EMSL

Data Ecosystem

Integrating Temporal and Spatial Scales

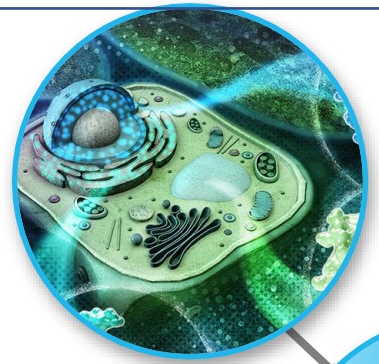
Ratna Saripalli
Chief Data Officer

Agenda

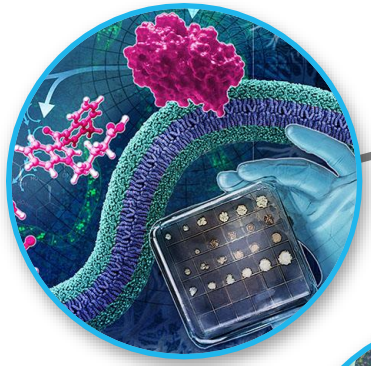
- Science Areas
- User Personas
- Instruments
- Data
- Challenges
- Data Ecosystem Framework

EMSL Science Areas

Functional & Systems Biology



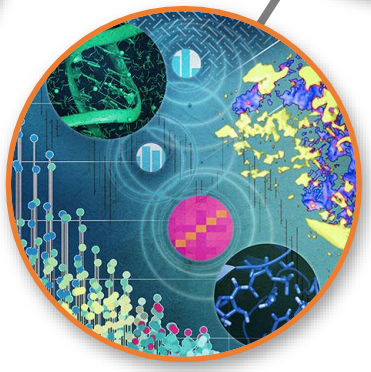
Structural Biology



Biomolecular Pathways

Cell Signaling & Communications

Computing, Analytics & Modeling

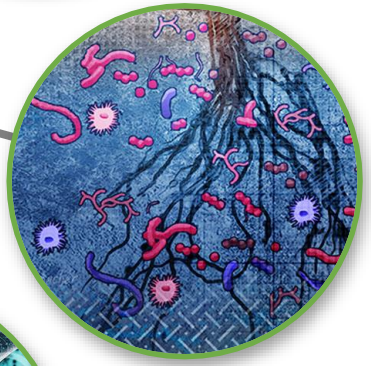


Systems Modeling

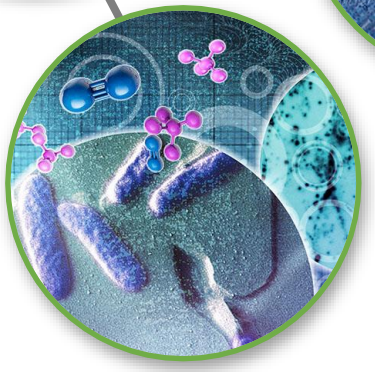
Data Transformations

Environmental Transformations & Interactions

Terrestrial-Atmospheric Process



Rhizosphere Functions



Bio-geo-chemical Transformations

User Personas

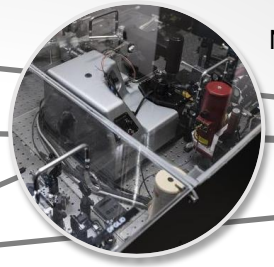
- Structural Biology
- Biomolecular Pathways
- Cell Signaling & Communications
- Terrestrial-Atmospheric Process
- Rhizosphere Functions
- Bio-geo-chemical Transformations
- Systems Modeling
- Data Transformations

Instruments

Optical Coherence Tomography

Raman Atomic Force Microscopy (AFM)

