



**Pacific
Northwest**
NATIONAL LABORATORY

A Network of Networks for User Facilities in 2050

Community of Interest Workshop on Future Scientific
Methodologies

February 2020

Douglas M Mans
Director, Environmental Molecular Sciences Laboratory

U.S. DEPARTMENT OF
ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <<https://www.ntis.gov/about>>
Online ordering: <http://www.ntis.gov>

Current State:

DOE SC currently operates 27 national user facilities that span the Office of Science. Together these facilities are foundational for helping support and execute the research critical to delivering on the DOE SC mission. Most of these facilities are large monolithic entities where a core of deep scientific expertise is localized and science from the research community is brought to the facilities. These facilities by and large operate as independent entities where the research community engages in a one-on-one and often one-way set of interactions with individual facilities until the breadth of capabilities needed to perform the science inquiry is obtained.

The Challenge:

The pace and capacity demand of scientific research is increasing rapidly as the challenges DOE SC address become increasingly more complex. The ever-increasing velocity and volume of experimentation is juxtaposed against (a) a core tenant of DOE SC to continually develop and provide to the science community the most advanced cutting-edge capabilities within the user facilities and (b) a rapidly accelerating commoditization of high-tech instrumentation. This creates a dual-pronged bottleneck for advancing science. On the one hand the advanced instrumentation available only at the user facilities requires dedicated experts to maintain, tune, operate and interpret the data and are inherently low throughput both in pace and capacity. Whereas the commoditized high-tech instruments provide enhanced capacity and pace for researchers but result in a plethora of data and metadata standards, types, quality and veracity issues. As multiple independent researchers acquire these capabilities reproducibility challenges are further compounded slowing science advances. This silo'ing also represents an inaccessible massive big data opportunity through the network of independent experimental silos.

Future State:

The user facilities of the future operate as a seamless fluid entity with each other and the research community. Researchers can remotely access any of the advanced capabilities housed across the user facilities and the user facilities themselves have diffused out into the research community creating a network of networks of user facilities. To accomplish this ASCR and the other offices in SC have joined forces to enable three key computational aims:

AIM 1: Realize a fully automated and remotely accessible facility through rigorous and purposeful application of AI and ML development combined with advanced automation controls concurrent with the development of advanced capabilities to enable 24/7/365 computer operation to provide greatly enhanced accessibility at a reduced operational cost. A key aspect that the user facilities bring is performing complex custom experiments. As part of augmenting an automated facility for routine work, the ability to remotely interact in real time via robotic assistants would be transformational in allowing real-time experimentation with geographically dispersed researchers.

AIM 2: Realize a network of networks of the user facilities' operational envelop ensuring experimental veracity and data provenance through creation of AI controlled remote, field-deployed and satellite instrumentation all connected through an advanced wireless communications network that continuously transmits and receives protocols, methods and operational standards (including for example automated data pre-processing workflows, meta-data capture through instrument situational awareness and data compression) and automatically determines which analyses, sampling or measurements to perform next.

AIM 3: Realize a seamless, lag-free computational capability to facilitate near-instantaneous data retrieval and processing from the remote/satellite instruments as well as all user facilities via a mesh network of ESNNet 400gbps (or higher??) optical lines paired with advanced 6G wireless networks to ensure every instrument connected has the ability to serve as a distributed data repository node alongside a set of super data archive facilities. The mesh network would provide connection directly to the LCFs as well as the mid-range HPC facilities supporting the individual SC offices such as EMSL for BER. This would enable maximum efficiency and velocity in data analyses (near real time data processing with complete remote interactions) as well as facilitate real-time exascale applications through instant access to massive data sets obtained from 1000's of "always-on" distributed instrument data repositories alongside the super data archives.

By accomplishing all three aims the user facilities of the future will be fundamentally transformed in their operation and scope to enable the hyper multi-disciplinary, hyper velocity and hyper capacity of experimentation and modeling required to fully deliver the DOE SC mission of 2050.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov