

Developing Research Data Life Cycle Strategy

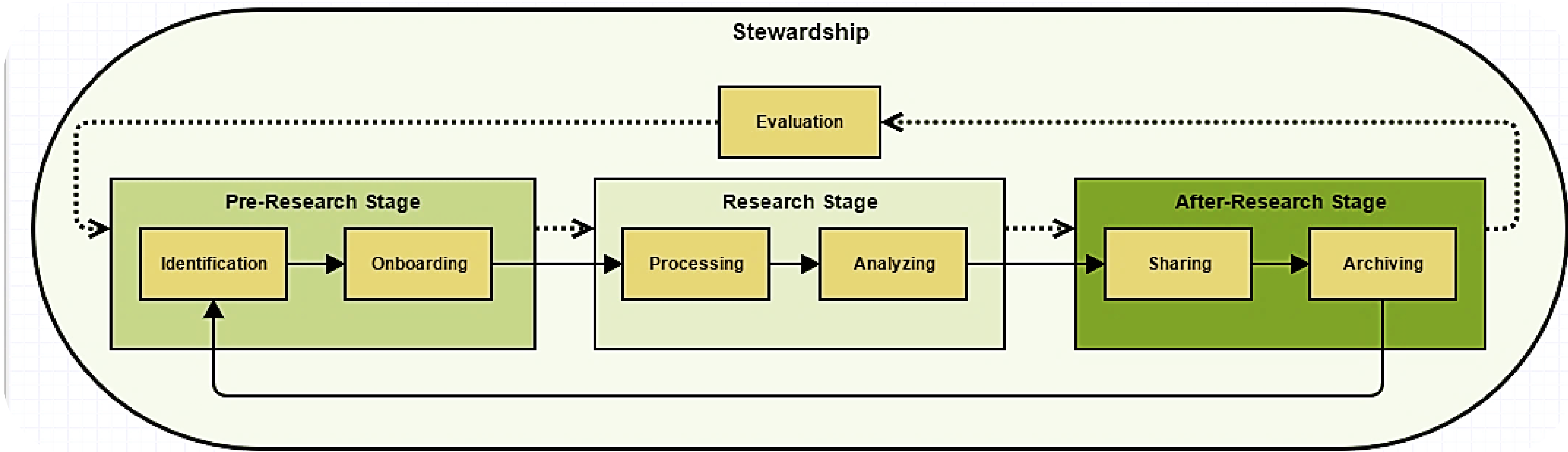
A Collaborative Approach at Fed Chicago

Background

The Computing and Data Committee (CDC) that consists of researchers, IT staff, data librarians in the Economic Research Department of Federal Reserve Bank of Chicago developed The Research Data Life Cycle Strategy (RDLCS) in mid 2016 to provide a strategic direction on data management and data governance. RDLCS has been endorsed by senior management of the Department and presented to other departments as a model to adopt.

Mission Statement

Research analysis is the core output of the Economic Research department. Nearly all policy and academic research work are built of this analysis. RDLCS will support researchers' empirical analysis by enabling them to identify, access, and use data quickly, efficiently, and conveniently while promoting sharing and appropriate preservation subject to Federal Reserve System policy, contract mandates, and budgetary considerations through effective data stewardship.



Identification

end user, data librarian

This is the stage where potential datasets, including publicly available ones as well as contractual ones, are identified and selected by researchers or data librarian for the use in research projects.

- Provide tools to identify data
- Provide training
- Provide appropriate communication concerning data-related topics from data librarian to Economic Research economists
- Develop data librarian expertise in data coverage, usage, and other questions related to data topics

Processing

end user, IT, data librarian

This is the stage where data are integrated into the infrastructure and prepared for general use in research projects.

- Develop procedures and assign clear responsibility for:
 - Preparation: Cleaning, validation, translation, and transcription
 - Loading: Provisioning data on appropriate platforms
 - Formatting: Making data available in a known file structure
 - Monitoring: Checking for updates to the data
 - Auditing: Performing access reviews

Sharing

end user, data librarian, editor, legal

This is the stage where research data and the analytical results are, to the greatest extent possible, disseminated and shared with the research community.

- Share Economic Research owned data to the greatest extent possible
- Collaborate with other Federal Reserve System entities to maximize Chicago Economic Research’s ability to access their data
- Ensure continued compliance with Federal Reserve System policy and contract constraints

Onboarding

data librarian, IT, procurement, legal

This is the stage where identified datasets are brought in through collection or acquisition by researchers or data librarian.

- Consult with technical support on the feasibility of using particular data
- Conduct speedy, responsible, and compliant data acquisition
- Discover and catalog new data

Analyzing

end User, IT

This is the stage where researchers actually employ the data as well as the derived data for analytical work.

- Provision data such that it can be efficiently used by economists subject to FRS policy and contract constraints

Archiving

data Librarian, end user, IT

This is the stage where research data itself as well as the associated work products are preserved based on standards to facilitate future identification and re-use.

- Work with researchers to develop best practices for preserving data and documenting data files for replicability
- Encourage and train researchers to follow such practices
- Leverage Federal Reserve System resources for data curation and preservation
- Make archived data available to the Identification stage

Action Item #1 (2017)

Identify Who Owns What in Economic Research

- Conducted a Research data inventory survey – to identify the data we have owned/used as well as what has been “hidden”
 - 59 out of 81 researchers responded; 34 “hidden” datasets were reported
- In the process of developing metadata
 - Sources consulted: Fed contracted data catalog, Fed Board enterprise data inventory; ICPSR; DDI
- In the process of evaluating application platforms to build an internal infrastructure to host the inventory

Action Item #2 (2017)

Assign Permanent Identification to Work Product

- Implement permanent identification (DOI) to Bank working papers and other research publications to promote the share of scholarly work and data

Action Item #3 (2017)

Promote RDLCS to Other Departments

- The Bank Data Capability group (consisting of IT and business units) endorsed data life-cycle concept
- Will work with Data Capability group to catalog non-contractual structured data.

