CATEGORY III: PLANNING FOR A NATIONAL SCALE DATA PLATFORM FOR PERSONAL DATA

NSF Award # 2609498

University at Buffalo The State University of New York

Team

Varun Chandola
UB CSE

Joe Grupp
UB CISO

Sharon Hewner
UB SON

Matt Jones
UB CCR

Kenneth Joseph
UB CSE, AIS

Danial Smith
UB SON

Erica Smith
UB SON

External Collaborators

MIMIC

ICPSR

HEALTHeLINK™

Pegasus
Workflow Management System

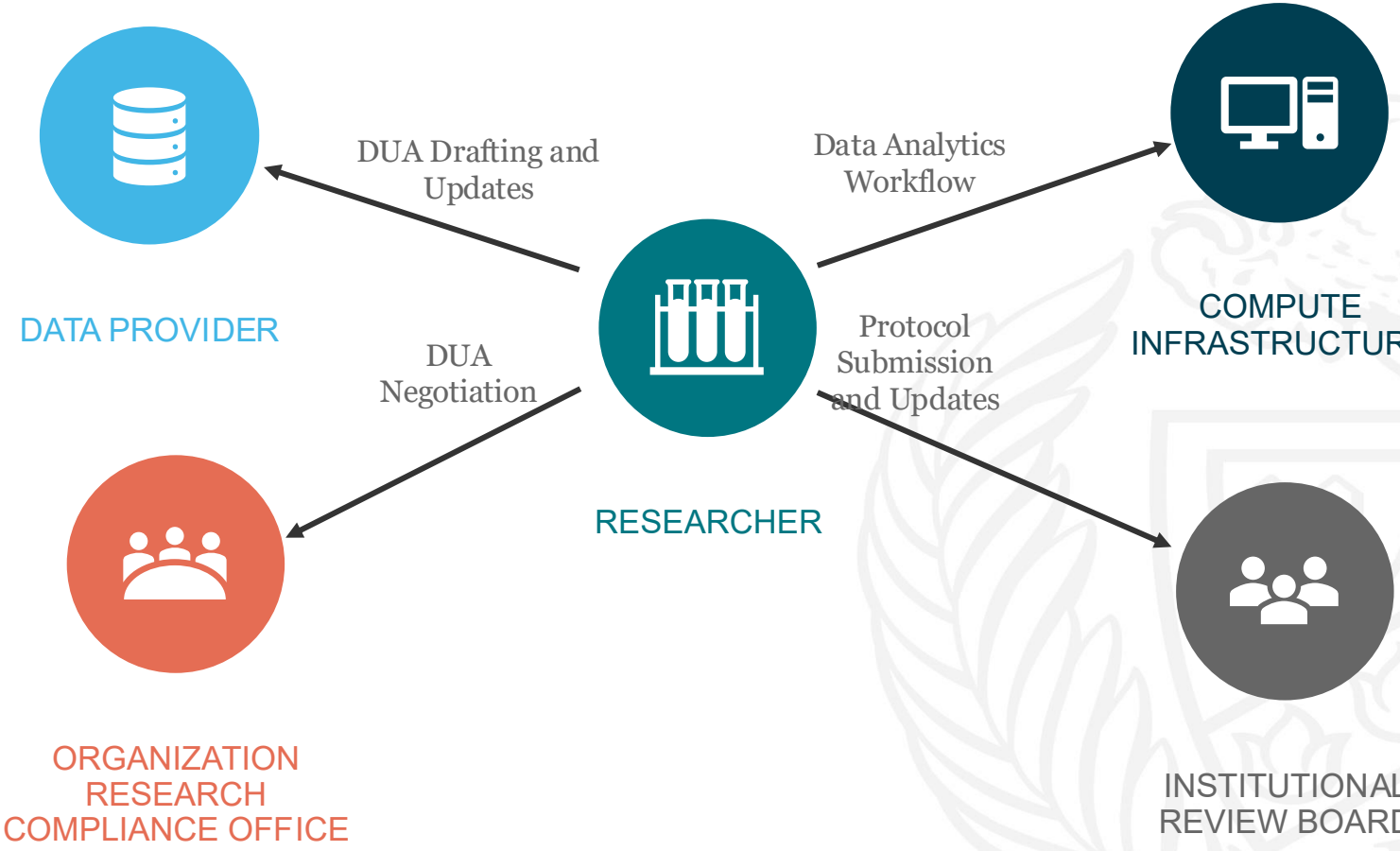
NSDF
National Science Data Fabric

Objective – Reducing the Time to Science with *Sensitive* Data

- A national scale data platform to support researchers working with ***sensitive*** data
  - Personal data
  - Or any data that carries significant legal, financial, or ethical risk, if exposed
- Streamlining data use agreements, user management, data anonymization workflows within secure and compliant enclaves
  - At low cost

Personal data - *any research data asset that contains direct identifiers, indirect attributes, or behavioral patterns that can be linked back to a specific living individual or sensitive sub-population.*

