# Hypernets III Statement of Work[[1]](#footnote-1)

December 8, 2020  
**Team:** Cliff Joslyn (PI), Brenda Praggastis (tech lead), Dustin Arendt, Tony Liu (WSU), Jin Young Yun, Andrew Lumsdaine

**Proposed budget:** $300K for a 12 month period of performance

The HyperNetX (HNX) v3.7 Python library[[2]](#footnote-2) contains core hypergraph methods including hypernetwork analytics, computational topology methods, and visualization. The focus of HNX is on exploratory data analysis of hypergraphs at human scale, with visualization support. Metrics, collapsing, and components can be run on hypergraphs of modest size up to O(10K) nodes and edges, with visualization recommended for sub-hypergraphs O(100) nodes and edges.

With HNX successfully released and reported in multiple publications, we are now preparing to proceed with the next steps needed to serve sponsor mission. Below we describe four tasks for new capability development, but first describe a critical overall prior task. Our deliverable will be a full release HyperNetX v1.0 with incremental releases to follow with full functionality including:

* Robust and generic use case exploration capability with the ability to pivot between multiple views on the same data set
* Improved scaling using state of the art HPC and Python acceleration techniques
* Topologically driven gap identification
* Addition of more *hypernetwork science* methods and metrics inspired by traditional network science methods and metrics
* Greatly improved interactive visualization capability for modest sized (sub)hypergraphs

## Task 1: Community Interaction and Applications

* Work closely with the sponsor to fully support and integrate HNX into their workflows, including maintenance of the HNX open source repository: bug fixes, feature requests, and robust test cases.
* Work with the user community at the sponsor and closely related organizations (e.g. TIMC, LAS).
* Work with the sponsor to perform validation and verification of the methods and software on their own mission-relevant data sets, including incorporating prior efforts in analyzing cyber data sets involving DNS and malware data.
* Work with the scientific and academic communities to foster technical communication and joint technical development.

## Task 2: HNX Overall

Continue to advance new methods in hypergraph analytics, releasing upgrades incrementally on GitHub and PIPy. Potential targets include:

* Directed hypergraphs
* Expansion of centrality metrics for undirected and directed hypergraphs
* Define and implement s-clustering coefficients
* Hypergraph generators
* Data Cube Interaction
* Domain specific HNX modules

## Task 3: High Performance Computing

Focus is on development of a higher performant HNX in these steps:

* Establishing a static hypergraph class in addition to the current dynamic class.
* Using PyBind to link the NWHy C++ hypergraph classes for back end support of shared memory scaling
* Develop efficient algorithms for s-metrics and homology in NWHy for import into HNX
* Explore multi-node scaling of HNX using PNNL HPC projects such as SHAD/SHADES or an existing OpenSHMEM Implementation.

## Task 4: Human-Computer Interaction

Focus is on improvements to HNX interfaces.

* Create seamless interface within the Python-Jupyter environment to query and analyze hypergraphs at scale.
* Implement interactive jupyter widgets for visual exploration of data
* Develop domain specific jupyter notebooks

1. The material contained here is submitted for informational purposes and is not binding on Pacific Northwest National Laboratory, or the U.S. Department of Energy. Binding commitments can only be made by the submission of a formal proposal which sets forth a specific Statement of Work, estimated cost, and contract documents, and which is signed by a Pacific Northwest National Laboratory Contracting Officer and transmitted by the U.S. Department of Energy. [↑](#footnote-ref-1)
2. <https://github.com/pnnl/HyperNetX> [↑](#footnote-ref-2)