

---

# METADATA & DATA CURATION SERVICES

THU-MAI CHRISTIAN, DATA ARCHIVIST

DATAVERSE COMMUNITY MEETING

JUNE 10, 2015



UNC

THE ODUM INSTITUTE

THE ODUM INSTITUTE FOR RESEARCH IN SOCIAL SCIENCE

DAVIS LIBRARY, 2<sup>ND</sup> FLOOR, CB# 3355

UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

[WWW.ODUM.UNC.EDU](http://WWW.ODUM.UNC.EDU)

# METADATA & DATA CURATION SERVICES

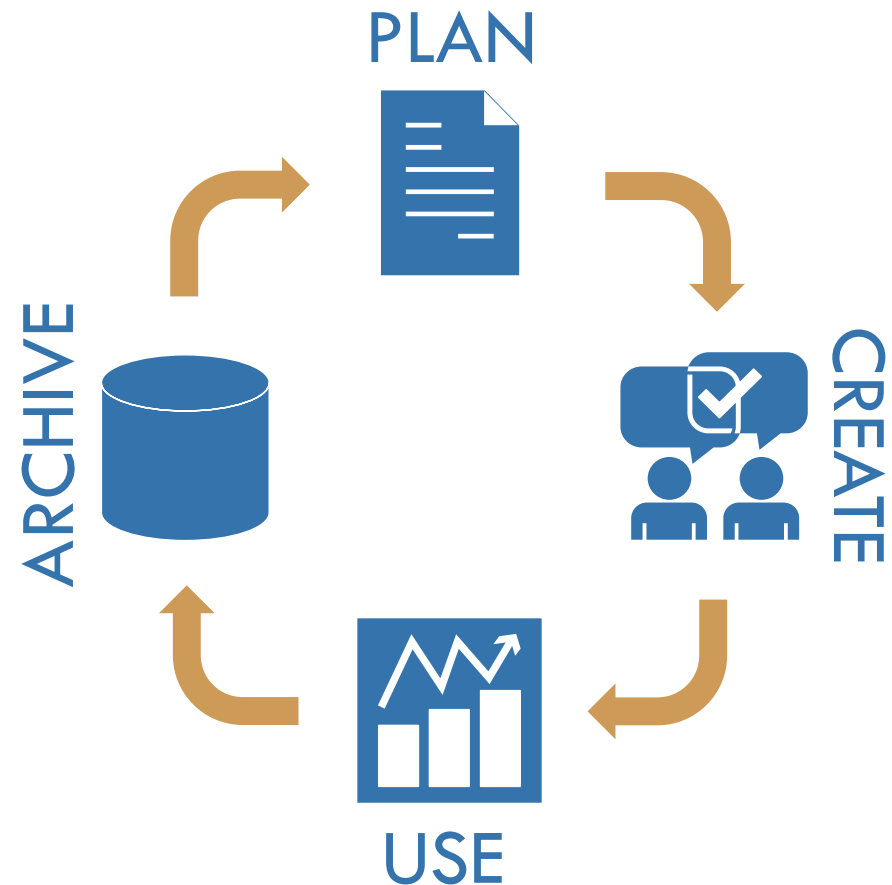
- The Odum Institute for Research in Social Science
- Odum Institute Data Archive
- Data Curation Pipeline + Dataverse
- Data Curation Service Levels

# THE ODUM INSTITUTE FOR RESEARCH IN SOCIAL SCIENCE

- Founded in 1924 by Howard W. Odum
- Oldest university-based interdisciplinary social science research institute in the U.S.
- *To help grow and lead a world-class **social science research infrastructure** at the University of North Carolina at Chapel Hill to ensure that researchers can conduct **scientifically rigorous research** that contributes to better lives of the citizens of North Carolina and the World...*

