Proposal Title: Convergence Accelerator Pilot (RAISE): Credible Open Knowledge Network
Institution: University of Texas at Arlington
Abstract Date: 08/09/19

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 develop new capabilities to ensure the quality and credibility of information collected in large assemblages of data known as knowledge graphs or knowledge networks. Many of today's intelligent software products are powered by massive knowledge assemblages which are often proprietary. Developing an openly available infrastructure based on public data is one of the overarching tracks of the overall Convergence Accelerator Pilot in 2019. The goal of this specific project is to ensure credibility -- integrity, completeness and truthfulness -- in developing such an open knowledge network. As such, this project's efforts are likely to provide insights and value to many of the Convergence Accelerator Phase I efforts initiated in 2019. This project plans to develop a resource for debunking misinformation, which is important to decision making by individuals, organizations, communities, and policy makers. The project's initial use-cases will be healthcare, helping to create methodologies that ensure that health-related data assembled is reliable to support healthcare understanding and decision making, and mitigating security threats from software vulnerabilities by looking at the validity and reliability of indicators of compromise within collections of cyber threat intelligence.

The project supports a multi-institutional and multidisciplinary team that has valuable expertise spanning computer science, economics, journalism and communication, political science, psychology, and public health. The team includes researchers from industrial partners, government research organizations, academic institutions, and international organizations with complementary expertise relevant to information credibility. Phase I of the project will focus on team formation, research planning, and developing proof-of-concept. If the project successfully proceeds to phase II it would develop a sustainable ecosystem of datasets, algorithms, software, as well as a stakeholder community for creating and fully utilizing credibility tools across an open knowledge network.

The team will conduct use-inspired, convergence research on several cross-cutting problems: (1) Modeling credible knowledge graphs, deciding what types of knowledge need to be captured in order to promote credibility, and how such knowledge should be represented in order to enable computational approaches. (2) Developing data-driven understanding of what factors contribute to the persuasiveness of factual statements, what signals help gauge credibility, and how to employ that information to envision countermeasures against misinformation. (3) Designing computational methods that leverage knowledge graphs to vet statements and generate verifiable and reliable

explanations. (4) Procedures and mechanisms for initiating, developing, and maintaining a credible open knowledge network over time. The findings of this research have the potential to significantly impact understanding of information and mis-information creation and consumption and may trigger new lines of investigation while also helping to create more reliable information in knowledge networks generally.

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.