Proposal Title: Convergence Accelerator Phase I (RAISE): Exploiting Financial and Economic Data - Business OKN
Institution: University of Southern California
Abstract Date: 08/12/19

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 is to lay the foundation for capturing the essential knowledge about businesses, innovation, and markets and to use the latest techniques in computer science to make this knowledge freely available in easily usable forms. The project is a partnership between faculty in business schools and computer science departments and will engage partners in regulatory agencies as well as financial technology companies. The proposed Business Open Knowledge Network (BOKN) will provide the resources necessary for entrepreneurs to fully understand the competitive landscape as they create small businesses, allow regulators to quickly identify issues to help prevent the next financial crisis, and enable researchers to develop and test theories to transform our nation's business practices. Using the BOKN resource, a new generation of students and scholars will be able to blend computational solutions with theories, models, and methodologies from finance, economics, mathematics, and statistics leading to increased understanding as well as broader opportunities for scholarship.

The project efforts to develop the BOKN will require the development of new research approaches that can combine state-of-the-art computational approaches for extracting, representing, linking, and analyzing data with complex and nuanced knowledge about the business domain. The project team will develop business and finance-specific computational tools that can leverage a wealth of unstructured data on the Web, as well as semi-structured data and time series datasets provided for regulatory or legal purposes, and reference datasets with standard identifiers and metadata that enable cross-resource federation. Business expertise will drive these computational tools by defining a concrete ontology of concepts, identifying the key entities of interest, and validating the extracted knowledge and downstream predictions in a series of practical use cases. One expected technical result is the creation of a hybrid knowledge graph that supports traditional symbolic knowledge representation and reasoning enhanced by high-dimensional vector space embeddings capturing temporal evolution and semantic relationships that support machine learning 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 criteria.