# THE ODUM INSTITUTE FOR RESEARCH IN SOCIAL SCIENCE

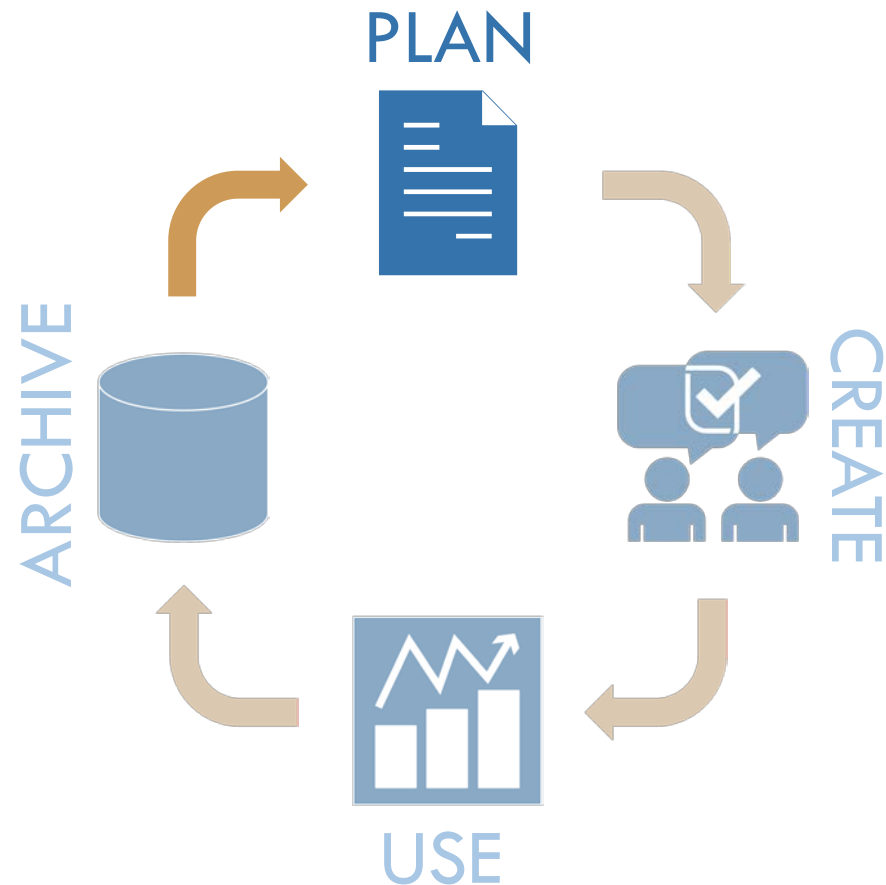
## SERVICES ACROSS THE RESEARCH LIFECYCLE



# THE ODUM INSTITUTE FOR RESEARCH IN SOCIAL SCIENCE

## SERVICES ACROSS THE RESEARCH LIFECYCLE

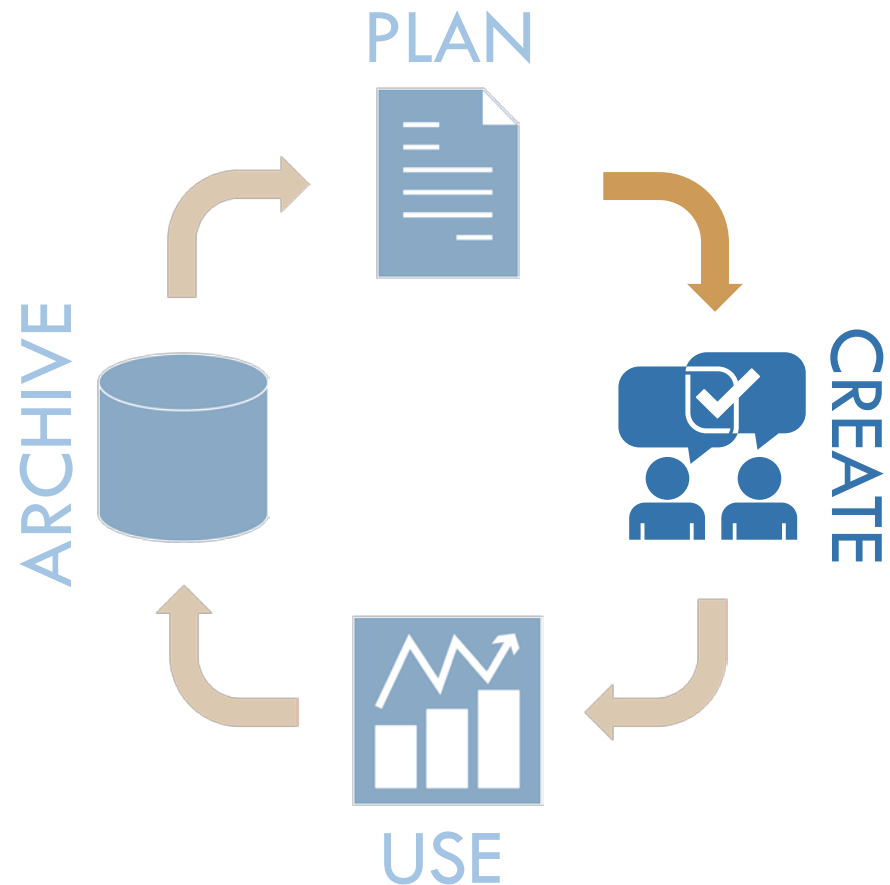
- ➔ Grant services
- ➔ Data management plans



