Note on the 'ILLC Core Area Labels' in the bibliometric analysis

October 14, 2024

On the basis of the titles, all ILLC research papers in OpenAlex have been categorized (by automatic machine learning methods) into many topics, and each one comes with a label. The ILLC (during 2018–2023) has publications categorized into 200 topics (each paper is categorized into exactly 1 topic). Below you see the list of all OpenAlex labels of topics with at least 10 ILLC publications. The Core Area for each institute is then determined as the 15 topics in which the institute has the most publications. The (automatically generated) OpenAlex labels are included at the left-hand side of the table below. Some of these labels are somewhat surprising for the ILLC, and we have given the Core Area labels that are more recognizable for the ILLC, listed under the heading 'ILLC labels'. However, we do not have the specific information about which lists of ILLC papers actually have been categorized under a specific OpenAlex label.

The ILLC Core Area, as determined by the mentioned 15 labels, is important for a comparison with other institutes. As mentioned in the self-evaluation document, in terms of impact w.r.t. the Core Area, the ILLC is ranked 8^{th}. However, this comparative analysis does depend on the specific labelling that the AI-system has used. In our view, the presented results could have been significantly different in case the OpenAlex categories would be more in line with the ILLC labels. If the content of the OpenAlex categories would have been more in accordance with our ILLC labels, less world-wide publications would fall under it and ILLC's percentage would presumably be much higher. This holds, we guess, for many of the categories in the ILLC Core Area. So while the comparative rankings can be a useful tool for traditional well-demarcated research disciplines, such a comparative analysis is much more difficult to conduct and to interpret when we take interdisciplinary research domains into account.

aantal ILLC pubs	OpenALex label	ILLC label
188	Logic Programming and Knowledge Representation	Logic, Epistemology, and Social Dynamics
129	Natural Language Processing	Advanced Research in NLP, Language Models, and Cross-Lingual Studies
126	Statistical Machine Translation and Natural Language Processing	Advanced Computational Linguistics and Language Processing
59	Quantum Computing and Simulation	Quantum Algorithms and Computational Complexity
38	Visual Question Answering in Images and Videos	Visual Question Answering and Multimodal Learning
37	Consciousness and Phenomenal Experience	Philosophical Logic and Metaphysics
37	Mechanism Design and Matching Theory in Economics	Social Choice Theory and Participatory Budgeting
35	Program Analysis and Verification Techniques	Modal Logic and Mu-Calculus Research
32	Fuzzy Logic and Residuated Lattices	Intermediate Logics and Algebraic Methods
30	Neural Mechanisms of Auditory Processing and Perception	Music Cognition and Cognitive Performance
26	Cultural Evolution and Language Development	Language Evolution and Semantic Change
25	Metaphorical Representation of Time in Language and Cognition	Metaphor and Discourse in Cognitive Linguistics
24	Theoretical Foundations of Linguistic Studies	Semantics and Pragmatics in Language
22	Mechanistic Explanation in Scientific Discovery and Evolution	Scientific Explanation and Philosophy of Science
20	Epistemology and Philosophical Knowledge Studies	Conditionals, Epistemic Modality, and Assertion
17	Automated Detection of Hate Speech and Offensive Language	
17	Computational Complexity and Algorithmic Information Theory	
16	Mechanism Design in Auctions and Procurement Contracts	
16	Learning and Inference in Bayesian Networks	
15	Formal Methods in Software Verification and Control	
15	Semantic Web and Ontology Development	
14	Acquisition and Development of Sign Language	
14	Explainable Artificial Intelligence	
13	Advanced Cryptographic Schemes and Protocols	
13	Audio Signal Classification and Analysis	
12	Dialogue Act Modeling for Spoken Language Systems	
12	Active Learning in Machine Learning Research	
12	Collaboration and Dynamics in Wikipedia Research	
12	Automatic Text Simplification and Readability Assessment	
10	Quantum Information and Computation	
10	Methods and Techniques for Agent-Based Modeling	
10	The Spread of Misinformation Online	