

Ian H Witten

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/683573/publications.pdf>

Version: 2024-02-01

180
papers

10,108
citations

147726

31
h-index

53190

85
g-index

185
all docs

185
docs citations

185
times ranked

5939
citing authors

#	ARTICLE	IF	CITATIONS
1	Arithmetic coding for data compression. Communications of the ACM, 1987, 30, 520-540.	3.3	2,287
2	Data mining in bioinformatics using Weka. Bioinformatics, 2004, 20, 2479-2481.	1.8	793
3	Learning to link with wikipedia. , 2008, , .		759
4	The zero-frequency problem: estimating the probabilities of novel events in adaptive text compression. IEEE Transactions on Information Theory, 1991, 37, 1085-1094.	1.5	450
5	Arithmetic coding revisited. ACM Transactions on Information Systems, 1998, 16, 256-294.	3.8	337
6	Using Model Trees for Classification. Machine Learning, 1998, 32, 63-76.	3.4	316
7	Mining meaning from Wikipedia. International Journal of Human Computer Studies, 2009, 67, 716-754.	3.7	250
8	Multiple viewpoint systems for music prediction. Journal of New Music Research, 1995, 24, 51-73.	0.6	244
9	An open-source toolkit for mining Wikipedia. Artificial Intelligence, 2013, 194, 222-239.	3.9	229
10	Towards the digital music library. , 1996, , .		180
11	Human-competitive tagging using automatic keyphrase extraction. , 2009, , .		179
12	Modeling for text compression. ACM Computing Surveys, 1989, 21, 557-591.	16.1	178
13	Technical Note: Naive Bayes for Regression. Machine Learning, 2000, 41, 5-25.	3.4	164
14	Improving browsing in digital libraries with keyphrase indexes. Decision Support Systems, 1999, 27, 81-104.	3.5	157
15	Thesaurus based automatic keyphrase indexing. , 2006, , .		139
16	Interactive machine learning: letting users build classifiers. International Journal of Human Computer Studies, 2001, 55, 281-292.	3.7	137
17	The Reactive Keyboard: a predictive typing aid. Computer, 1990, 23, 41-49.	1.2	129
18	Applying machine learning to agricultural data. Computers and Electronics in Agriculture, 1995, 12, 275-293.	3.7	116

#	ARTICLE	IF	CITATIONS
19	An adaptive optimal controller for discrete-time Markov environments. <i>Information and Control</i> , 1977, 34, 286-295.	1.3	109
20	A Compression-based Algorithm for Chinese Word Segmentation. <i>Computational Linguistics</i> , 2000, 26, 375-393.	2.5	108
21	Integrating error detection into arithmetic coding. <i>IEEE Transactions on Communications</i> , 1997, 45, 1-3.	4.9	106
22	Adaptive personalized interfaces – A question of viability. <i>Behaviour and Information Technology</i> , 1985, 4, 31-45.	2.5	104
23	A knowledge-based search engine powered by wikipedia. , 2007, , .		97
24	Using Peano Curves for Bilevel Display of Continuous-Tone Images. <i>IEEE Computer Graphics and Applications</i> , 1982, 2, 47-52.	1.0	94
25	Clustering Documents with Active Learning Using Wikipedia. , 2008, , .		89
26	One-Class Classification by Combining Density and Class Probability Estimation. <i>Lecture Notes in Computer Science</i> , 2008, , 505-519.	1.0	88
27	Mining Domain-Specific Thesauri from Wikipedia: A Case Study. , 2006, , .		72
28	Greenstone. , 2000, , .		71
29	The New Zealand Digital Library MELody inDEX. <i>D-Lib Magazine</i> , 1997, 3, .	0.5	68
30	Clustering Documents Using a Wikipedia-Based Concept Representation. <i>Lecture Notes in Computer Science</i> , 2009, , 628-636.	1.0	58
31	On the generation and use of space-filling curves. <i>Software - Practice and Experience</i> , 1983, 13, 519-525.	2.5	55
32	Domain-independent automatic keyphrase indexing with small training sets. <i>Journal of the Association for Information Science and Technology</i> , 2008, 59, 1026-1040.	2.6	55
33	Learning a concept-based document similarity measure. <i>Journal of the Association for Information Science and Technology</i> , 2012, 63, 1593-1608.	2.6	53
34	Text mining in a digital library. <i>International Journal on Digital Libraries</i> , 2004, 4, 56-59.	1.1	45
35	Realistic books. , 2004, , .		43
36	Textual image compression: two-stage lossy/lossless encoding of textual images. <i>Proceedings of the IEEE</i> , 1994, 82, 878-888.	16.4	40