# THE ODUM INSTITUTE FOR RESEARCH IN SOCIAL SCIENCE

## SERVICES ACROSS THE RESEARCH LIFECYCLE

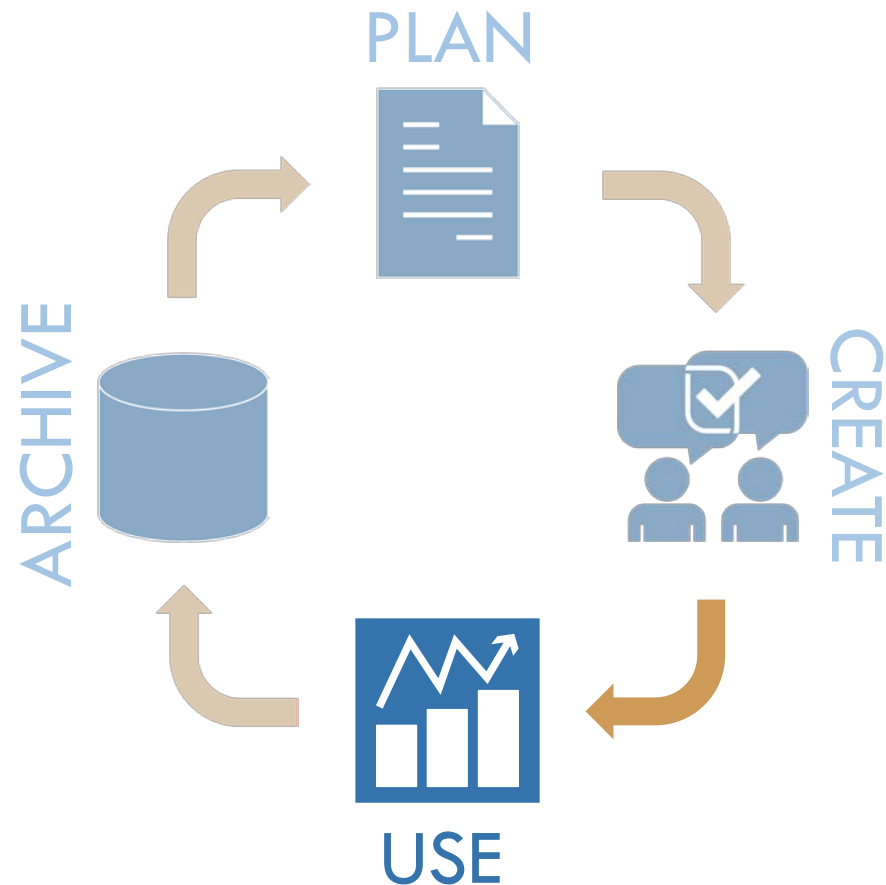
- ➔ Survey methodology
- ➔ Data collection
- ➔ Education & training



