

# Roman Kern

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/7277580/publications.pdf>

Version: 2024-02-01

61  
papers

626  
citations

758635

12  
h-index

752256

20  
g-index

70  
all docs

70  
docs citations

70  
times ranked

633  
citing authors

#	ARTICLE	IF	CITATIONS
1	Understanding the true effects of the COVID-19 lockdown on air pollution by means of machine learning. <i>Environmental Pollution</i> , 2021, 274, 115900.	3.7	54
2	PySpark and RDKit: Moving towards Big Data in Cheminformatics. <i>Molecular Informatics</i> , 2019, 38, e1800082.	1.4	47
3	Of categorizers and describers. , 2010, , .		42
4	Understanding why users tag: A survey of tagging motivation literature and results from an empirical study. <i>Web Semantics</i> , 2012, 17, 1-11.	2.2	36
5	Authorship identification of documents with high content similarity. <i>Scientometrics</i> , 2018, 115, 223-237.	1.6	26
6	Machine learning in prediction of intrinsic aqueous solubility of drug-like compounds: Generalization, complexity, or predictive ability?. <i>Journal of Chemometrics</i> , 2021, 35, e3349.	0.7	25
7	Polarity Classification for Target Phrases in Tweets: A Word2Vec Approach. <i>Lecture Notes in Computer Science</i> , 2016, , 217-223.	1.0	22
8	Recommending Tags for Pictures Based on Text, Visual Content and User Context. , 2008, , .		19
9	Analysis of structural relationships for hierarchical cluster labeling. , 2010, , .		18
10	Feature Extraction From Analog Wafermaps: A Comparison of Classical Image Processing and a Deep Generative Model. <i>IEEE Transactions on Semiconductor Manufacturing</i> , 2019, 32, 190-198.	1.4	18
11	Deep learning—a first meta-survey of selected reviews across scientific disciplines, their commonalities, challenges and research impact. <i>PeerJ Computer Science</i> , 2021, 7, e773.	2.7	18
12	Big data as a promoter of industry 4.0: Lessons of the semiconductor industry. , 2017, , .		17
13	Ubiquitous Access to Digital Cultural Heritage. <i>Journal on Computing and Cultural Heritage</i> , 2017, 10, 1-27.	1.2	17
14	Unsupervised document structure analysis of digital scientific articles. <i>International Journal on Digital Libraries</i> , 2014, 14, 83-99.	1.1	16
15	Deriving Public Transportation Timetables with Large-Scale Cell Phone Data. <i>Procedia Computer Science</i> , 2015, 52, 67-74.	1.2	14
16	QZTool—Automatically Generated Origin-Destination Matrices from Cell Phone Trajectories. <i>Advances in Intelligent Systems and Computing</i> , 2017, , 823-833.	0.5	14
17	Efficient linear text segmentation based on information retrieval techniques. , 2009, , .		13
18	Map-Matching Cell Phone Trajectories of Low Spatial and Temporal Accuracy. , 2015, , .		11

#	ARTICLE	IF	CITATIONS
19	Extraction of References Using Layout and Formatting Information from Scientific Articles. D-Lib Magazine, 2013, 19, .	0.5	11
20	Gaussian Process Surrogates for Modeling Uncertainties in a Use Case of Forging Superalloys. Applied Sciences (Switzerland), 2022, 12, 1089.	1.3	11
21	Mesh-Free Surrogate Models for Structural Mechanic FEM Simulation: A Comparative Study of Approaches. Applied Sciences (Switzerland), 2021, 11, 9411.	1.3	10
22	Constructing robust health indicators from complex engineered systems via anticausal learning. Engineering Applications of Artificial Intelligence, 2022, 113, 104926.	4.3	10
23	Aspects of Broad Folksonomies. , 2007, , .		9
24	A comparison of layout based bibliographic metadata extraction techniques. , 2012, , .		9
25	Predicting Treatment Outcomes Using Explainable Machine Learning in Children with Asthma. Children, 2021, 8, 376.	0.6	9
26	Should We Embed in Chemistry? A Comparison of Unsupervised Transfer Learning with PCA, UMAP, and VAE on Molecular Fingerprints. Pharmaceuticals, 2021, 14, 758.	1.7	9
27	A Comparison of Supervised Approaches for Process Pattern Recognition in Analog Semiconductor Wafer Test Data. , 2018, , .		8
28	Predictive Capability of QSAR Models Based on the CompTox Zebrafish Embryo Assays: An Imbalanced Classification Problem. Molecules, 2021, 26, 1617.	1.7	8
29	Extending Folksonomies for Image Tagging. , 2008, , .		7
30	SAZED: parameter-free domain-agnostic season length estimation in time series data. Data Mining and Knowledge Discovery, 2019, 33, 1775-1798.	2.4	7
31	Parasitic resistance as a predictor of faulty anodes in electro galvanizing: a comparison of machine learning, physical and hybrid models. Advanced Modeling and Simulation in Engineering Sciences, 2020, 7, .	0.7	7
32	Driver's dashboard " using social media data as additional information for motorway operators. IET Intelligent Transport Systems, 2018, 12, 1116-1122.	1.7	6
33	A Comparison of Two Unsupervised Table Recognition Methods from Digital Scientific Articles. D-Lib Magazine, 2014, 20, .	0.5	6
34	Improving the Consistency of the Failure Mode Effect Analysis (FMEA) Documents in Semiconductor Manufacturing. Applied Sciences (Switzerland), 2022, 12, 1840.	1.3	6
35	Markov random fields for pattern extraction in analog wafer test data. , 2017, , .		5
36	Self- and Cross-Excitation in Stack Exchange Question & Answer Communities. , 2019, , .		5