#	ARTICLE	IF	CITATIONS
37	The bubble of web visibility. <i>Communications of the ACM</i> , 2005, 48, 115-117.	3.3	40
38	On the privacy afforded by adaptive text compression. <i>Computers and Security</i> , 1988, 7, 397-408.	4.0	39
39	Data compression in full-text retrieval systems. <i>Journal of the Association for Information Science and Technology</i> , 1993, 44, 508-531.	1.2	39
40	Towards a digital library of popular music. , 1999, , .		39
41	User modelling for a computer coach: a case study. <i>International Journal of Man-Machine Studies</i> , 1985, 23, 729-750.	0.7	38
42	Scalable browsing for large collections. , 2000, , .		38
43	Bonsai: A compact representation of trees. <i>Software - Practice and Experience</i> , 1993, 23, 277-291.	2.5	36
44	Using concept learning for knowledge acquisition. <i>International Journal of Man-Machine Studies</i> , 1988, 29, 171-196.	0.7	35
45	On-line and off-line heuristics for inferring hierarchies of repetitions in sequences. <i>Proceedings of the IEEE</i> , 2000, 88, 1745-1755.	16.4	35
46	The apparent conflict between estimation and control—a survey of the two-armed bandit problem. <i>Journal of the Franklin Institute</i> , 1976, 301, 161-189.	1.9	34
47	Lexically-generated subject hierarchies for browsing large collections. <i>International Journal on Digital Libraries</i> , 1999, 2, 111-123.	1.1	32
48	Displaying 3D images: algorithms for single-image random-dot stereograms. <i>Computer</i> , 1994, 27, 38-48.	1.2	31
49	The promise of digital libraries in developing countries. <i>Electronic Library</i> , 2002, 20, 7-13.	0.8	30
50	Utilizing lexical data from a Web-derived corpus to expand productive collocation knowledge. <i>ReCALL</i> , 2010, 22, 83-102.	3.2	28
51	Source models for natural language text. <i>International Journal of Man-Machine Studies</i> , 1990, 32, 545-579.	0.7	27
52	NIUPEPA. <i>Communications of the ACM</i> , 2001, 44, 86-87.	3.3	27
53	On frequency-based menu-splitting algorithms. <i>International Journal of Man-Machine Studies</i> , 1984, 21, 135-148.	0.7	26
54	Browsing in digital libraries. , 1997, , .		25

#	ARTICLE	IF	CITATIONS
55	Liveware: a new approach to sharing data in social networks. <i>International Journal of Man-Machine Studies</i> , 1991, 34, 337-348.	0.7	24
56	Seeking information in realistic books. , 2008, , .		24
57	An empirical evaluation of coding methods for multi-symbol alphabets. <i>Information Processing and Management</i> , 1994, 30, 791-804.	5.4	23
58	Tune Retrieval in the Multimedia Library. <i>Multimedia Tools and Applications</i> , 2000, 10, 113-132.	2.6	23
59	Adaptive text mining: inferring structure from sequences. <i>Journal of Discrete Algorithms</i> , 2004, 2, 137-159.	0.7	23
60	Realistic electronic books. <i>International Journal of Human Computer Studies</i> , 2012, 70, 588-610.	3.7	23
61	Comparing Human and Computational Models of Music Prediction. <i>Computer Music Journal</i> , 1994, 18, 70.	0.3	21
62	Managing complexity in a distributed digital library. <i>Computer</i> , 1999, 32, 74-79.	1.2	21
63	The promise of digital libraries in developing countries. <i>Communications of the ACM</i> , 2001, 44, 82-85.	3.3	21
64	Text Mining. <i>Chapman & Hall/CRC Computer and Information Science Series</i> , 2004, , .	0.4	21
65	A public library based on full-text retrieval. <i>Communications of the ACM</i> , 1998, 41, 71-75.	3.3	20
66	Supporting collocation learning with a digital library. <i>Computer Assisted Language Learning</i> , 2010, 23, 87-110.	4.8	20
67	Greenstone: open-source digital library software with end-user collection building. <i>Online Information Review</i> , 2001, 25, 288-298.	2.2	19
68	Refining the use of the web (and web search) as a language teaching and learning resource. <i>Computer Assisted Language Learning</i> , 2009, 22, 249-268.	4.8	19
69	Constructing a Focused Taxonomy from a Document Collection. <i>Lecture Notes in Computer Science</i> , 2013, , 367-381.	1.0	19
70	Dynamic Digital Library Construction and Configuration. <i>Lecture Notes in Computer Science</i> , 2004, , 1-13.	1.0	19
71	Can We Avoid High Coupling?. <i>Lecture Notes in Computer Science</i> , 2011, , 204-228.	1.0	19
72	Supporting command reuse: empirical foundations and principles. <i>International Journal of Man-Machine Studies</i> , 1993, 39, 353-390.	0.7	18