Key Entities and Interactions



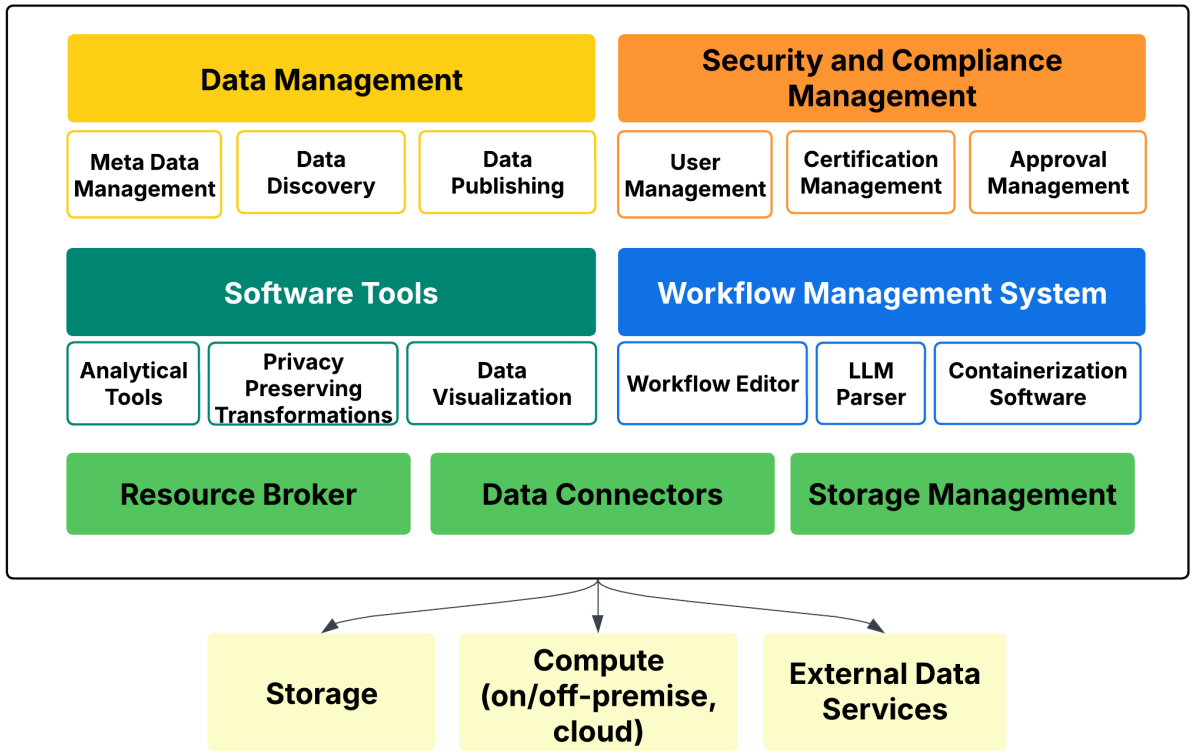
Key Challenges – Streamlining Interactions at Low Cost

- Diverse needs of a broad research community
- Support for regulatory standards
  - HIPAA, CCPA, SOC2, GDPR
- Supporting regulatory boards and data providers
- Integrating compute with data
- Integrating with national data CI ecosystem
- Sustainability

Key Capabilities

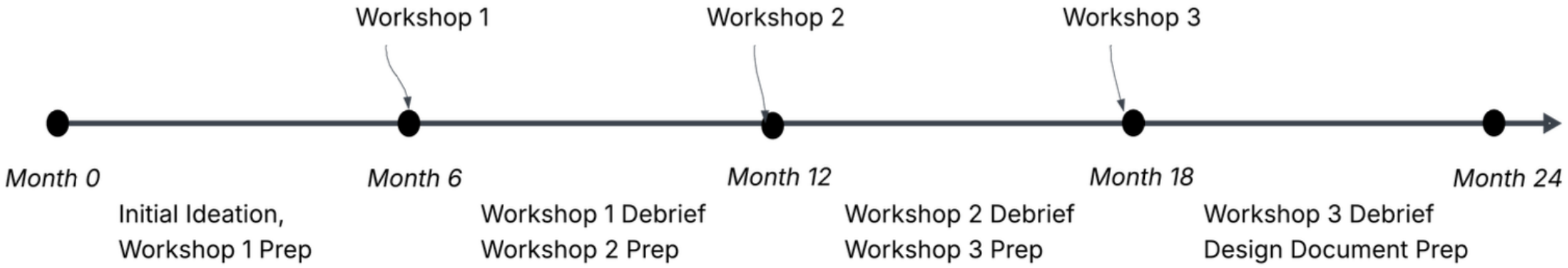
- Provisioning **secure and compliant enclaves** for a research team working with sensitive data
- Secure user and data management tools for data stewards
- Secure workflow management tools, deidentification workflows and other analytical layers
- Workflow deployment on underlying compute
- Connectors to external data and analytics endpoints
- Open marketplace to share results with larger community

Preliminary Design



Timeline

- Planned workshops



Workshop stakeholders:

- Workshop 1 – Researchers
- Workshop 2 – Data Providers and Institutional Review Boards
- Workshop 3 – Compute providers and IDSS/NSF CI Ecosystem

Thanks for listening

- We welcome your thoughts on how we can collaborate with and benefit from the larger IDSS ecosystem

chandola@buffalo.edu