



VIRTUAL ASTRONOMICAL OBSERVATORY

Data Curation and Preservation Interest Group

Robert Hanisch
STScI/VAO



The VAO is operated by the VAO, LLC.



Presentations

- Arnold Rots: linking
 - Identifiers
 - Repository for data behind the figures in journals
 - iRODS
- Rob Seaman: data compression
 - Rice algorithm is optimal combination of compression ratio and computational speed
 - Compression is needed both for storage and bandwidth
- Matthew Graham: data storage in the Cloud
 - features of Cloud-based storage
 - not all Clouds are the same
 - Cloud-base storage is not inexpensive
 - VOSpace implementations could interface to Cloud storage



Presentations, cont'd

- Mike Fitzpatrick: data access
 - separate data discovery from data access
 - importance of inventory and footprint services
 - responsiveness of services important factor for users, data providers --> proxy services, VOTable stream and binary formats, develop inventory VOSI standard
- Andre Schaaf: Cloud storage costs
 - Vizier data could be stored at Amazon S3 for ~\$6,000/year, not including mirror sites, remote computation, etc.



Synthesis

- DCP is a key topic in the VO
 - Data providers: efficiency, cost
 - Metadata
 - Discovery and access
 - Linking
- VAO plans to provide repository service, details TBD
- VAO Science Council urges collaboration with journal publishers for data capture and linking