

Data Curation

Sayed Choudhury
Johns Hopkins University
GWU PASIG Event

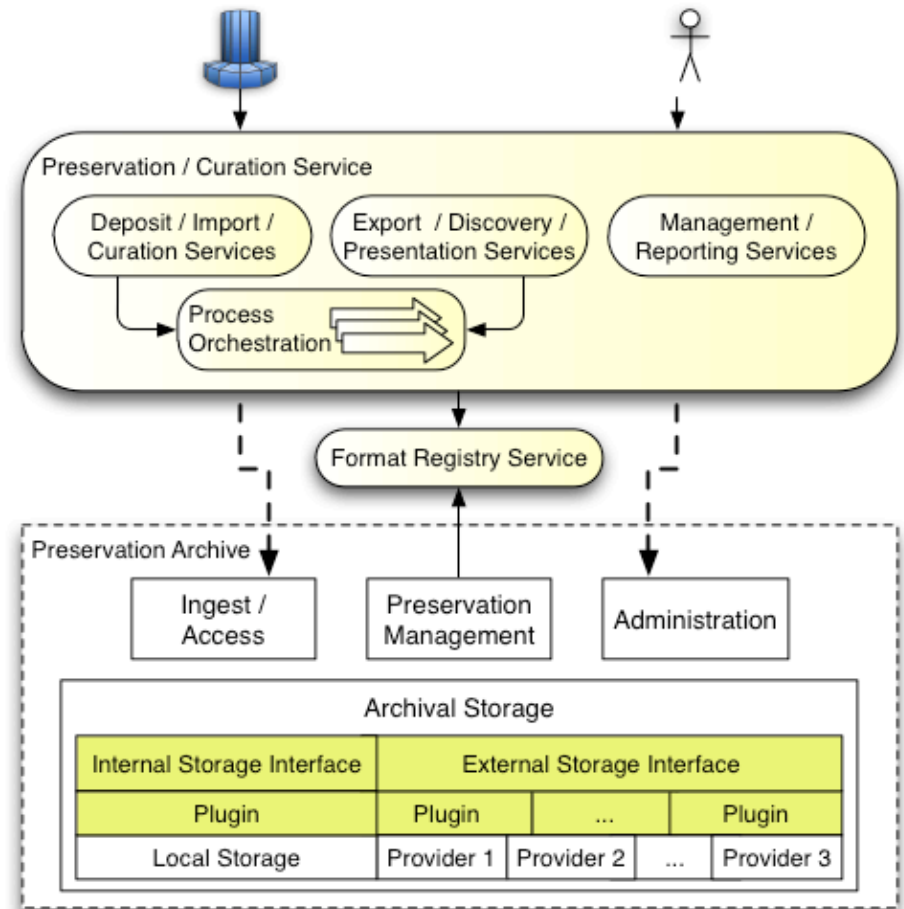
Data Curation

NSF DataNet award: The Data Conservancy
\$20 million, five years – nine partners

