Postdoctoral Position in How Programming Languages Shape Thought

The Knowledge Lab at the University of Chicago seeks to hire outstanding candidates for a postdoctoral research opportunity with support from the Sloan Foundation that explores the degree to which programming languages and data science environments shape how individuals, groups and communities “think”—how they construct code, analyze data and solve computational problems together. The project, titled “The Impact of Programming Languages and Datascience Frameworks on Thinking, Software, and Science” is inspired by the longstanding Sapir/Whorf Hypothesis that natural languages influence how speakers think, which has garnered new evidence with computational methods and large-scale language data. Our project involves analysis of all public GitHub and other code repositories using statistical and machine learning approaches that generate insights linking programming language properties to individual and group behavior to coding and analytical outputs. Based on insights from these large-scale analyses and ongoing surveys of programming communities, we will also generate programming experiments (e.g., with the Jupyter interface) to test whether discovered associations are causal—whether changing languages can predictably improve the efficiency, collaboration, and creativity of coders and coding communities.

Postdoctoral candidates will design and conduct independent research, in collaboration with UChicago Professor and Knowledge Lab Director James Evans, and Gary Lupyan, a computational psychologist from the University of Wisconsin-Madison. Candidates must have substantial computational and data science background and a Ph.D. in Computer Science, Statistics, Applied Math, Sociology or another Social Science, Linguistics, Informatics, (statistical) Physics or a related field, and a strong publishing background.

Specifically, the successful candidate(s) will be responsible for managing and analyzing a massive collection of version controlled source code with Machine Learning (ML) and Natural language Processing (NLP) techniques. Candidates must understand and will need to maintain long running web scraping tasks, via APIs and HTML parsing and have knowledge regarding state of the art in NLP (specifically neural language models, context free grammars and auto encoders), which they will extend to new domains, primarily programming code. This development of new techniques for understanding source code will likely benefit from knowledge of compiler design, static analysis, complex systems and network analysis. Candidates must have knowledge of Python and experience running large scale computational tasks on UNIX systems. Proficiency in multiple other programming languages, including a functional language, will be a benefit. Positions could begin anytime within the coming year, and as early as September 2018. Competitive salary & benefits.

To apply, please send CV and names for letters from at least two references to Candice Lewis, cllewis@uchicago.edu.