

# Supplementary Information: The Impact of Responsible AI Research on Innovation and Development

Ali Akbar Septiandri<sup>1</sup>, Marios Constantinides<sup>1</sup>, Daniele Quercia<sup>1, 2</sup>

<sup>1</sup>Nokia Bell Labs, Cambridge, UK

<sup>2</sup>Kings College London, London, UK

{ali.septiandri, marios.constantinides, daniele.quercia}@nokia-bell-labs.com

## Keywords to Filter RAI Papers

We developed a set of 25 keywords based on the NIST framework to identify RAI papers in our corpus. For all the listed keywords in Table A.4, we prepended “artificial intelligence” or “machine learning” to ensure the resulting papers are about RAI.

Topic	Keywords
Fairness	fairness, equality, equity, equitable privacy, anonymity, confidentiality, confidential
Privacy	
Explainability	explainability, explainable accountable, accountability, trans- parency, auditability, governance, compliance, accountability mecha- nisms, algorithmic accountability
Accountability	
Sustainability	green, energy-efficient, carbon footprint, environmental impact, eco-friendly, energy consumption, green computing

Table A.4: Selected NIST keywords to filter RAI papers. From the list of 31 NIST keywords, we found that some of them are non-discriminatory. For example, the keyword ‘validation’ is frequently used as a generic term in AI, but it does not serve as a specific keyword for identifying RAI papers. After eliminating these non-discriminatory terms, we retained 18 keywords and included an additional 7 keywords related to ‘Sustainability,’ since it is a significant topic in the public RAI debate.

**Removed keywords.** These keywords were removed from our query because they were non-discriminative as they are often used to describe generally AI. The 19 keywords included: validation, reliability, correctness, accuracy, robustness, generalizability, authenticity, quality, measurability, dependability, capability, safety, security, resilience, confidentiality, integrity, availability, usability, controllability.

## Analyzed Venues

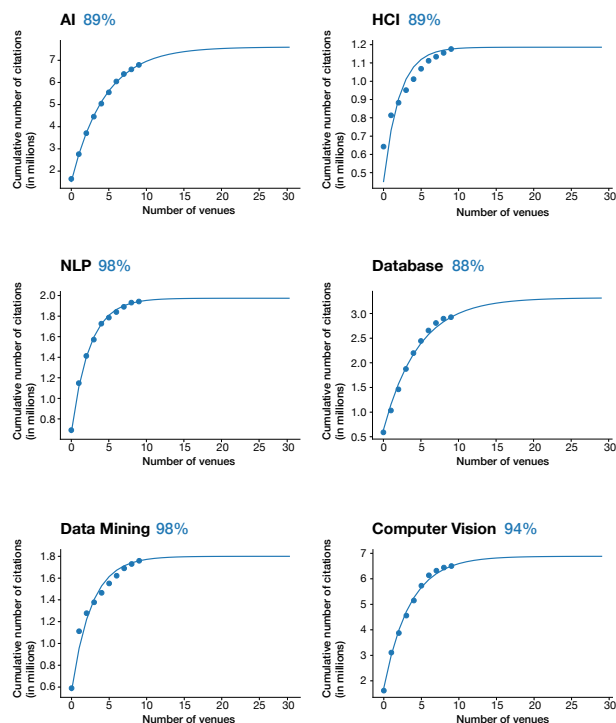


Figure A.6: The cumulative number of citations in millions as a function of top venues considered in each research area. The top 10 venues in the six research areas (AI, Database, NLP, HCI, Data Mining, and Computer Vision) cover more than 88% of the total citations in each area. The numbers at the top of each subplot show the total coverage of citations from the top 10 venues in each area.

Category	Venue	# Papers
Artificial Intelligence	Neural Information Processing Systems (NeurIPS)	16049
	International Conference on Learning Representations (ICLR)	5679
	International Conference on Machine Learning (ICML)	9139
	AAAI Conference on Artificial Intelligence (AAAI)	15120
	Expert Systems with Applications	17051
	IEEE Transactions on Neural Networks and Learning Systems	5419
	IEEE Transactions On Systems, Man And Cybernetics Part B, Cybernetics	1917
	Neurcomputing	18004
	International Joint Conference on Artificial Intelligence (IJCAI)	11721
Human-Computer Interaction	Applied Soft Computing	8756
	Computer Human Interaction (CHI)	10640
	Proceedings of the ACM on Human-Computer Interaction (CSCW)	2039
	Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)	1408
	IEEE Transactions on Affective Computing	883
	International Journal of Human-Computer Studies	2268
	Behaviour and Information Technology	2160
	ACM/IEEE International Conference on Human Robot Interaction (HRI)	3475
	International Journal of Human-Computer Interaction	2145
Computer Vision	Virtual Reality	1279
	International Conference on Intelligent User Interfaces (IUI)	1849
	IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	9812
	European Conference on Computer Vision (ECCV)	5908
	IEEE/CVF International Conference on Computer Vision (ICCV)	4775
	IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)	7398
	IEEE Transactions on Image Processing	8912
	Pattern Recognition	9837
	IEEE/CVF Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)	3684
Data Mining & Analytics	Medical Image Analysis	2530
	IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)	2815
	International Journal of Computer Vision (IJCV)	3313
	ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)	5195
	IEEE Transactions on Knowledge and Data Engineering	8546
	International Conference on Artificial Intelligence and Statistics (AISTATS)	3783
	ACM International Conference on Web Search and Data Mining (WSDM)	1565
	Journal of Big Data	1350
	Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery	1474
Natural Language Processing	IEEE International Conference on Big Data (Big Data)	6016
	IEEE International Conference on Data Mining (ICDM)	1544
	ACM Conference on Recommender Systems (RecSys)	1649
	Knowledge and Information Systems	3009
	Meeting of the Association for Computational Linguistics (ACL)	9314
	Conference on Empirical Methods in Natural Language Processing (EMNLP)	7565
	Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL)	3518
	Transactions of the Association for Computational Linguistics (TACL)	214
	International Conference on Computational Linguistics (COLING)	5793
International Conference on Language Resources and Evaluation (LREC)	7994	
Database & Information Systems	Workshop on Machine Translation (WMT)	1330
	International Workshop on Semantic Evaluation (SemEval)	2582
	Conference on Computational Natural Language Learning (CoNLL)	989
	Computer Speech and Language	1254
	International World Wide Web Conferences (WWW)	7958
	IEEE Transactions on Knowledge and Data Engineering	8546
	ACM SIGIR Conference on Research and Development in Information Retrieval	5840
	Information Processing & Management	3519
	ACM International Conference on Information and Knowledge Management (CIKM)	7106
Responsible AI	International Conference on Very Large Databases (VLDB)	6248
	ACM International Conference on Web Search and Data Mining (WSDM)	1565
	ACM SIGMOD International Conference on Management of Data	4940
	Journal of Big Data	1350
	International Conference on Data Engineering (ICDE)	5575
	ACM Conference on Fairness, Accountability, and Transparency (FAccT)	339
	AAAI/ACM Conference on AI, Ethics, and Society (AIES)	474

Table A.5: Distribution of papers (including non-RAI papers) over the years for each research area. Venues are ranked based on their h5 index impact factor. According to Google Scholar, h5 index is the h-index for articles published in the last five complete years; it is the largest number  $h$  such that  $h$  articles published in 2018-2022 have at least  $h$  citations each.