#	ARTICLE	IF	CITATIONS
73	A retrospective look at Greenstone. , 2007, , .		18
74	Greenstone. D-Lib Magazine, 2001, 7, .	0.5	18
75	Adaptive predictive text generation and the reactive keyboard. Interacting With Computers, 1991, 3, 27-50.	1.0	17
76	Learning text editing tasks from examples: a procedural approach. Behaviour and Information Technology, 1992, 11, 32-45.	2.5	17
77	Searching digital music libraries. Information Processing and Management, 2005, 41, 41-56.	5.4	17
78	A new framework for building digital library collections. , 2005, , .		17
79	A Flexible Scheme for Assigning Timing and Pitch To Synthetic Speech. Language and Speech, 1977, 20, 240-260.	0.6	16
80	Complexity-based induction. Machine Learning, 1994, 16, 203-225.	3.4	16
81	The relationship between greedy parsing and symbolwise text compression. Journal of the ACM, 1994, 41, 708-724.	1.8	16
82	The MG retrieval system. Communications of the ACM, 1995, 38, 41-42.	3.3	16
83	Measuring inter-indexer consistency using a thesaurus. , 2006, , .		16
84	The New Zealand Digital Library Project. D-Lib Magazine, 1996, 2, .	0.5	16
85	A system for interactive viewing of structured documents. Communications of the ACM, 1985, 28, 280-288.	3.3	15
86	A framework for knowledge acquisition through techniques of concept learning. IEEE Transactions on Systems, Man, and Cybernetics, 1989, 19, 499-512.	0.9	15
87	Digital libraries and minority languages. New Review of Hypermedia and Multimedia, 2005, 11, 139-155.	0.9	15
88	The Fastest Fourier Transform in the South. IEEE Transactions on Signal Processing, 2013, 61, 4707-4716.	3.2	15
89	Reflexive CSCW: supporting long-term personal work. Interacting With Computers, 1990, 2, 330-336.	1.0	14
90	Explaining cryptographic systems. Computers and Education, 2003, 40, 199-215.	5.1	14

#	ARTICLE	IF	CITATIONS
91	Creating digital library collections with Greenstone. Library Hi Tech, 2005, 23, 541-560.	3.7	14
92	StoneD. D-Lib Magazine, 2005, 11, .	0.5	14
93	Supporting command reuse: mechanisms for reuse. International Journal of Man-Machine Studies, 1993, 39, 391-425.	0.7	13
94	APPROXIMATE, NON-DETERMINISTIC MODELLING OF BEHAVIOUR SEQUENCES. International Journal of General Systems, 1979, 5, 1-12.	1.2	12
95	Concepts of cooperation in artificial life. IEEE Transactions on Systems, Man, and Cybernetics, 1995, 25, 1166-1171.	0.9	12
96	Creating and Reading Realistic Electronic Books. Computer, 2009, 42, 72-81.	1.2	12
97	Greenstone: A Platform for Distributed Digital Library Applications. Lecture Notes in Computer Science, 2001, , 137-148.	1.0	12
98	Examples of Practical Digital Libraries. D-Lib Magazine, 2003, 9, .	0.5	12
99	A Planning Mechanism for Generating Story Text. Literary and Linguistic Computing, 1991, 6, 119-126.	0.6	11
100	Subject metadata support powered by Maui. , 2010, , .		11
101	Teaching agents to learn: from user study to implementation. Computer, 1997, 30, 36-44.	1.2	10
102	Greenstone digital library software. , 2004, , .		10
103	Building digital library collections with greenstone. , 2005, , .		10
104	Digital libraries for the developing world. Interactions, 2006, 13, 20-21.	0.8	10
105	Human operators and automatic adaptive controllers: A comparative study on a particular control task. International Journal of Man-Machine Studies, 1973, 5, 75-104.	0.7	9
106	Lossless Compression for Text and Images. International Journal of High Speed Electronics and Systems, 1997, 08, 179-231.	0.3	9
107	Data compression in full-text retrieval systems. , 1993, 44, 508.		9
108	Finite-Time Performance of Some Two-Armed Bandit Controllers. IEEE Transactions on Systems, Man, and Cybernetics, 1973, SMC-3, 194-197.	0.9	8