# THE ODUM INSTITUTE FOR RESEARCH IN SOCIAL SCIENCE

## SERVICES ACROSS THE RESEARCH LIFECYCLE

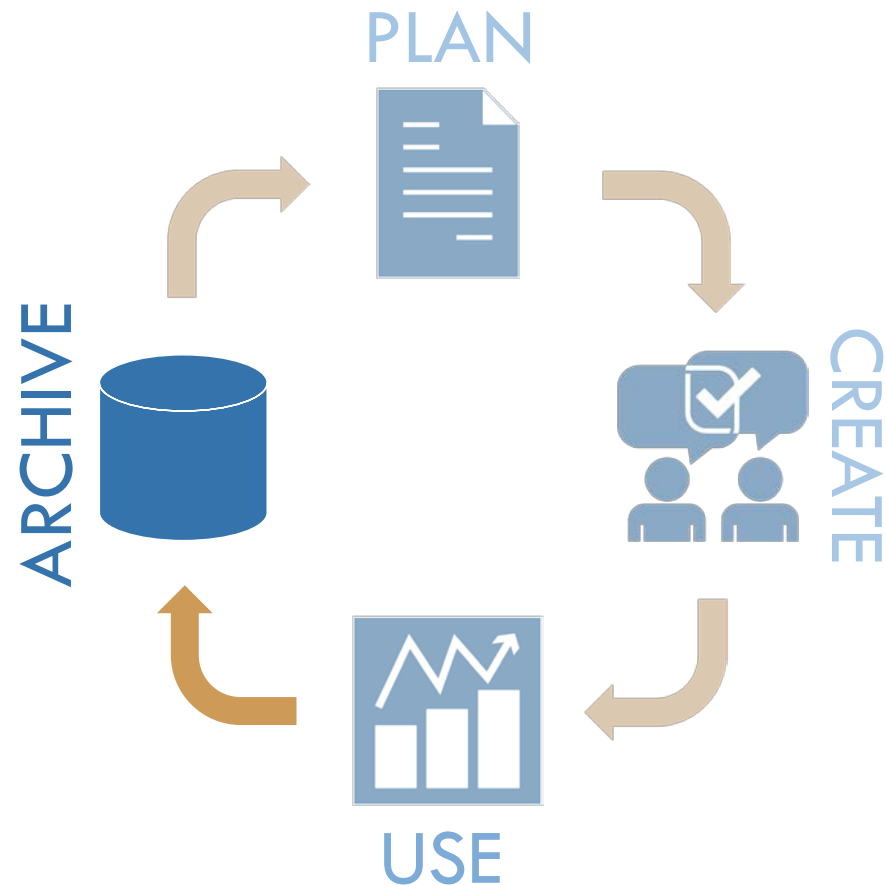
- ➔ Consultations
- ➔ Education & Training
- ➔ Facilities



# THE ODUM INSTITUTE FOR RESEARCH IN SOCIAL SCIENCE

## SERVICES ACROSS THE RESEARCH LIFECYCLE

- ➔ Data curation
- ➔ Data preservation
- ➔ Data sharing





# ODUM INSTITUTE DATA ARCHIVE

- One of the oldest and largest catalogs of machine readable data
  - Louis Harris Data Center
  - North Carolina Vital Statistics
  - National Network of State Polls (NNSP)
  - North Carolina Online Collection of Knowledge and Scholarship (NC DOCKS)
- Over 33,000 studies with over 690,000 files

