EESSD - ESS Cyberinfrastructure Working Groups

Current Scope and White Paper Plans

Data Management

Management and Archival of DOE climate and environmental datasets

- Data **Preservation, Sharing, and Publication**
- Common Data and Metadata **Standards**
- Data **Citation and Attribution**
- Data **Federation** across different data catalogs

Data **Synthesis** across ESS and other relevant Datasets

Development of common **Tools** for data usage

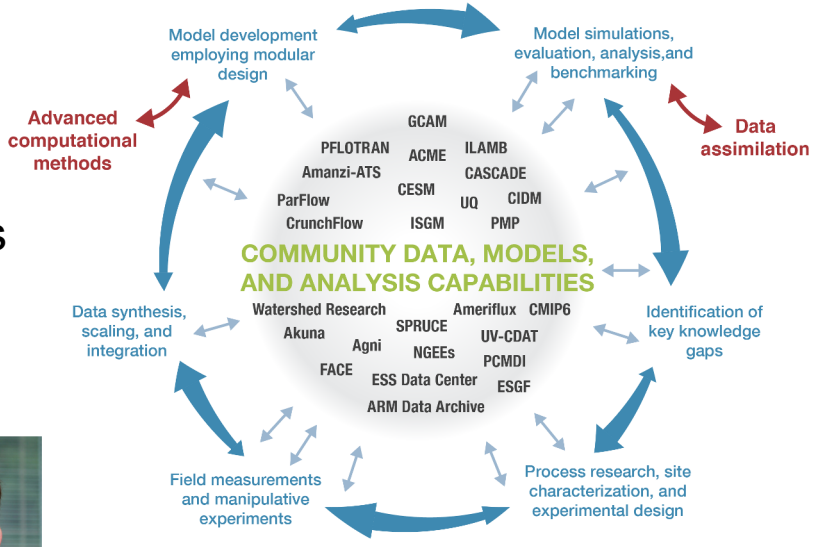
QA/QC, processing, analysis, mining and visualization data to prepare them for use in new research projects.

Leads: Danielle Christianson (LNBL)

Terri Velliquette (ORNL)

Model–Data Integration

- Model–data comparison, model evaluation and benchmarking
- Uncertainty quantification (UQ) and data assimilation (DA)
- Management of model results and observational data (with Data Management Working Group)
- Geospatial and remote sensing data analysis
- Data analytics methods and techniques, e.g.,
  - Data mining
  - Machine learning, neural network models
  - Genetic algorithms
  - Visual analytics
- Hybrid ML/process-based models & data-driven models



Co-Leads:

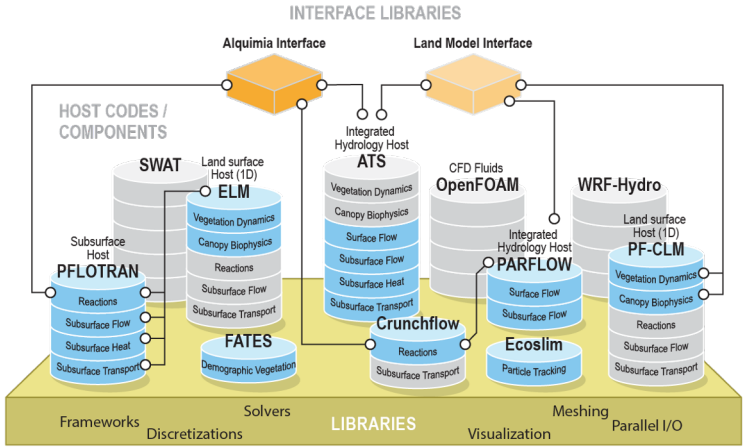
Xingyuan Chen (PNNL)

Forrest M. Hoffman (ORNL)

Software Engineering & Interoperability

A high quality community driven open-source software ecosystem of interoperable components that can be assembled in flexible configurations within a common framework supporting ModEx and the Virtual Laboratory:

- integration of legacy and new capabilities across projects
- rigorous – but rapid – testing and validation of model-data integration capabilities
- changing architectures and programming models
- complex multiscale models (coupling, interoperability)
- performing quantitative and formalized UQ
- diverse interdisciplinary teams, and training
- increased scientific productivity



Co-Leads: David Moulton (LANL) Ethan Coon (ORNL)

