AMIA Summits on Translational Science Proceedings, 2016, 2016, 160-6.	0.4	5
187	Identifying Information Gaps in Electronic Health Records by Using Natural Language Processing: Gynecologic Surgery History Identification. Journal of Medical Internet Research, 2022, 24, e29015.	4.3	5
188	Mapping Gene/Protein Names in Free Text to Biomedical Databases. , 2007, , .		4
189	Risk, Mechanisms and Implications of Asthma-Associated Infectious and Inflammatory Multimorbidities (AIMs) among Individuals With Asthma: a Systematic Review and a Case Study. Allergy, Asthma and Immunology Research, 2021, 13, 697.	2.9	4
190	Facilitating post-surgical complication detection through sublanguage analysis. AMIA Summits on Translational Science Proceedings, 2014, 2014, 77-82.	0.4	4
191	Towards Transforming Expert-based Content to Evidence-based Content. AMIA Summits on Translational Science Proceedings, 2014, 2014, 83-90.	0.4	4
192	Leveraging Terminology Services for Extract-Transform-Load Processes: A User-Centered Approach. AMIA ... Annual Symposium proceedings, 2016, 2016, 1010-1019.	0.2	4
193	A Text-Mining Framework for Supporting Systematic Reviews. , 2016, 1, 1-9.		4
194	Deep Denoising of Raw Biomedical Knowledge Graph From COVID-19 Literature, LitCovid, and Pubtator: Framework Development and Validation. Journal of Medical Internet Research, 2022, 24, e38584.	4.3	4
195	The Implication of Latent Information Quality to the Reproducibility of Secondary Use of Electronic Health Records. Studies in Health Technology and Informatics, 2022, , .	0.3	4
196	Automatically Extracting Sentences from Medline Citations to Support Clinicians' Information Needs. , 2012, , .		3
197	Integrating information retrieval with distant supervision for Gene Ontology annotation. Database: the Journal of Biological Databases and Curation, 2014, 2014, bau087-bau087.	3.0	3
198	Automated Chart Review for Asthma Ascertainment: An Innovative Approach for Asthma Care and Research in the Era of Electronic Medical Record. Journal of Allergy and Clinical Immunology, 2016, 137, AB196.	2.9	3

#	Article	IF	Citations
199	On Mapping Textual Queries to a Common Data Model. , 2017, , .		3
200	Advances in Text Mining and Visualization for Precision Medicine. , 2018, , .		3
201	Predicting Practice Setting Using Topic Modeling. , 2018, , .		3
202	Integrating word embedding neural networks with PubMed abstracts to extract keyword proximity of chronic diseases. , 2019, , .		3
203	Investigating the impact of disease and health record duration on the eMERGE algorithm for rheumatoid arthritis. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 601-605.	4.4	3
204	Probing Patient Messages Enhanced by Natural Language Processing: A Top-Down Message Corpus Analysis. Health Data Science, 2021, 2021, .	2.3	3
205	Genome-Wide Analysis of Loss of Heterozygosity in Breast Infiltrating Ductal Carcinoma Distant Normal Tissue Highlights Arm Specific Enrichment and Expansion across Tumor Stages. PLoS ONE, 2014, 9, e95783.	2.5	3
206	Correlating Lab Test Results in Clinical Notes with Structured Lab Data: A Case Study in HbA1c and Glucose. AMIA Summits on Translational Science Proceedings, 2017, 2017, 221-228.	0.4	3
207	Evaluating the Impact of Dictionary Updates on Automatic Annotations Based on Clinical NLP Systems. AMIA Summits on Translational Science Proceedings, 2019, 2019, 714-721.	0.4	3
208	iProLINK: A Framework for Linking Text Mining with Ontology and Systems Biology. , 2008, , .		2
209	Medical concept intersection between outside medical records and consultant notes: A case study in transferred cardiovascular patients. , 2017, , .		2
210	Leveraging Association Rule Mining to Detect Pathophysiological Mechanisms of Chronic Kidney Disease Complicated by Metabolic Syndrome. , 2018, , .		2
211	Generalized ensemble model for document ranking in information retrieval. Computer Science and Information Systems, 2017, 14, 123-151.	1.0	2
212	Characterizing Chronic Pain Episodes in Clinical Text at Two Health Care Systems: Comprehensive Annotation and Corpus Analysis. JMIR Medical Informatics, 2020, 8, e18659.	2.6	2
213	Analysis of medication and indication occurrences in clinical notes. AMIA ... Annual Symposium proceedings, 2014, 2014, 1046-55.	0.2	2
214	Granular Quality Reporting for Cervical Cytology Testing. AMIA Summits on Translational Science Proceedings, 2015, 2015, 178-82.	0.4	2
215	Mining Hierarchies and Similarity Clusters from Value Set Repositories. AMIA ... Annual Symposium proceedings, 2017, 2017, 1372-1381.	0.2	2
216	Incorporating Knowledge-Driven Insights into a Collaborative Filtering Model to Facilitate the Differential Diagnosis of Rare Diseases. AMIA ... Annual Symposium proceedings, 2018, 2018, 1505-1514.	0.2	2