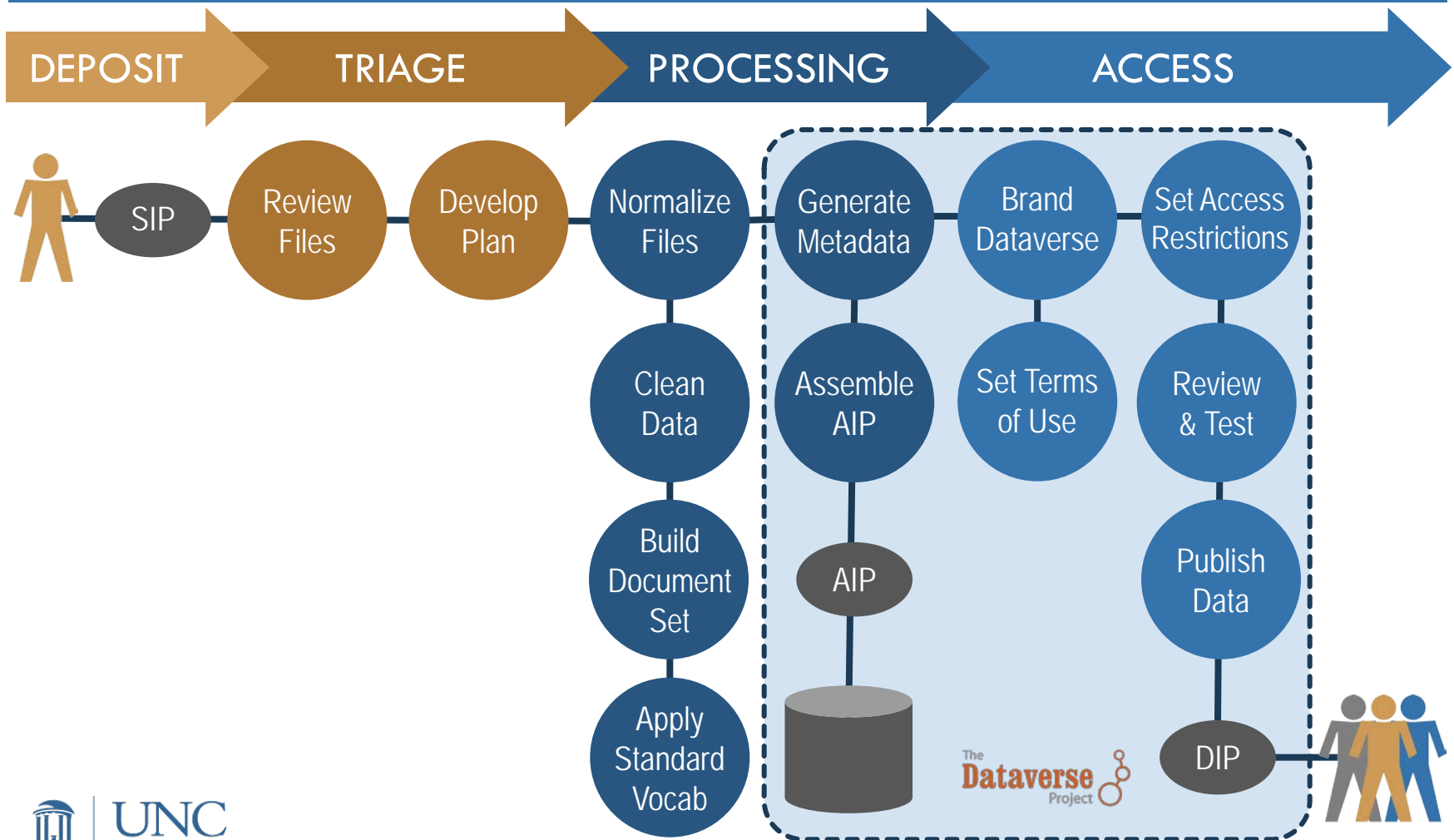
# ODUM INSTITUTE DATA ARCHIVE

- Member of the Data Preservation Alliance for the Social Sciences (Data-PASS)
- Awarded the 2014-2015 Data Seal of Approval