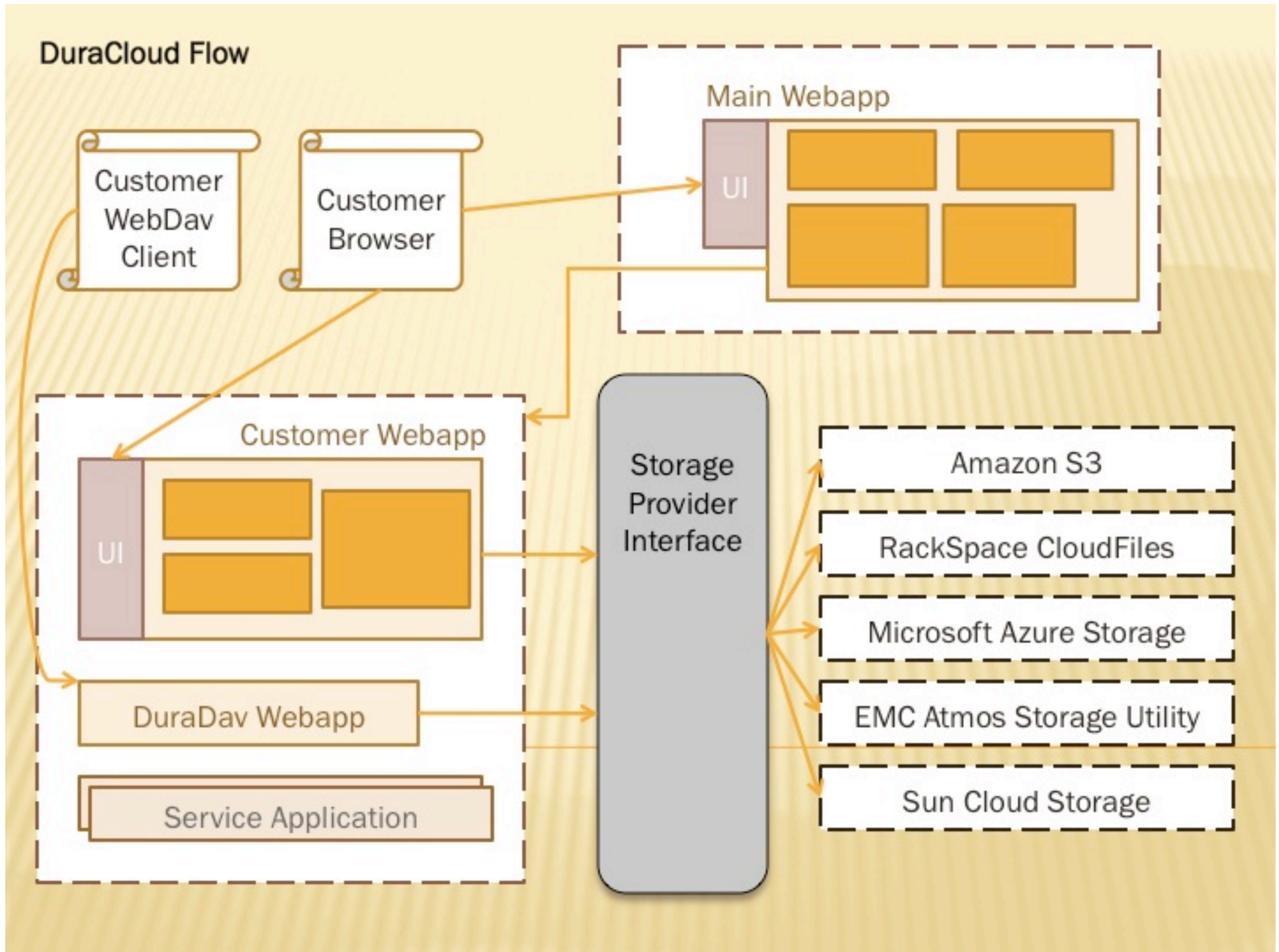
The Data Conservancy embraces a shared vision: data curation is not an end, but rather a means to collect, organize, validate, and preserve data to address grand research challenges that face society.