Our Vision From the 2016 White Paper

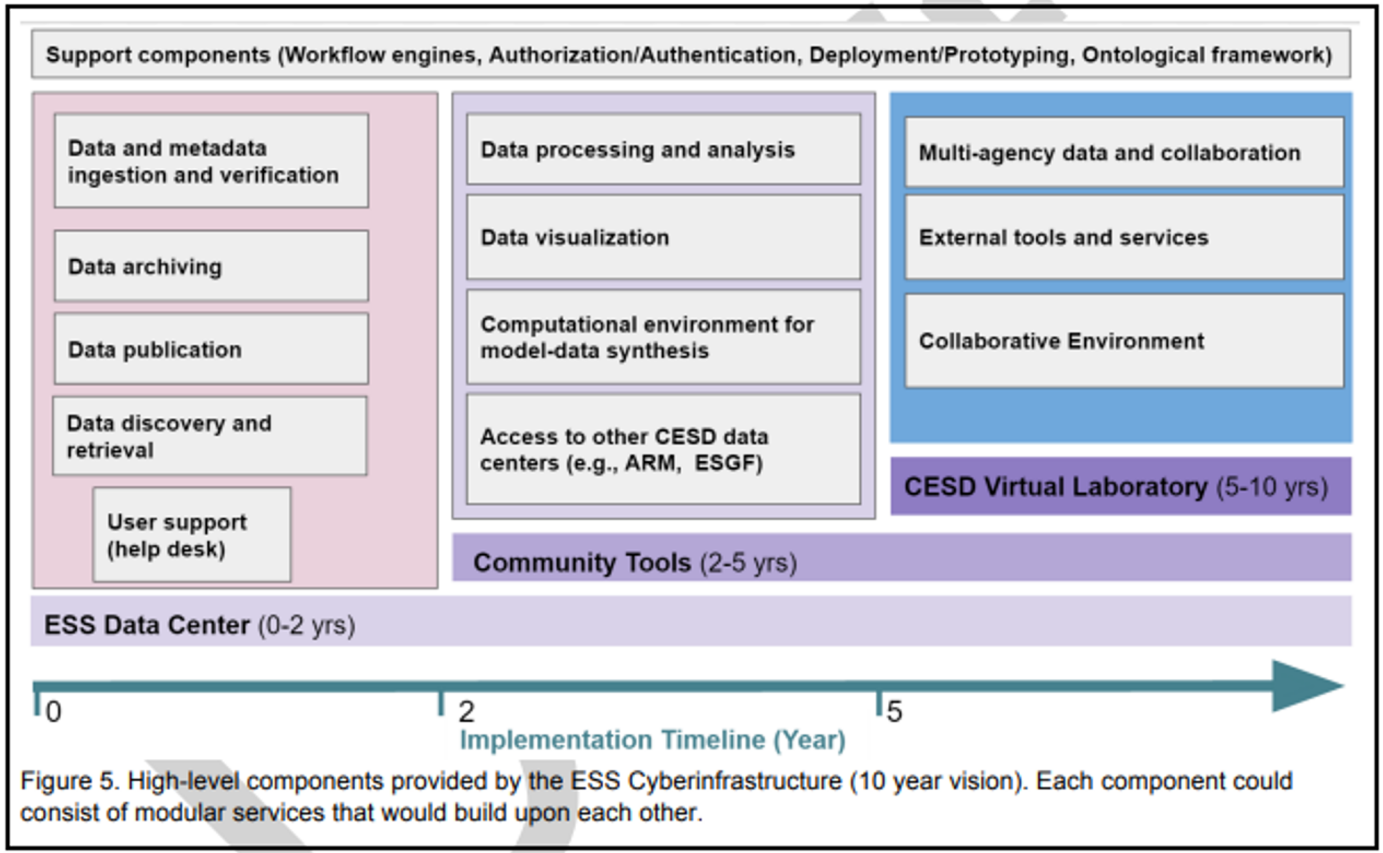


Figure 5. High-level components provided by the ESS Cyberinfrastructure (10 year vision). Each component could consist of modular services that would build upon each other.

In the years since that White Paper

What has Changed?

- Advances across a wide range of disciplines, agencies and projects.
- New advisory board and workshop reports
- Greater collaboration across projects, increased potential for shared resources, workflows, and tools
- View of cyberinfrastructure is expanding beyond data to include modeling and model data integration.

What has been Accomplished?

- Updates from each working group this afternoon

What is next?

We need to create a new white paper which lays out a vision for next 10 years in phases (e.g., 0-2 years, 2-5 years, 6-10 years)