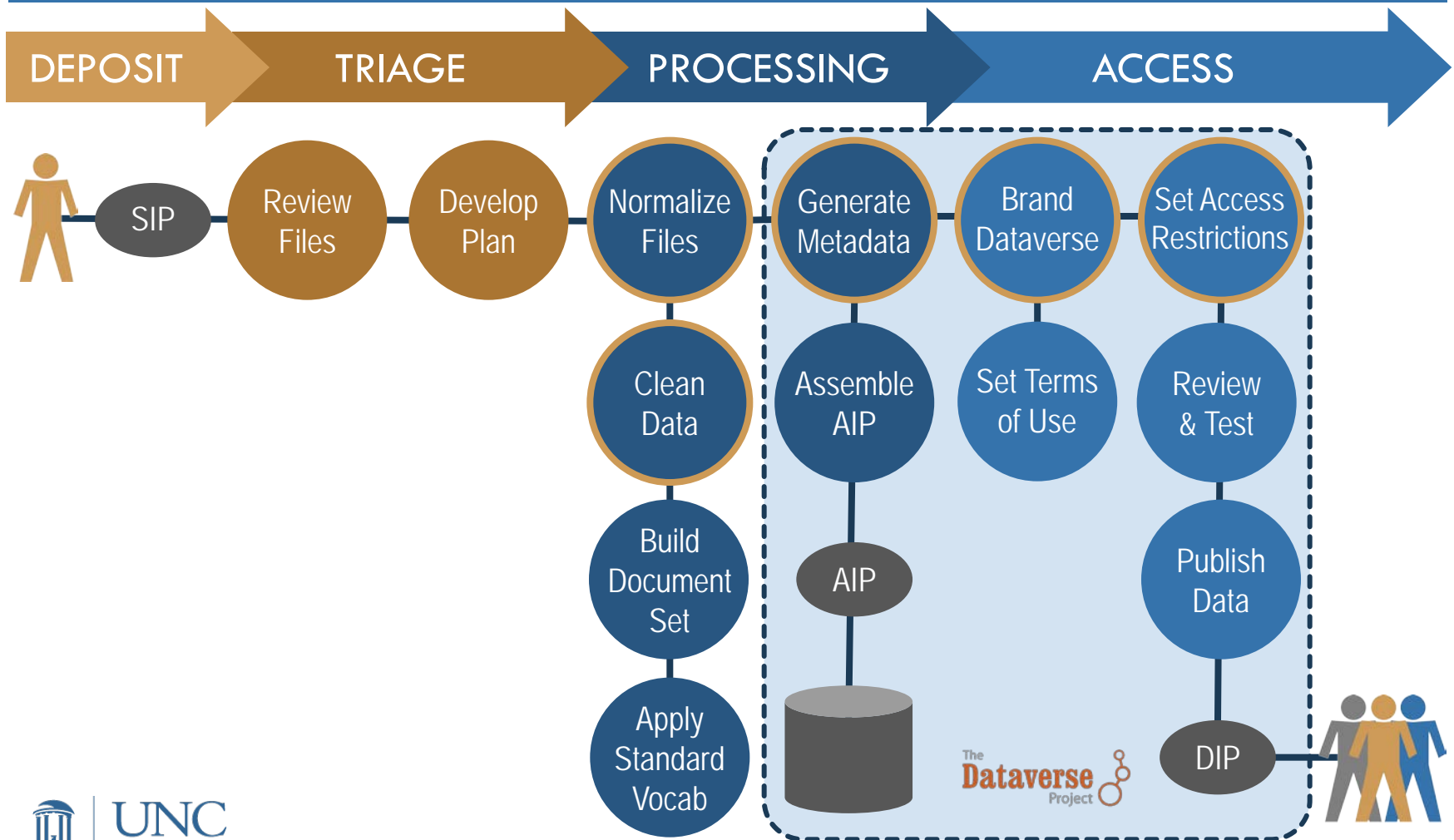
# ODUM INSTITUTE DATA ARCHIVE

- Adopted the Dataverse virtual archive platform in 2005
- Dataverse 3.6.2 ❤️ Data curation services

# DATA CURATION PIPELINE



# DATA CURATION PIPELINE



# DATA CURATION PIPELINE

## PROCESSING

Normalize  
Files

Convert files to Dataverse-compatible formats to enable subsetting and analysis

# DATA CURATION PIPELINE

## PROCESSING

Clean Data

Edit/Add labels to variables and value codes to enhance variable-level search and usability in the Dataverse



# DATA CURATION PIPELINE

## PROCESSING

Generate  
Metadata

Create metadata templates to ensure completeness of cataloging information and to encourage consistent use of metadata standards across dataset records in the Dataverse



# DATA CURATION PIPELINE

ACCESS

Brand  
Dataverse

Preserve individual or organizational identity by customizing the Dataverse look and feel by incorporating logos and HTML/CSS



