NATIONAL SCIENCE FOUNDATION
2415 EISENHOWER AVENUE
ALEXANDRIA, VIRGINIA 22314

NSF 22-017

Dear Colleague Letter: Encouraging Research on Open Knowledge Networks

November 18, 2021

Dear Colleagues:

Knowledge graphs constitute a powerful approach for organizing, representing, integrating, and accessing data and information from multiple structured and unstructured sources. Large-scale knowledge graphs are now employed at enterprise scales by the private sector. The success of such proprietary knowledge networks offers the potential for open knowledge networks (OKNs) that can support broad public use cases and can be shared in a manner that is secure and preserves privacy. In other words, tools and knowledge bases that would be freely available could reasonably address a broad set of challenges of national importance, spanning science, engineering, health, and commerce. Indeed, the recently–released report from the National Security Commission on AI stated, "Open knowledge networks (or repositories) with massive amounts of world knowledge could fuel the next wave of [artificial intelligence] exploration, driving innovations from scientific research to the commercial sector."

With this Dear Colleague Letter (DCL), NSF invites principal investigators (PIs) to submit proposals that contribute to OKN research and practice. Proposals can be submitted to core programs as Small projects within the Directorate for Computer and Information Science and Engineering (CISE). This spans the following divisions and programs: Computing and Communication Foundations (CCF) and Information and Intelligent Systems (IIS) divisions and the Office of Advanced Cyberinfrastructure (OAC), as well as the Secure and Trustworthy Cyberspace (SaTC) program.

This DCL builds on work on OKNs by the National Science Foundation's (NSF's) Convergence Accelerator program. Recently NSF funded a cohort of OKN projects focused on addressing convergent research challenges associated with OKNs (see the Convergence Accelerator Portfolio for details of funded projects). While these projects are working to create OKNs that address user needs in specific domains, they have also identified additional research directions associated with creating, scaling, and effectively using OKNs. These

research directions include but are not limited to the categories below.

OKN AUTHORING AND LEARNING

- Automatic generation of data workflows for ingesting data into knowledge graphs from specific sources;
- Automated learning of knowledge graphs from heterogeneous and distributed data sources; and
- Programming systems and integrated software development environments that incorporate data workflows, data provenance, social/collaborative programming, and user-friendly interfaces as native capabilities.

OKN REPRESENTATION

- Representation of information and knowledge at different levels of abstraction to serve diverse communities of users and match the level of sophistication and cognition of diverse end users/applications;
- Consistent representation of metadata, context, data, and data-driven rules in a common scheme, transparent to the user;
- Combining open as well as privacy-sensitive or access-controlled information within a single, common knowledge graph scheme;
- Common schemes for combining data and computational models within a knowledge graph, and for addressing representation and propagation of heterogeneous error bounds, data quality and privacy risks; and
- The ability to track provenance, and support transparency and accountability.

OKN USES

- Automated, privacy-aware and semi-automated query processing and inferencing, using logic-based systems as well as machine-learning approaches, including foundational machine learning, including with humans-in-the-loop, for deriving and interpreting complex information and knowledge;
- Question-answering with knowledge graphs, including introspection/interrogation of the information represented in the graph; and
- Secure, privacy-preserving and robust mechanisms to adapt knowledge graphs as information changes over time.

OKN SCALABILITY

- Scaling to very large, online, real-time scenarios;
- Interacting with very large knowledge graphs including graph visualization techniques, abstraction of content (what is contained in the graph), and abstraction of query

capability (what you can ask of the graph); and
- Cyberinfrastructure to support secure, scalable, performant knowledge graph implementations.

This DCL is not a special competition nor a new program. Prospective PIs should meet all requirements associated with the Small size class of the CISE core programs solicitation to which they are responding. Additionally, to call attention to responsiveness to this DCL, project summaries should include the acronym "OKN" in the keyword list. Note that some programs may already require specific keywords to be used; in such a case, please use “OKN” as the last keyword used.

Proposals responsive to this DCL will be reviewed with other Small proposals submitted to CISE’s core program solicitations and in accordance with NSF’s merit review criteria as well as any additional solicitation-specific review criteria identified in the corresponding solicitation.

Please note that the proposal submission window for Small projects to the OAC Core Program is December 1 – December 22, 2021. Other CISE Core Programs accept proposals anytime. However, for full consideration under this DCL in fiscal year (FY) 2022, proposals should be submitted before **April 15, 2022**.

For further information about proposing pursuant to this DCL, PIs may contact:

- Hector Munoz–Avila (CISE/IIS), hmunoz@nsf.gov
- Tevfik Kosar (CISE/OAC), tkosar@nsf.gov
- James Joshi (CISE/CNS, CISE/SaTC), jjoshi@nsf.gov
- Peter Brass (CISE/CCF), pbrass@nsf.gov

For questions about previous Convergence Accelerator–funded OKN projects, PIs may contact:

- Lara A. Campbell (Convergence Accelerator), lcampbel@nsf.gov

Sincerely,

Margaret Martonosi
Assistant Director, CISE
National Science Foundation