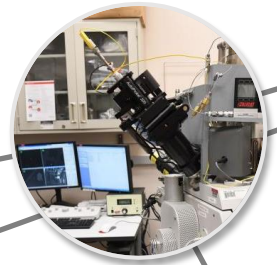
Nano-FTR



Biogeochemical Transformations

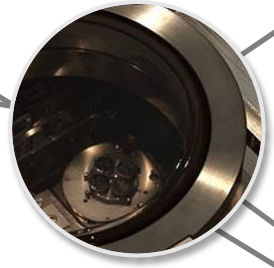
Cryogenic Transmission Electron Microscopy (TEM)

Environmental Scanning Electron Microscopy (ESEM)



Structural Biology

Liquids NMR for Organic Matter/ Complex Mixtures



Helium Ion Microscope

X-Ray Diffraction

Mossbauer Spectroscopy

Soil Incubation

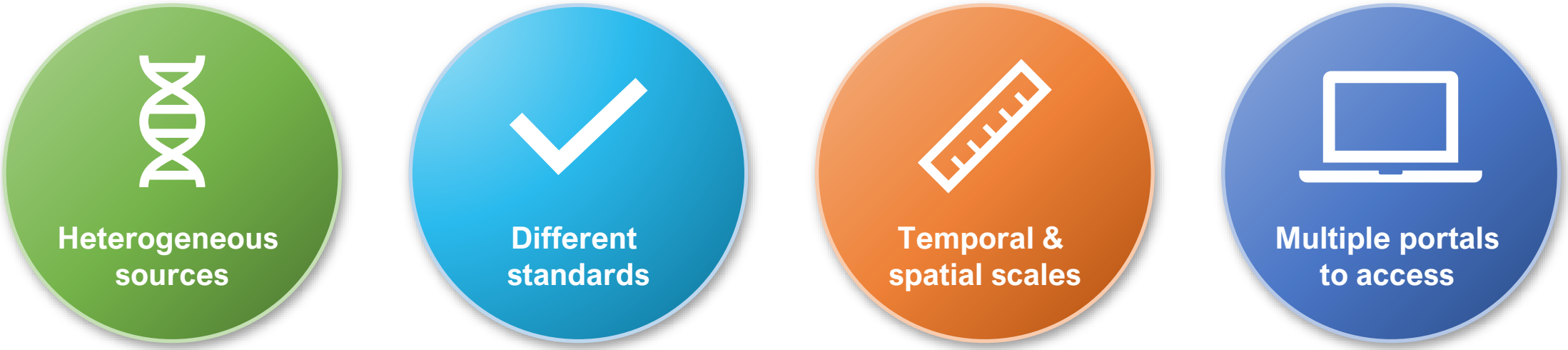
Dispersive X-Ray Spectroscopy

Electron Paramagnetic Resonance

Terrestrial-Atmospheric Processes


Imaging
Metadata
Text
Table
Video
Time
Audio
Modalities
DATA
Heterogeneity
Sources
Instruments
Users
Study
Researchers
APIs
Portal
Findable Accessible Interoperable Reusable