Technical Architecture

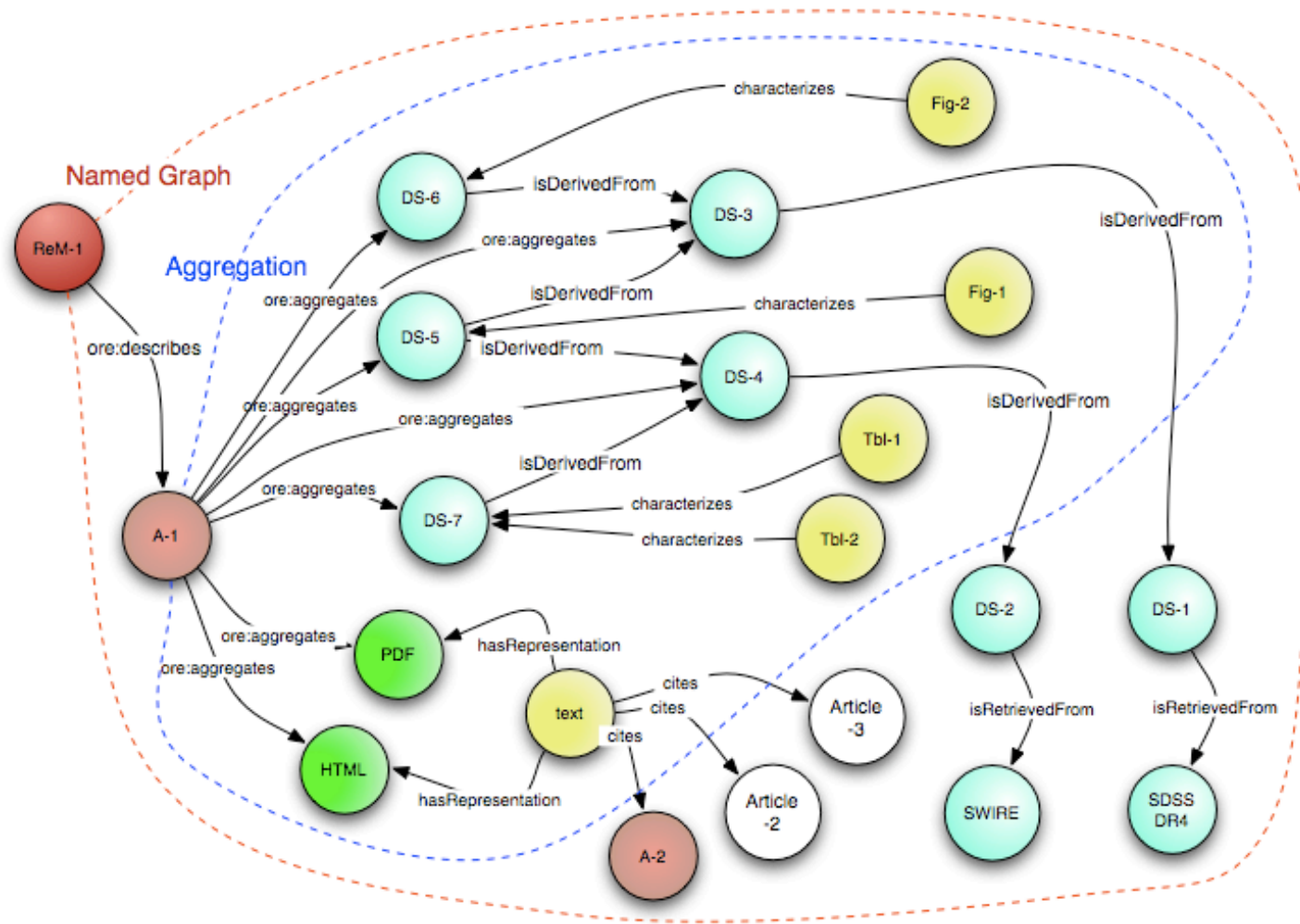
- Well-defined interfaces
- Emphasis on modularity and interchangeability
- Reference implementation using Fedora, but other implementations desirable
- Storage layer features
 - Access to data stored / preserved elsewhere
 - Interface to third-party storage services



DuraCloud Flow



Data Model using OAI-ORE



Organizational Dimension

- DataNet solicitation mentioned the importance of organizations that are both risk aware and risk averse
- JHU Libraries is the lead organization for Data Conservancy
- Digital Research and Curation Center (aka Digital Knowledge Center) is an embedded R&D group within JHU Libraries

Acknowledgements



DataNet award for “The Data Conservancy”



NLG grant award LG0606018206



- Tim DiLauro and David Reynolds (OAI-ORE slide)
- Tim DiLauro (Technical architecture slide)
- DuraSpace (DuraCloud slide)