#	ARTICLE	IF	CITATIONS
109	Inferring Graphical Procedures: The Compleat Metamouse. <i>Human-Computer Interaction</i> , 1992, 7, 47-89.	3.1	8
110	Cima: An interactive concept learning system for end-user applications. <i>Applied Artificial Intelligence</i> , 1997, 11, 653-671.	2.0	8
111	Extracting text from PostScript. <i>Software - Practice and Experience</i> , 1998, 28, 481-491.	2.5	8
112	Knowledge Visualization Techniques for Machine Learning. <i>Intelligent Data Analysis</i> , 1998, 2, 333-347.	0.4	8
113	Constructing a Collocation Learning System from the Wikipedia Corpus. <i>International Journal of Computer-Assisted Language Learning and Teaching</i> , 2016, 6, 18-35.	0.5	8
114	Managing Change in a Digital Library System with Many Interface Languages. <i>Lecture Notes in Computer Science</i> , 2003, , 350-361.	1.0	8
115	Elements of computer typography. <i>International Journal of Man-Machine Studies</i> , 1985, 23, 623-687.	0.7	7
116	The New Zealand Digital Library: Collections and experience. <i>Electronic Library</i> , 1997, 15, 495-504.	0.8	7
117	Complexity-Based Induction. <i>Machine Learning</i> , 1994, 16, 203-225.	3.4	6
118	Portable digital libraries on an ipod. , 2008, , .		6
119	A competitive environment for exploratory query expansion. , 2008, , .		6
120	A fedora librarian interface. , 2008, , .		6
121	Jade. <i>Operating Systems Review (ACM)</i> , 1983, 17, 10-23.	1.5	5
122	Phrase hierarchy inference and compression in bounded space. , 0, , .		5
123	Digital libraries for creative communities. <i>Digital Creativity</i> , 2004, 15, 110-125.	0.8	5
124	Bin Encoding: A User-Centric Secure Full-Text Searching Scheme for the Cloud. , 2015, , .		5
125	Privacy Preserving Computation by Fragmenting Individual Bits and Distributing Gates. , 2016, , .		5
126	Managing Personal Documents with a Digital Library. <i>Lecture Notes in Computer Science</i> , 2005, , 195-206.	1.0	5

#	ARTICLE	IF	CITATIONS
127	Towards a Digital Library for Language Learning. Lecture Notes in Computer Science, 2006, , 341-352.	1.0	5
128	Exploring, modelling and controlling discrete sequential environments. International Journal of Man-Machine Studies, 1977, 9, 715-735.	0.7	4
129	Indexing and compressing full-text databases for CD-ROM. Journal of Information Science, 1991, 17, 265-271.	2.0	4
130	Learning from Batched Data: Model Combination Versus Data Combination. Knowledge and Information Systems, 1999, 1, 83-106.	2.1	4
131	The development and usage of the Greenstone digital library software. Bulletin of the American Society for Information Science, 2009, 35, 31-38.	0.3	4
132	Computer graphics techniques for modeling page turning. International Journal on Digital Libraries, 2009, 10, 93-121.	1.1	4
133	Developing and Evaluating a Learner-Friendly Collocation System With User Query Data. International Journal of Computer-Assisted Language Learning and Teaching, 2019, 9, 53-78.	0.5	4
134	A Combined Phrase and Thesaurus Browser for Large Document Collections. Lecture Notes in Computer Science, 2001, , 25-36.	1.0	4
135	Beyond the Client-Server Model: Self-contained Portable Digital Libraries. Lecture Notes in Computer Science, 2008, , 294-303.	1.0	4
136	PROBABILISTIC BEHAVIOUR/STRUCTURE TRANSFORMATIONS USING TRANSITIVE MOORE MODELS. International Journal of General Systems, 1980, 6, 129-137.	1.2	3
137	On the power of traps and diversions in a document preparation language. Software - Practice and Experience, 1982, 12, 1119-1131.	2.5	3
138	Explanation-based learning: its role in problem solving. Journal of Experimental and Theoretical Artificial Intelligence, 1989, 1, 27-49.	1.8	3
139	Modeling music: Systems, structure, and prediction. Interface, 1990, 19, 53-66.	0.2	3
140	Searching Digital Music Libraries. Lecture Notes in Computer Science, 2002, , 129-140.	1.0	3
141	Applying machine learning to programming by demonstration. Journal of Experimental and Theoretical Artificial Intelligence, 2004, 16, 161-188.	1.8	3
142	Lightweight realistic books. , 2007, , .		3
143	Running greenstone on an ipod. , 2008, , .		3
144	F-Lingo: Integrating lexical feature identification into MOOC platforms for learning professional and academic English. , 2019, , .		3