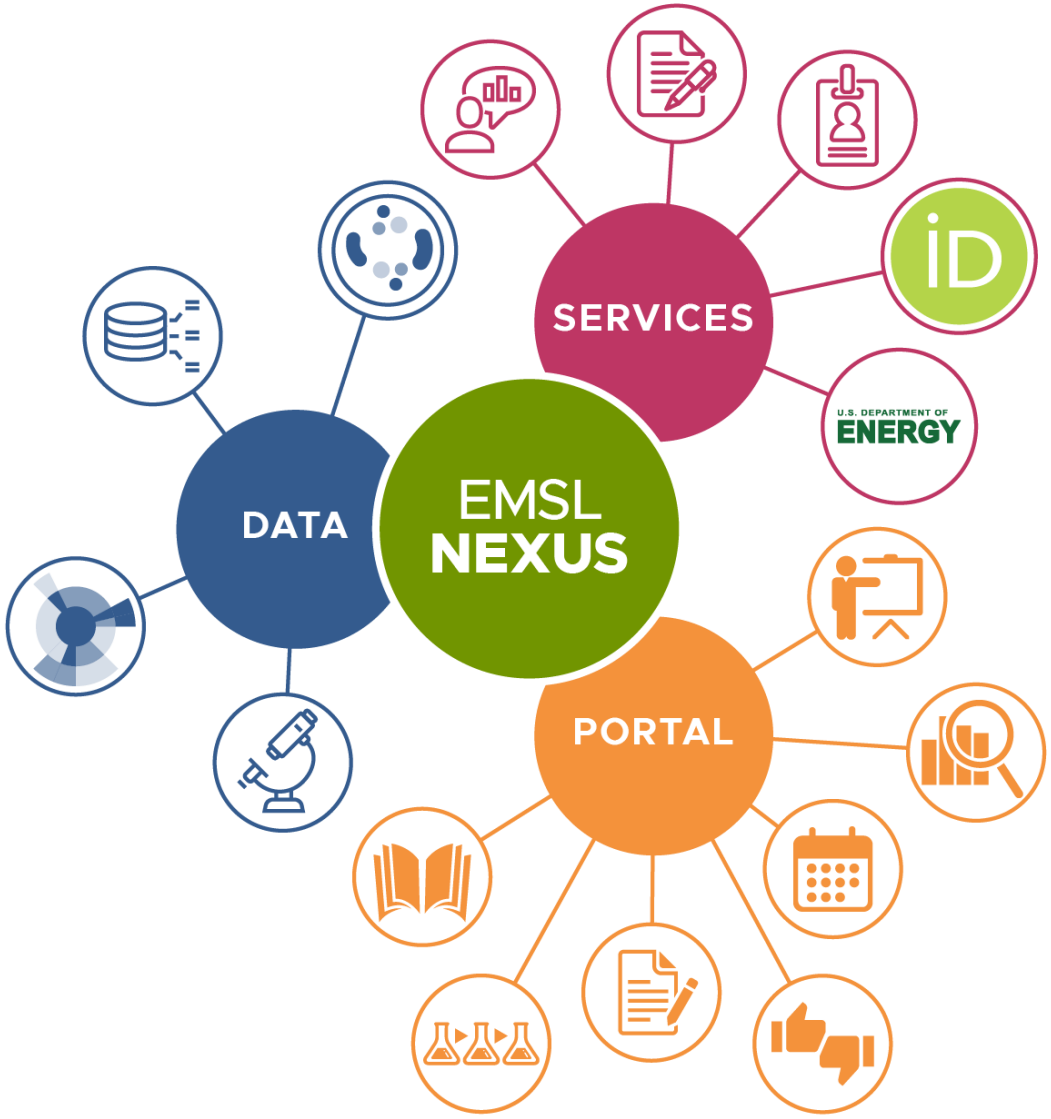
Challenges



Solvable: Intentional designed data ecosystem

EMSL Data Ecosystem Design Framework

- FAIR Design principles
- Data Integration
- Data governance
- Metadata management
- User-centered design
- Search & Discovery
- Collaboration & Sharing
- Data Visualization



CAM: Computing, Analytics & Modeling

Translational modeling

Advances in

Agricultural Sciences
AI/ML
Biomedicine
Computational Math (UQ)

