Project Summary—*Collaborative research: Elements: Shared data-delivery infrastructure to enable discovery with the next generation of Dark Matter and Computational Astrophysics experiments*

Overview

We propose to build infrastructure that allows analysis across experiments in two fields whose next-generation science goals can only be achieved with unprecedented coordination: direct-detection dark matter searches and computational astrophysics. The proposed infrastructure will not require experiments to adopt common data formats or move from their existing data catalog infrastructure. Instead, collaborations will need to describe their data format and their existing data catalog API. Collaborations will be able to leverage funded support from our team to do this work.

The PI and Co-PI on this proposal are scientists who will actively use the proposed infrastructure. PI Roberts is a member of the Super Cryogenic Dark Matter Search collaboration and serves as the Data Quality Technical Coordinator as well as the Software Working Group Chair. She leads the push for advanced cyberinfrastructure that supports an accessible, web-based analysis environment for the SuperCDMS collaboration. Co-PI Turk has a background in computational astrophysics and currently leads the development of `yt`, a volumetric analysis tool that is used in multiple domains for data investigation. Both Roberts and Turk are uniquely situated to develop the proposed infrastructure and use it for active analysis.

Intellectual Merit

Dark Matter: For the next generation of dark matter experiments to be able to discover dark matter, experiments have to understand their backgrounds to an unprecedented level. This will require advances in simulations, detector characterization, and maximizing experimental constraints. Combining data from detectors that are located in the same place can provide much-needed insight on the identity of possible dark matter signals. But this is currently impossible because the data formats and analysis software are incompatible.

Astrophysics: To realize the science accessible through multi-messenger astrophysics, it is critical that astrophysical simulations and experimental data be accessible for easy cross-analysis. `yt` has made strides toward this and the proposed work would lower the barrier to including new data sources, increase the maintainability of the `yt` codebase, and increase the discoverability and accessibility of data sets across the field.

Broader Impacts

There is a clear need for broadening participation in research at the undergraduate level, as well as increasing data and computational literacy for STEM majors [1]–[3]. Accessible tools are a recommended practice for broadening participation in data science [2]; these web-accessible tools lower the barrier to entry for all students. Accessible entry to research is particularly critical for students from underrepresented groups—groups that are often less likely to use university resources [4]—and the existing analysis workflows in both dark matter and computational astrophysics require students to be willing to demand significant amounts of an experts' time.

The fundamental goal of the proposed work is to make data easily accessible to individuals who want to answer science questions. We anticipate increasing the participation and productivity of undergraduate students in research with the Cryogenic Dark Matter Search collaboration and expanding this accessible platform to additional dark matter experiments, computational astrophysics, and other data-intensive science domains.