

NATIONAL SCIENCE FOUNDATION

Proposal Abstract

Proposal:1936940

PI Name:Cafarella , Michael

Proposal Title: Convergence Accelerator Phase I (RAISE): Simultaneous Knowledge Network Programming and Extraction

Institution: University of Michigan Ann Arbor

Abstract Date: 07/31/19

The NSF Convergence Accelerator supports team-based, multidisciplinary efforts that address challenges of national importance and show potential for deliverables in the near future.

The broader impact and potential societal benefit of this Convergence Accelerator Phase I project will include better use and growth of knowledge networks. Today's knowledge networks, for example Wikidata, include high-quality structured information about a very wide range of topics. Knowledge networks make many new and compelling applications possible, such as structured search engine results and voice assistants. Unfortunately, today's knowledge networks and applications have been very difficult and expensive to construct, making it extremely burdensome to create them for novel topics. This project will take advantage of the research team's expertise in data management, artificial intelligence, and economics to create a combination of software and data that should make novel knowledge network systems dramatically easier to produce. The first effort will be an Economics-focused integrated knowledge network-and-tool system, which has the potential to dramatically improve the ease of performing higher-quality economic measurement and analysis. This knowledge network tool can improve economic-oriented research efforts that will benefit national prosperity. However, the even greater value of the effort will be a tool that allows knowledge networks on any topic to be developed more easily and with less programming expertise.

Although knowledge networks are thought to be key to future data-enabled discovery, knowledge network-driven applications have generally not been developed using a reproducible system. This project will begin to build a knowledge application development system that should make knowledge applications easier to write, existing knowledge networks easier to improve, and entirely novel knowledge networks easier to construct. The team's effort is based on a novel and extremely succinct form of programming that they have developed that allows simultaneous programming and extraction of relevant information to contribute to a knowledge network. The proposed simultaneous programming and extraction system will help construct knowledge networks, but will also improve knowledge network data quality, by providing additional weak supervision for the information extraction pipelines that are commonly used to produce the networks. The system will be tested on real data and users in the Economics domain. However, the methods and tools will not be topic-specific, but rather should be widely applicable to knowledge networks in many topical domains. This work will generate research as well as practical downloadable software, datasets, and applications.

This award reflects NSF's statutory mission and has been deemed worthy of support through evaluation using the Foundation's intellectual merit and broader impacts review

NATIONAL SCIENCE FOUNDATION
Proposal Abstract

Proposal:1936940

PI Name:Cafarella , Michael

criteria.