New Capabilities for BER Science

Data-driven modeling

Enabling automation

Data products

Embedding modeling capabilities

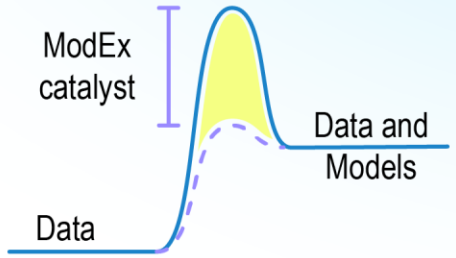
SYSTEMS MODELING

DATA TRANSFORMATIONS

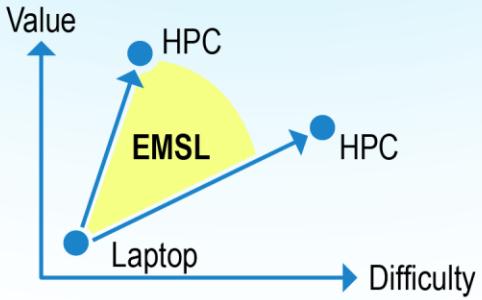
COMPUTING PLATFORMS

Analysis, integration and visualization

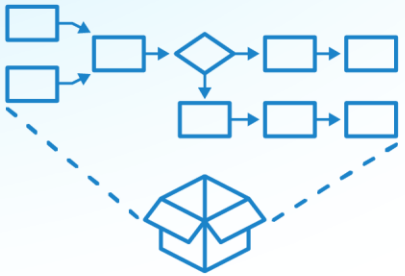
ModEx app and data store



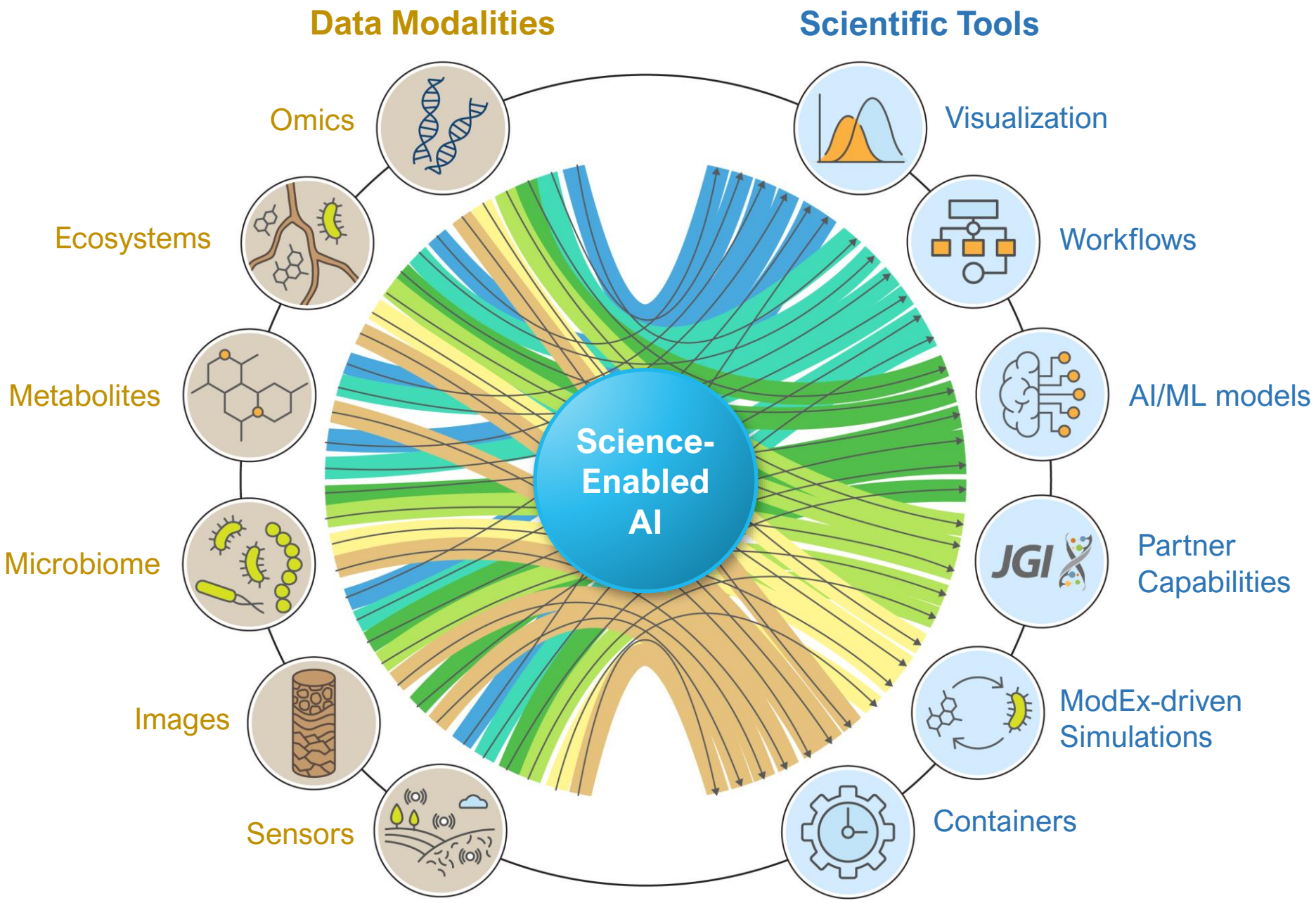
Scale beyond laptops



Embedding workflows

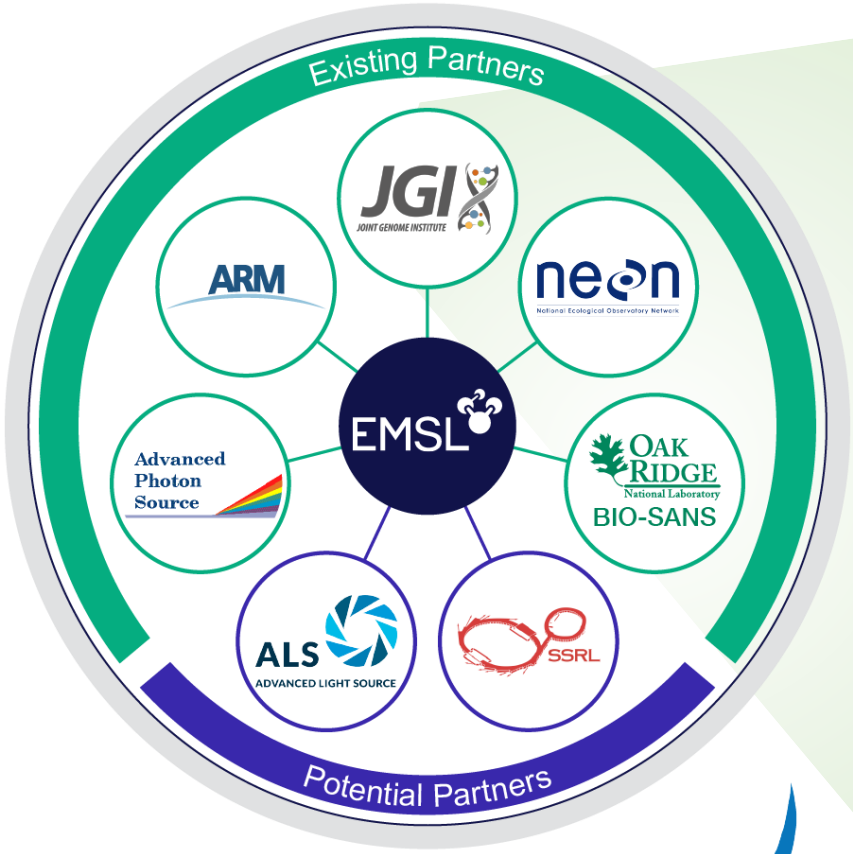


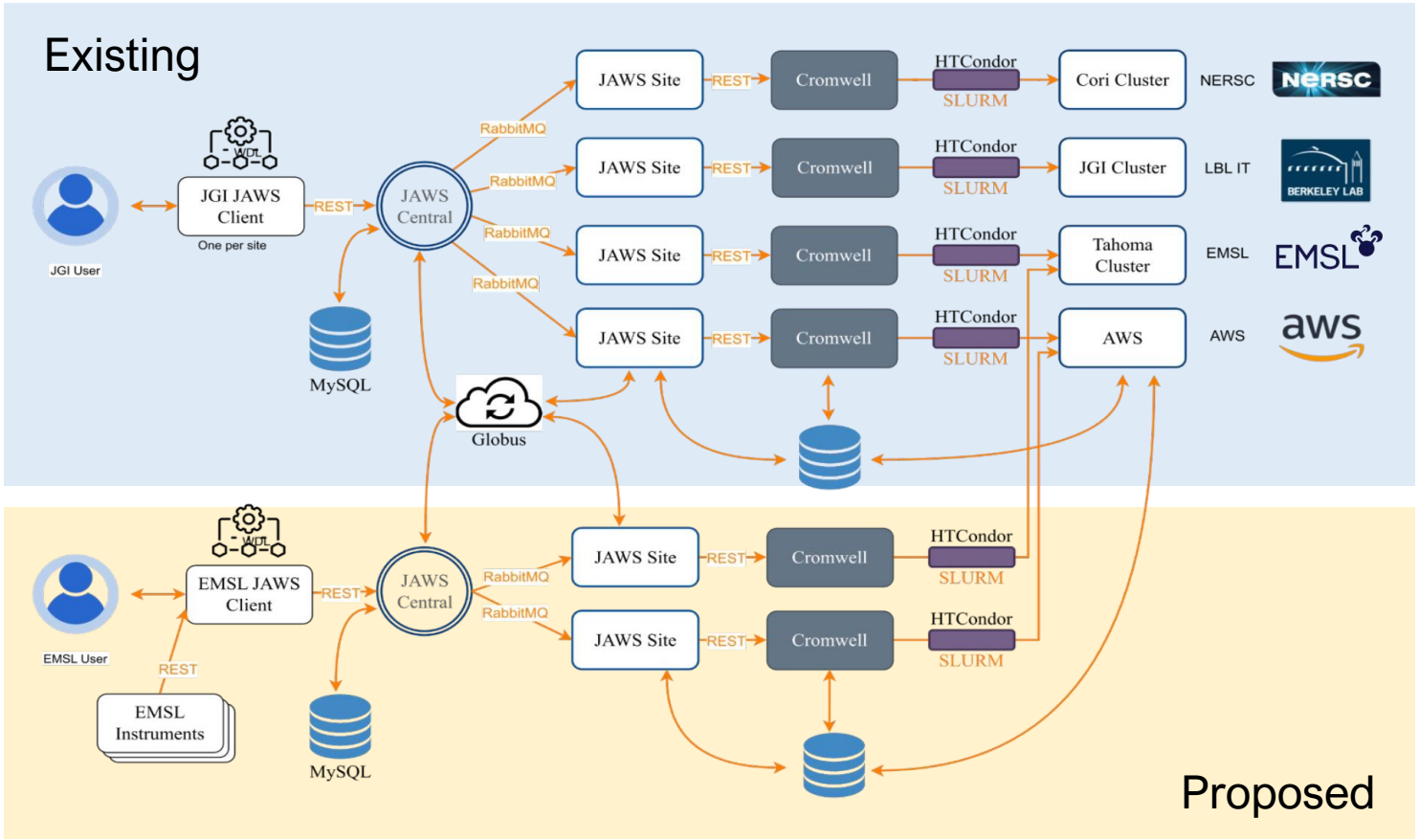
Lowering Barriers through: Scale-up • Reproducibility • Sustainability • "On ramps"



Source reference: https://www.nature.com/articles/s41591-022-01981-2/figures/1

Accelerators





KBase

nmdc
National Microbiome Data Collaborative

ESS-DIVE

Enable new scientific discovery

- Easier access, understanding, & reuse
- Increased efficiency
- Increased data quality
- Enabling innovation
- Enhanced compliance


Thank you!
Q&A