#	ARTICLE	IF	CITATIONS
37	Unleashing Semantics of Research Data. Lecture Notes in Computer Science, 2014, , 103-112.	1.0	5
38	Source selection of long tail sources for federated search in an uncooperative setting. , 2018, , .		4
39	A semantic federated search engine for domain-specific document retrieval. , 2017, , .		4
40	Efficient Search Result Diversification via Query Expansion Using Knowledge Bases. , 2015, , .		3
41	Evaluation of Pseudo Relevance Feedback Techniques for Cross Vertical Aggregated Search. Lecture Notes in Computer Science, 2015, , 91-102.	1.0	3
42	KOMPOS: Connecting Causal Knots in Large Nonlinear Time Series with Non-Parametric Regression Splines. ACM Transactions on Intelligent Systems and Technology, 2021, 12, 1-27.	2.9	3
43	Distributed Web2.0 crawling for ontology evolution. , 2007, , .		2
44	Model Selection Strategies for Author Disambiguation. , 2011, , .		2
45	Is enterprise search useful at all?. , 2014, , .		2
46	A Health Factor for Process Patterns Enhancing Semiconductor Manufacturing by Pattern Recognition in Analog Wafermaps. , 2019, , .		2
47	Efficient Table Annotation for Digital Articles. D-Lib Magazine, 2015, 21, .	0.5	2
48	A generative semi-supervised classifier for datasets with unknown classes. , 2020, , .		2
49	Towards a Marketplace for the Scientific Community: Accessing Knowledge from the Computer Science Domain. D-Lib Magazine, 2014, 20, .	0.5	2
50	A hybrid system for German encyclopedia alignment. International Journal on Digital Libraries, 2010, 11, 75-89.	1.1	1
51	An Information Retrieval Based Approach for Multilingual Ontology Matching. Lecture Notes in Computer Science, 2016, , 433-439.	1.0	1
52	Cluster Purging: Efficient Outlier Detection based on Rate-Distortion Theory. IEEE Transactions on Knowledge and Data Engineering, 2021, , 1-1.	4.0	1
53	Exploring the Influence of Tagging Motivation on Tagging Behavior. Lecture Notes in Computer Science, 2010, , 461-465.	1.0	1
54	Know-Center at SemEval-2016 Task 5: Using Word Vectors with Typed Dependencies for Opinion Target Expression Extraction. , 2016, , .		1

#	ARTICLE	IF	CITATIONS
55	A Neural-based Architecture For Small Datasets Classification. , 2020, , .		1
56	Enhanced Active Learning of Convolutional Neural Networks: A Case Study for Defect Classification in the Semiconductor Industry. , 2020, , .		1
57	Query Splitting for Context-Driven Federated Recommendations. , 2016, , .		0
58	Evaluation of Contextualization and Diversification Approaches in Aggregated Search. , 2017, , .		0
59	Generating Tailored Classification Schemas for German Patents. Lecture Notes in Computer Science, 2016, , 230-238.	1.0	0
60	Do Ambiguous Words Improve Probing for Federated Search?. Lecture Notes in Computer Science, 2016, , 438-441.	1.0	0
61	ManEx: The Visual Analysis of Measurements for the Assessment of Errors in Electrical Engines. IEEE Computer Graphics and Applications, 2022, 42, 68-80.	1.0	0