

Scientific Annotation Middleware

DOE National Laboratories Program

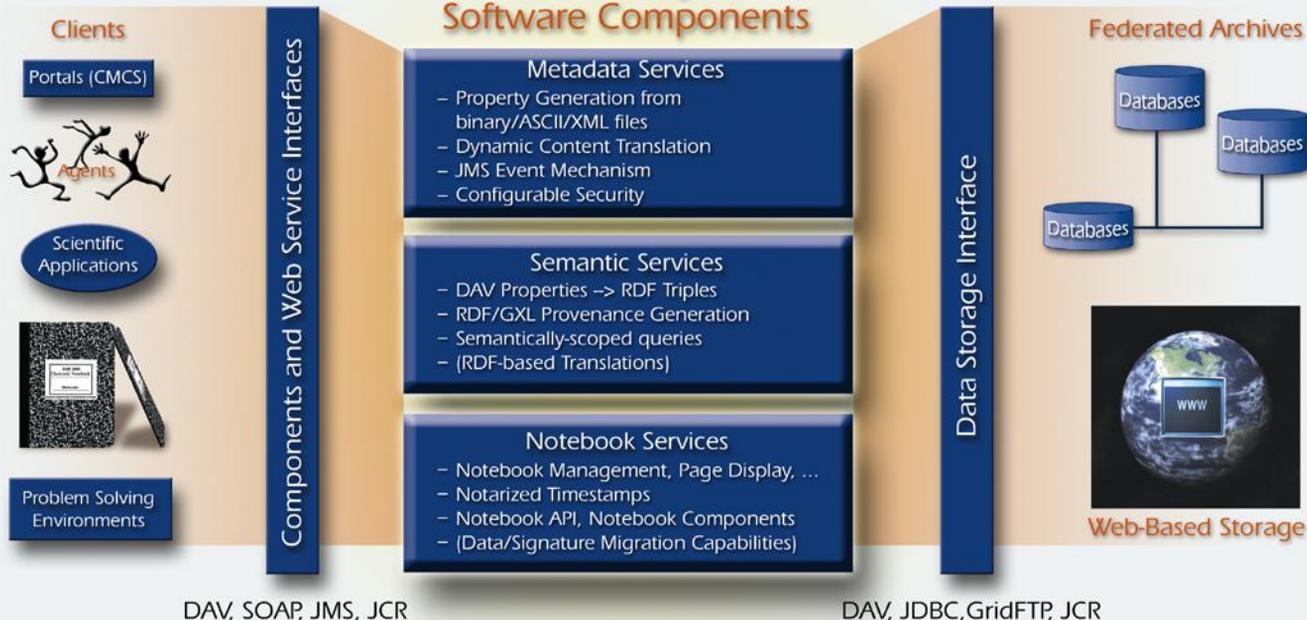


U.S. DEPARTMENT OF ENERGY

Principal Investigators: James Myers, Tara Talbott, Michael Peterson, Alan Chappell, Carina Lansing (PNNL)
AI Geist, Jens Schwidder (ORNL)

Program Manager: Mary Anne Scott

SAM's Major Software Components



Benefits

- Integrated electronic notebook, semantic relation (e.g. provenance) tracking, and third-party annotation services
- Rich, accessible, integrated scientific records
- Increased automation of metadata capture and data/metadata translation
- Support for systems-science approaches and collaboration across disciplines
- Open source, standards-based scientific content management services

Recent Advances

- JAAS-based Single Sign-on Capabilities
- Simple Web-based SAM and ELN Administration
- Metadata and Content Caching and Data Store Performance Tuning
- Binary Format Description (BFD) Language, Web-service, and XSLT-based Metadata Extraction and Data Translation Capabilities
- Semantically-scoped DASL-based Search
- New Web-based Notebook Interface (interoperable with ELN notebook)

Software

<http://www.scidac.org/SAM/>
BSD/Apache-style open source license
Version 1.2 released [April 2004]
Version 2.0 released [August 2004]

Community Interactions

- Collaboratory for Multiscale Chemical Science - using SAM to support a portal-based community knowledge grid
- NEESgrid - SAM-based ELN as an internationalized, grid-capable e-notebook
- Data Format Description Language (DFDL) Standardization within the Global Grid Forum
- Java Content Repository (JSR 170) Standardization within the Java Community Process (JCP)
- Public - 1600+ registrations

Selected Papers

- Re-Integrating the Research Record, James D. Myers, Alan R. Chappell, Matthew Elder, AI Geist, Jens Schwidder, Computing in Science and Engineering, May/June 2003, pp 44-50
- Multi-Scale Science: Supporting Emerging Practice with Semantically Derived Provenance, James D. Myers, Carmen Panzerella, Carina Lansing, Karen L. Schuchardt, and Brett Didier, Proceedings of the Semantic Web Technologies for Searching and Retrieving Scientific Data Workshop, Sanibel Island, FL, October 20, 2003
- A Collaborative Informatics Infrastructure for Multi-scale Science, James D. Myers et al., Proceedings of the Challenges of Large Applications in Distributed Environments (CLADE) Workshop, June 7, 2004, Honolulu, HI