#	Article	IF	Citations
217	Automating the Transformation of Free-Text Clinical Problems into SNOMED CT Expressions. AMIA Summits on Translational Science Proceedings, 2020, 2020, 497-506.	0.4	2
218	Developing Customizable Cancer Information Extraction Modules for Pathology Reports Using CLAMP. Studies in Health Technology and Informatics, 2019, 264, 1041-1045.	0.3	2
219	Midwest ruralâ€urban disparities in use of patient online services for COVIDâ€19. Journal of Rural Health, 2022, , .	2.9	2
220	Risk Factors for Silent Brain Infarcts and White Matter Disease in a Real-World Cohort Identified by Natural Language Processing. Mayo Clinic Proceedings, 2022, 97, 1114-1122.	3.0	2
221	Stratifying Future Stroke Risk with Incidentally Discovered White Matter Disease Severity and Covert Brain Infarct Site. Cerebrovascular Diseases, 2023, 52, 117-122.	1.7	2
222	Comparison of classification methods on protein-protein interaction document classification. , 2008, , .		1
223	Ontology-Based Temporal Relation Modeling with MapReduce Latent Dirichlet Allocations for Big EHR Data. , 2012, , .		1
224	Enhancing Clinical Information Retrieval through Context-Aware Queries and Indices. , 2019, , .		1
225	Subgrouping Rare Disease Patients Leveraging the Human Phenotype Ontology Embeddings. , 2020, , .		1
226	Enrich Rare Disease Phenotypic Characterizations via a Graph Convolutional Network Based Recommendation System. , 2020, , .		1
227	Abstract MP15: Validation of Phenotyping Algorithms for Stroke From Electronic Health Records Using Natural Language Processing. Circulation, 2019, 139, .	1.6	1
228	Establishing an expert consensus for the operational definitions of asthma-associated infectious and inflammatory multimorbidities for computational algorithms through a modified Delphi technique. BMC Medical Informatics and Decision Making, 2021, 21, 310.	3.0	1
229	A Bibliometric Analysis on Cancer Population Science with Topic Modeling. AMIA Summits on Translational Science Proceedings, 2015, 2015, 102-6.	0.4	1
230	Drug Normalization for Cancer Therapeutic and Druggable Genome Target Discovery. AMIA Summits on Translational Science Proceedings, 2015, 2015, 72-6.	0.4	1
231	Operationalizing Semantic Medline for meeting the information needs at point of care. AMIA Summits on Translational Science Proceedings, 2015, 2015, 152-6.	0.4	1
232	Assessing the Need of Discourse-Level Analysis in Identifying Evidence of Drug-Disease Relations in Scientific Literature. Studies in Health Technology and Informatics, 2015, 216, 539-43.	0.3	1
233	Distinction between medical and non-medical usages of short forms in clinical narratives. AMIA ... Annual Symposium proceedings, 2017, 2017, 1302-1311.	0.2	1
234	Natural Language Processing for the Evaluation of Methodological Standards and Best Practices of EHR-based Clinical Research. AMIA Summits on Translational Science Proceedings, 2020, 2020, 171-180.	0.4	1

#	Article	IF	Citations
235	Data-driven Sublanguage Analysis for Cancer Genomics Knowledge Modeling: Applications in Mining Oncological Genetics Information from Patients' Genetic Reports. AMIA Summits on Translational Science Proceedings, 2020, 2020, 720-729.	0.4	1
236	Risk of pneumonia in asthmatic children using inhaled corticosteroids: a nested case-control study in a birth cohort. BMJ Open, 2022, 12, e051926.	1.9	1
237	Relationship Between Very Cold Outside Weather and Surgical Outcome: Integrating Shallow and Deep Artificial Neural Nets. Studies in Health Technology and Informatics, 2019, 264, 1783-1784.	0.3	1
238	Development of dbOGAP: A bioinformatics resource of O-GlcNAcylated proteins and site prediction. , 2009, , .		0
239	Imbalanced text classification on host pathogen protein-protein interaction documents. , 2010, , .		0
240	Learning from positive and unlabeled documents for automated detection of alternative splicing sentences in medline abstracts. , 2011, , .		0
241	Evidence based computational drug repositioning candidate screening pipeline design: Case Study. , 2014, , .		0
242	LabLinker: Finding Links of Laboratory Test Results between Structured Records and Clinical Notes. , 2016, , .		0
243	Answering diabetic patients' questions using expert-vetted online resources: A case study. , 2016, , .		0
244	Identification of Genetic Causality Statements in Medline Abstracts Leveraging Distant Supervision. , 2018, , .		0
245	Information Extraction for Populating Lung Cancer Clinical Research Data. , 2019, 2019, .		0
246	A Deep Profiling and Visualization Framework to Audit Clinical Assessment Variation. , 2020, , .		0
247	Classification Systems for Bacterial Protein-Protein Interaction Document Retrieval. International Journal of Computational Models and Algorithms in Medicine, 2010, 1, 34-44.	0.4	0
248	Classification Systems for Bacterial Protein-Protein Interaction Document Retrieval. , 2012, , 33-43.		0
249	Abstract 16554: Rapid Identification of Familial Hypercholesterolemia From Electronic Health Records. Circulation, 2015, 132, .	1.6	0
250	PSB 2019 Workshop on Text Mining and Visualization for Precision Medicine. , 2018, , .		0
251	Annotating breast cancer microarray samples using ontologies. AMIA ... Annual Symposium proceedings, 2008, , 414-8.	0.2	0
252	An automated approach for ranking journals to help in clinician decision support. AMIA ... Annual Symposium proceedings, 2014, 2014, 757-66.	0.2	0

#	Article	IF	Citations
253	The Sublanguage of Clinical Problem Lists: A Corpus Analysis. AMIA ... Annual Symposium proceedings, 2018, 2018, 1451-1460.	0.2	0
254	Early Detection of Post-Surgical Complications using Time-series Electronic Health Records. AMIA Summits on Translational Science Proceedings, 2021, 2021, 152-160.	0.4	0
255	An Examination of the Statistical Laws of Semantic Change in Clinical Notes. AMIA Summits on Translational Science Proceedings, 2021, 2021, 515-524.	0.4	0
256	POKR: Building a Computable Heterogeneous Knowledge Resource for Precision Oncology. Studies in Health Technology and Informatics, 2022, , .	0.3	0