#	ARTICLE	IF	CITATIONS
145	A User Evaluation of Hierarchical Phrase Browsing. Lecture Notes in Computer Science, 2003, , 313-324.	1.0	3
146	A course on "expert systems" for electrical engineering students. SIGCSE Bulletin, 1987, 19, 257-260.	0.1	3
147	Programming by example: The human face of AI. AI and Society, 1992, 6, 166-180.	3.1	2
148	An experiment in the application of similarity-based learning to programming by example. International Journal of Intelligent Systems, 1994, 9, 341-364.	3.3	2
149	Document level interoperability for collection creators. , 2006, , .		2
150	A Systematic Review of Using Discipline-Specific Corpora for Lexico-Grammatical Pattern Learning. International Journal of Computer-Assisted Language Learning and Teaching, 2018, 8, 31-49.	0.5	2
151	Content-Based Language Learning in a Digital Library. , 2007, , 424-433.		2
152	Getting research students started. SIGCSE Bulletin, 1993, 25, 165-169.	0.1	2
153	First Person Singular. International Journal of Digital Library Systems, 2010, 1, 24-43.	0.1	2
154	Token Identification Using HMM and PPM Models. Lecture Notes in Computer Science, 2003, , 173-185.	1.0	2
155	Automatically Augmenting Academic Text for Language Learning. Advances in Educational Technologies and Instructional Design Book Series, 2018, , 512-537.	0.2	2
156	Driving the Votrax speech synthesizer from a wide phonetic transcription with high-level prosodic markers. International Journal of Man-Machine Studies, 1982, 16, 393-403.	0.7	1
157	Accelerating search in function induction. Journal of Experimental and Theoretical Artificial Intelligence, 1990, 2, 131-150.	1.8	1
158	Greenstone in practice: Implementations of an open source digital library system. Proceedings of the American Society for Information Science and Technology, 2006, 42, n/a-n/a.	0.2	1
159	Perambulating libraries. , 2011, , .		1
160	Digital Libraries: Developing Countries, Universal Access, and Information for All. Lecture Notes in Computer Science, 2004, , 35-44.	1.0	1
161	Collocation Games from a Language Corpus. , 2012, , 209-229.		1
162	Learning Structure from Sequences, with Applications in a Digital Library. Lecture Notes in Computer Science, 2002, , 42-56.	1.0	1

#	ARTICLE	IF	CITATIONS
163	F-Lingo: Leveraging Smart CALL for massive open online courses. , 2022, , 293-319.		1
164	Impact of the new microprocessors on management structures. IEE Proceedings A: Physical Science Measurement and Instrumentation Management and Education Reviews, 1982, 129, 244.	0.1	0
165	Towards Distributed Document Preparation With Interactive And Non-Interactive Viewing. Infor, 1985, 23, 365-388.	0.5	0
166	Evaluating programs formed by example: An informational heuristic. , 1991, , 388-402.		0
167	Predicting query times. , 1998, , .		0
168	Practical digital library interoperability standards. , 2005, , .		0
169	Multilingual digital libraries: Research and practice. Proceedings of the American Society for Information Science and Technology, 2006, 42, n/a-n/a.	0.2	0
170	How the dragons work. , 2007, , .		0
171	A Bookmaker's Workbench. , 2011, , .		0
172	A link-based visual search engine for Wikipedia. , 2011, , .		0
173	Exploring Wikipedia with HMpara. , 2011, , .		0
174	Browsing around a Digital Library: Today and Tomorrow. Lecture Notes in Computer Science, 2000, , 12-26.	1.0	0
175	First Person Singular. , 2012, , 22-40.		0
176	Exploring, Modelling, and Controlling Discrete Sequential Environments. , 1978, , 161-174.		0
177	Statistical Models of Language Use. , 1990, , 269-280.		0
178	LOSSLESS COMPRESSION FOR TEXT AND IMAGES. Selected Topics in Electornics and Systems, 1997, , 179-231.	0.2	0
179	Constructing a Collocation Learning System From the Wikipedia Corpus. , 2019, , 1018-1037.		0
180	Non-Deterministic Modelling and its Application in Adaptive Optimal Control. Lecture Notes in Statistics, 1983, , 213-226.	0.1	0