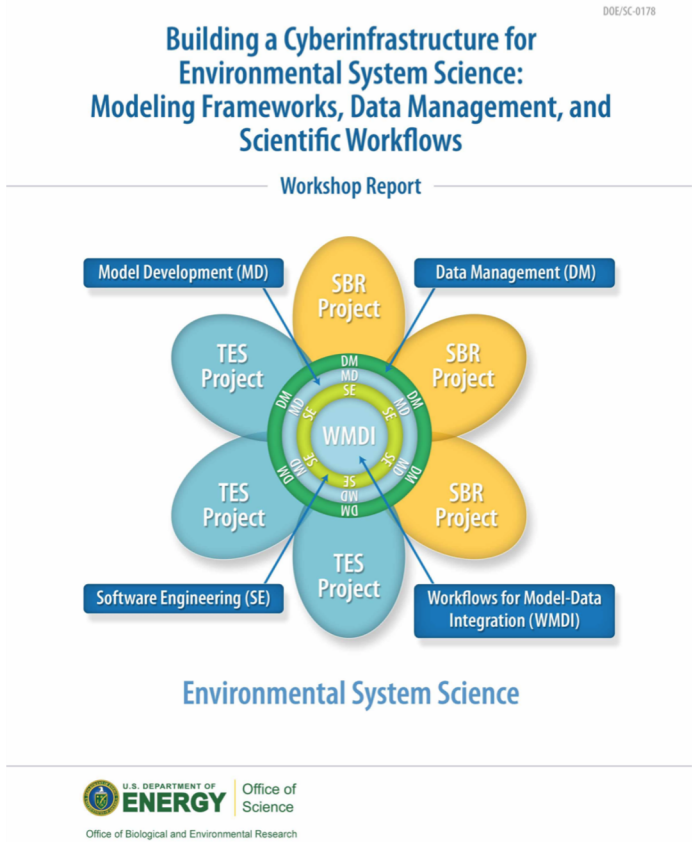
- Informed by talks this week and previous work
- Collaborative and open effort
- Four step effort
  - Provide input/thoughts (**today and next week**)
  - Dedicated writing/summary effort by Writing Team building on this input
  - Review process with the community and tweaking
  - Delivery to program managers (mid-summer)
- Input requested on aspects like
  - Overall concepts, requirements (functionality), priorities
  - Technical details, detailed roadmap
  - Business model
  - ….

Questions?

Working Groups Structure

Following recommendations of the workshop report and feedback from 2015 AGU Town Hall the Working Groups Kickoff was in 2016

- Established an Executive Committee (EC)
  - past and current membership on backup slide
- Established Three Working Groups
  - Data Management (DM)
  - Model–Data Integration (MDI)
  - Software Engineering and Interoperability (SEI)
- Reporting on Working Group Activities
  - Informal reporting to SC PMs and EC PIs
  - Annual reporting at ESS PI Meeting
- Established Annual Meeting
  - Significant part of the success of these meetings has been the informal face-to-face time.



https://doesbr.org/ESS-WorkingGroups/ESSWG_WorkshopReport-final.pdf

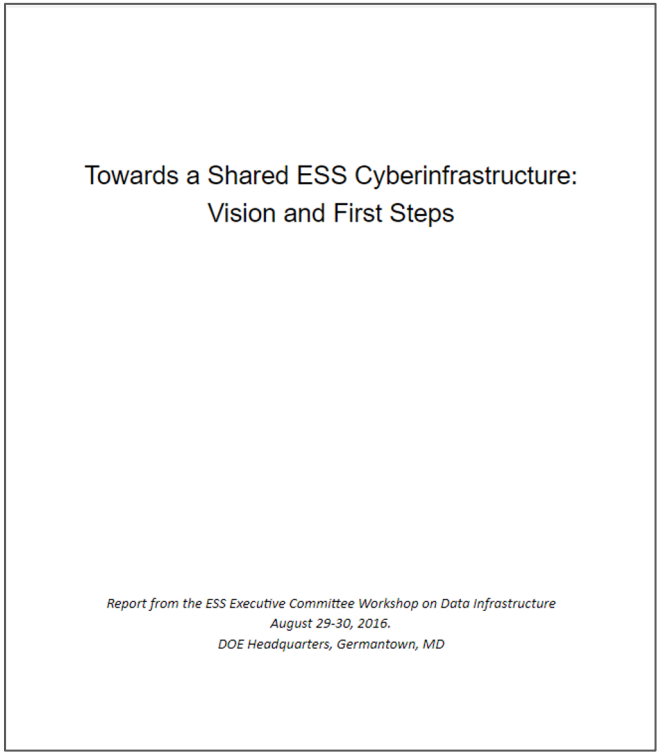
White Paper on Data Infrastructure

Motivation

- *"the innovation most needed is a framework that allows seamless integration of multiscale observations, experiments, theory, and process understanding into predictive models for knowledge discovery"* (BERAC 2013)
- Exponential growth in the amount, variety and complexity of scientific data.
- Significant fragmentation across projects and disciplines still remains.
- A need for a data center that would be a foundational part of a community cyberinfrastructure.

Outcome

- Report used as a reference in the call that led to ESS-DIVE.


Towards a Shared ESS Cyberinfrastructure:
Vision and First Steps

Report from the ESS Executive Committee Workshop on Data Infrastructure
August 29-30, 2016.
DOE Headquarters, Germantown, MD

https://doesbr.org/BERfiles/Towards_a_Shared_ESS_Cyberinfrastructure.pdf

ESS Cyberinfrastructure

- Concepts emerging in last ten years
- Formalized in multiple workshops and reports
- Right now need to update vision in community developed new white paper which should be provided to program managers by the end of June 2020