DFC DataNet FEDERATION CONSORTIUM

Odum Institute iRODS Policies to Support Preservation NDSLabs Demonstration

Jonathan Crabtree Reagan Moore Don Sizemore Craig Willis

NATIONAL DATA SERVICE

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

UNC THE ODUM INSTITUTE

renci

National Science Foundation
Cooperative Agreement:
OCI-0940841

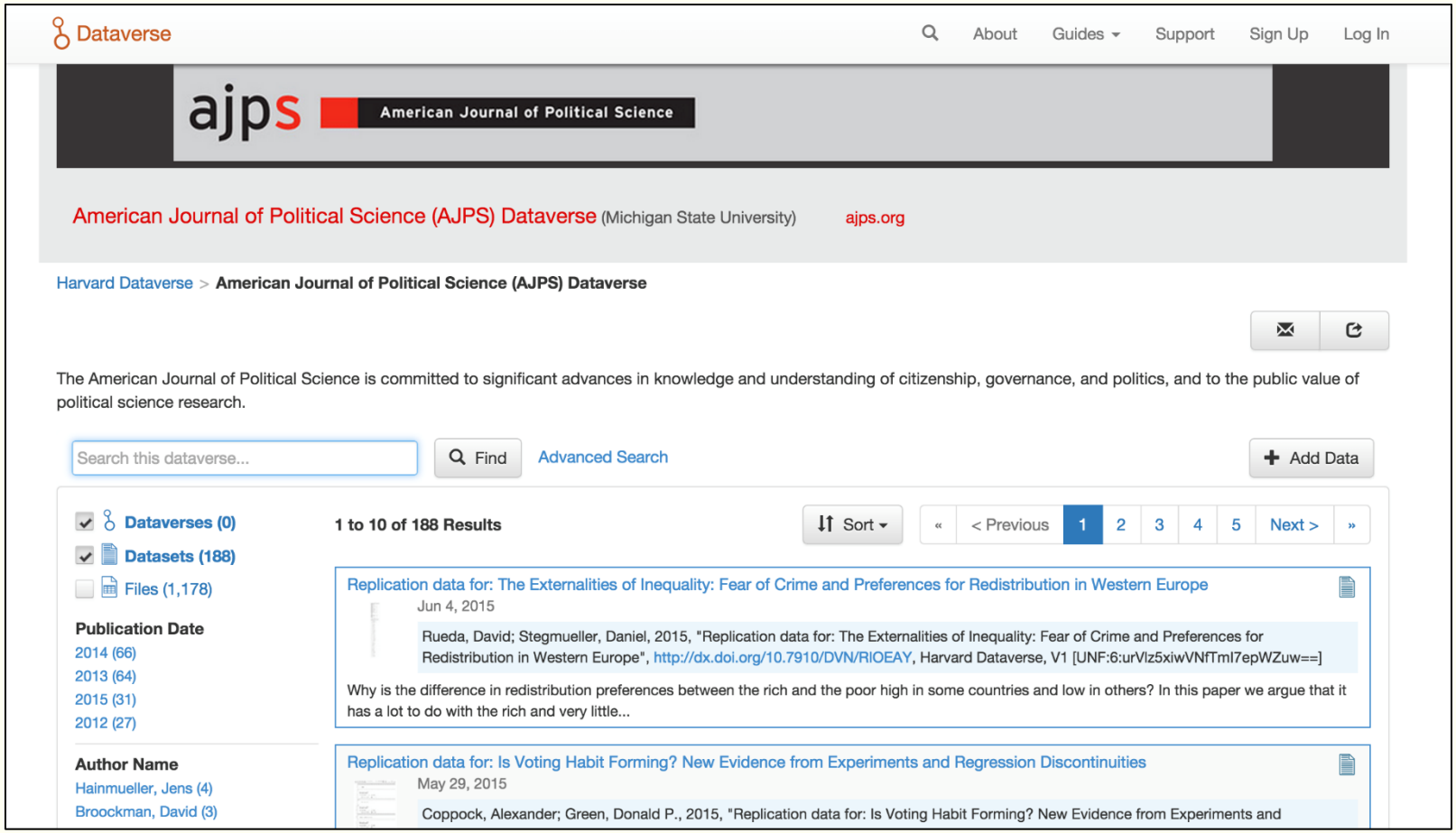
Project Goals

- Design curation/preservation workflow integration
- Connect research environment with archive
- Connect archive with national architecture
- Policy based
- Open source focused
- As “pluggable” as possible

Bringing great tools together

- iRODS
- Dataverse
- Modeshape
- BitCurator
- BagIt
- Preservation Audit Standards
  - ISO 16363
  - Data Seal of Approval
- Databook Architecture
  - Apache Service Mix
  - iRODS Rule Integration
  - Indexing Engine

Integration of the Dataverse



DFC DataNet FEDERATION CONSORTIUM

NSF

Leveraging iRODS

- Automate preservation actions
- Rule based policy management
- Scalable storage infrastructure
- Enhances secure data sharing possibilities
- Diversifies Dataverse storage options

Preservation Policy Enforcement

- Odum Current Production Efforts
  - iRODS configured to utilize BitCurator
  - Testing sensitive information identification utilizing iRODS/BitCurator/BulkExtractor
  - Selected ISO 16363 policies implemented with iRODS rules
  - Rules published to GitHub
  - Utilizing irsync client/server method to push data from Odum Dataverse to iRODS and policies to archive data in DFC
  - Integrated staging iRODS server to traverse UNC Firewall

Next Steps

- Stress test sensitive information identification and valid file formats
- Add more complex sensitive information formats
- Expand iRODS rules based policy reporting
- Complete testing of iRODS/Modeshape/ Dataverse integration
- Enable Dataverse/iRODS metadata lookup

Links

- https://github.com/DICE-UNC/dfc-dataverse-integration
- https://github.com/DICE-UNC/indexing-irods
- https://github.com/akio-sone/dvn/tree/Odum-Ext
- https://github.com/donsizemore/odum/tree/master/irods_rules
- https://en.wikipedia.org/wiki/BagIt
- http://datafed.org
- http://www.odum.unc.edu/odum/home2.jsp
- http://Dataverse.org
- http://nationaldataservice.org

Thanks

Demonstration of workflow and tools in NDSLabs

DFC DataNet FEDERATION CONSORTIUM

NSF