

# Paolo Merialdo

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/4240502/publications.pdf>

Version: 2024-02-01

40  
papers

620  
citations

759055

12  
h-index

677027

22  
g-index

41  
all docs

41  
docs citations

41  
times ranked

347  
citing authors

#	ARTICLE	IF	CITATIONS
1	Semistructured and structured data in the Web: going back and forth. SIGMOD Record, 1997, 26, 16-23.	0.7	70
2	Clustering Web pages based on their structure. Data and Knowledge Engineering, 2005, 54, 279-299.	2.1	59
3	Design and development of data-intensive web sites. ACM Transactions on Internet Technology, 2003, 3, 49-92.	3.0	50
4	Extraction and integration of partially overlapping web sources. Proceedings of the VLDB Endowment, 2013, 6, 805-816.	2.1	37
5	Data-Intensive Web Sites: Design and Maintenance. World Wide Web, 2001, 4, 21-47.	2.7	25
6	Myocardial Contrast Enhancement After Intravenous Injection of Sonicated Albumin Microbubbles: A Transesophageal Echocardiography Dipyridamole Study. Journal of the American Society of Echocardiography, 1994, 7, 337-346.	1.2	24
7	Crowdsourcing for data management. Knowledge and Information Systems, 2017, 53, 1-41.	2.1	24
8	WRAPPER INFERENCE FOR AMBIGUOUS WEB PAGES. Applied Artificial Intelligence, 2008, 22, 21-52.	2.0	22
9	Probabilistic Models to Reconcile Complex Data from Inaccurate Data Sources. Notes on Numerical Fluid Mechanics and Multidisciplinary Design, 2010, , 83-97.	0.2	22
10	A framework for learning web wrappers from the crowd. , 2013, , .		21
11	Wrapping-oriented classification of web pages. , 2002, , .		15
12	Crowdsourcing large scale wrapper inference. Distributed and Parallel Databases, 2015, 33, 95-122.	1.0	15
13	Web Content Extraction. SIGKDD Explorations: Newsletter of the Special Interest Group (SIG) on Knowledge Discovery & Data Mining, 2016, 17, 17-23.	3.2	15
14	Crawling programs for wrapper-based applications. , 2008, , .		11
15	Leveraging Wikipedia Table Schemas for Knowledge Graph Augmentation. , 2018, , .		11
16	Fine-grain web site structure discovery. , 2003, , .		10
17	Towards Knowledge Discovery from the Vatican Secret Archives. In Codice Ratio - Episode 1. , 2018, , .		10
18	Supporting the automatic construction of entity aware search engines. , 2008, , .		9

#	ARTICLE	IF	CITATIONS
19	Towards Annotating Relational Data on the Web with Language Models. , 2018, , .		8
20	An Automatic Data Grabber for Large Web Sites. , 2004, , 1321-1324.		7
21	Flint. , 2008, , .		6
22	In Codice Ratio: A crowd-enabled solution for low resource machine transcription of the Vatican Registers. Information Processing and Management, 2021, 58, 102606.	5.4	6
23	Redundancy-driven web data extraction and integration. , 2010, , .		6
24	Managing Web-based data - database models and transformations. IEEE Internet Computing, 2002, 6, 33-37.	3.2	5
25	Exploiting information redundancy to wring out structured data from the web. , 2010, , .		5
26	ALFRED. , 2013, , .		5
27	Self-supervised learning for medieval handwriting identification: A case study from the Vatican Apostolic Library. Information Processing and Management, 2022, 59, 102875.	5.4	5
28	Web-Scale Extension of RDF Knowledge Bases from Templated Websites. Lecture Notes in Computer Science, 2014, , 66-81.	1.0	4
29	Accurate fact harvesting from natural language text in wikipedia with Lector. , 2016, , .		3
30	Multikernel Activation Functions: Formulation and a Case Study. Proceedings of the International Neural Networks Society, 2020, , 320-329.	0.6	3
31	HOMER. SIGMOD Record, 2000, 29, 586.	0.7	2
32	Automatically building probabilistic databases from the web. , 2011, , .		2
33	Big Data Linkage for Product Specification Pages. , 2018, , .		2
34	In Codice Ratio: Machine Transcription of Medieval Manuscripts. Communications in Computer and Information Science, 2019, , 185-192.	0.4	2
35	Fine-grained semantic type discovery for heterogeneous sources using clustering. VLDB Journal, 2023, 32, 305-324.	2.7	2
36	Wrapper Generation for Overlapping Web Sources. , 2011, , .		1

#	ARTICLE	IF	CITATIONS
37	Hybrid Crowd-Machine Wrapper Inference. ACM Transactions on Knowledge Discovery From Data, 2019, 13, 1-43.	2.5	1
38	Web Data Reconciliation: Models and Experiences. Lecture Notes in Computer Science, 2012, , 1-15.	1.0	1
39	Characterizing the uncertainty of web data. , 2011, , .		1
40	Flint: From Web Pages to Probabilistic Semantic Data. Data-centric Systems and Applications, 2012, , 333-359.	0.2	0