# DATA CURATION PIPELINE

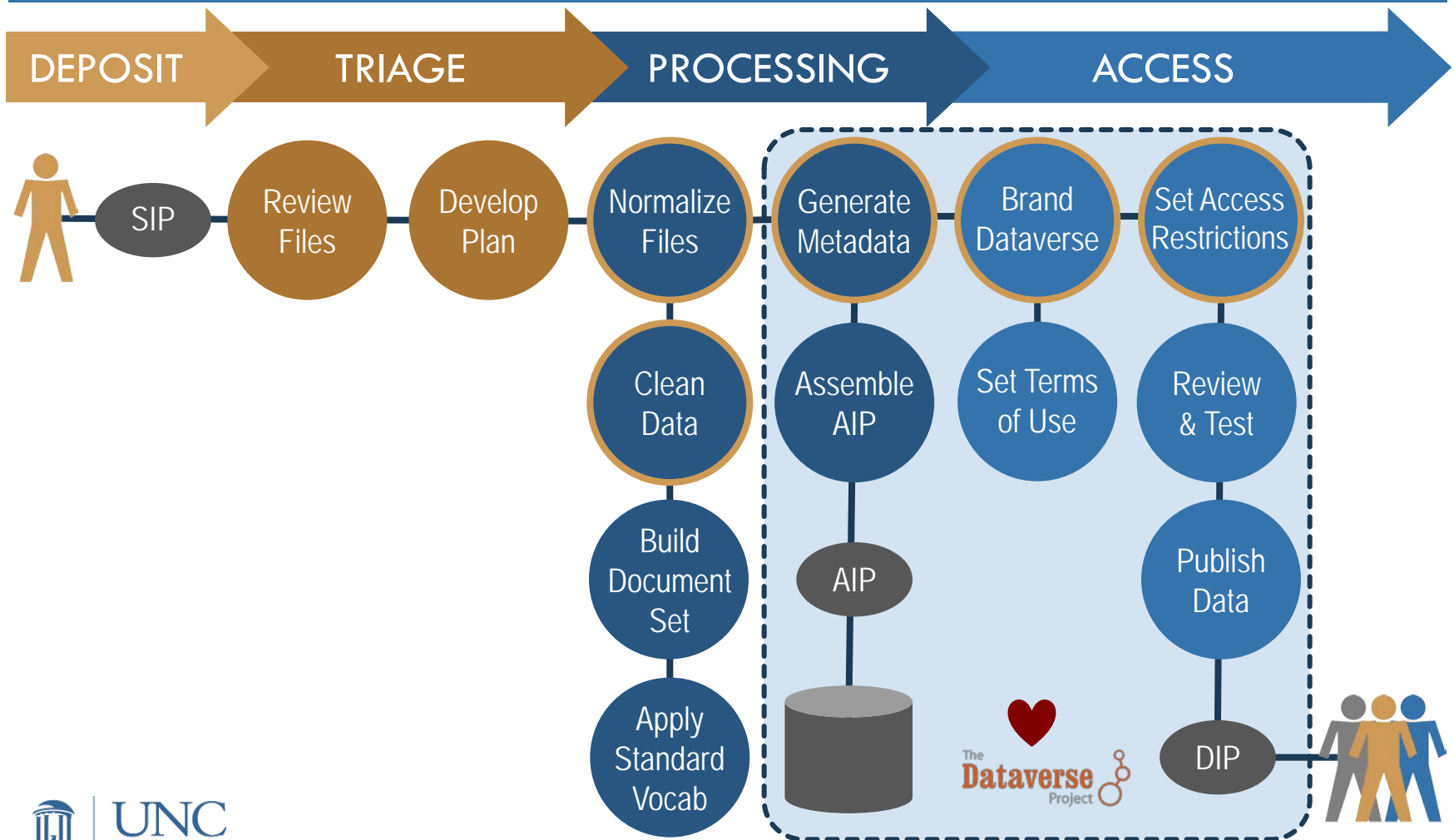
ACCESS

Set Access  
Restrictions

Set and enforce access restrictions in the Dataverse according to data access policies



# DATA CURATION PIPELINE



# DATA CURATION SERVICE LEVELS

**Self-Service**

**Guided Service**

**Lifecycle Service**



# DATA CURATION SERVICE LEVELS

## Self-Service

- + Dataverse tool access
- + Persistent identification
- + Data citation
- + Basic utilization reports
- + Long-term preservation
- + Standardized metadata
- + Limited user support

# DATA CURATION SERVICE LEVELS

## Guided Service

- + Introductory training
- + Dataverse branding
- + Collection arrangement
- + Metadata templates
- + Standard user support

# DATA CURATION SERVICE LEVELS

## Lifecycle Service

- + File normalization
- + Data file review
- + Access policy enforcement
- + Education programming
- + Dedicated user support

# DATA CURATION SERVICE LEVELS

## Special Services

- + De-identification
- + Data verification
- + Secure data services
- + Tool integration



# DATA CURATION SERVICE LEVELS

- Clear articulation of data curation services
- Emphasis on Dataverse value-added features
- Advocacy for the adoption of data curation standards and best practices
- *A model for sustainability*

THANK YOU!

**Thu-Mai Christian**  
[tlchristian@unc.edu](mailto:tlchristian@unc.edu)

**The Odum Institute**  
<http://www.odum.unc.edu>

 @Odum\_Institute

 The Odum Institute