

# Planning for Migration

## Format Transformations in DAITSS

# Florida Digital Archive

[: contact us :](#) [: sitemap :](#)



AFFILIATES



[: mission :](#)

[: menu :](#)

The mission of the Florida Digital Archive is to provide a cost-effective, long-term preservation repository for digital materials in support of teaching and learning, scholarship, and research in the state of Florida.

In support of this mission, the Florida Digital Archive guarantees that all files deposited by agreement with its affiliates remain available, unaltered, and readable from media. For those materials designated to receive full preservation treatment, the Florida Digital Archive will maintain a usable version using the best format migration tools available.

- [Digital Archive Information](#)
- [Format Information](#)
- [Publications and Presentations](#)
- [Software and Documentation](#)
- [Other Good Resources](#)
- [IMLS Grant Information](#)
- [News](#)

The Florida Digital Archive is based on DAITSS, a preservation repository management application, which is available as open source software under the GPL license. See <http://daitss.fcla.edu>.

 search[Login](#) [Settings](#) [Help/Guide](#) [About Trac](#)[Start Page](#) [Index by Title](#) [Index by Date](#) [Last Change](#)

## Wiki

[Timeline](#)[Roadmap](#)[Browse Source](#)[View Tickets](#)[Search](#)

## Welcome to DAITSS Navigation

- [Downloads](#)
- [Documentation](#)
- [About](#)
- [Contact](#)

## News

### **09.13.2007 -- DAITSS 1.2.4 now available**

Source and binary distributions of DAITSS 1.2.4 are now available for download.

- [Release Notes](#)

### **06.29.2007 -- DAITSS 1.2.3 now available**

Source and binary distributions of DAITSS 1.2.3 are now available for download.

- [Release Notes](#)

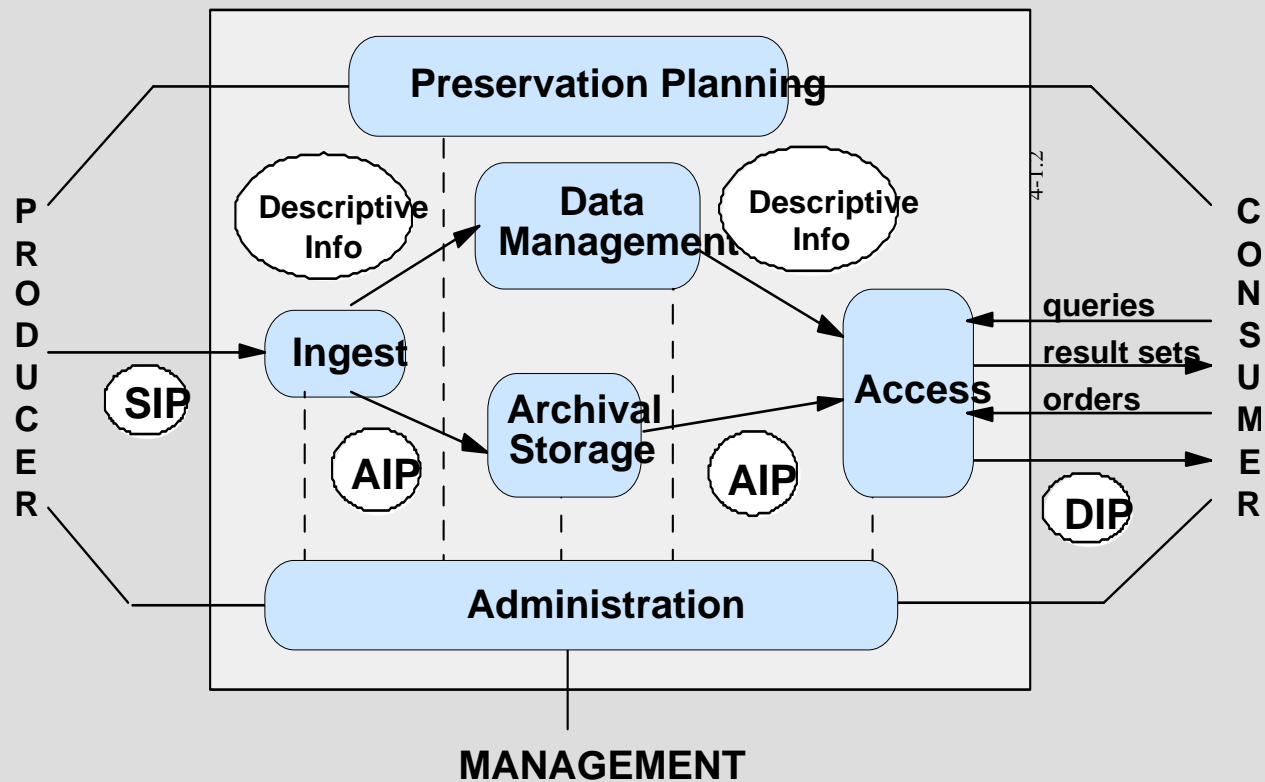
### **04.27.2007 -- DAITSS 1.2.2 now available**

Source and binary distributions of DAITSS 1.2.2 are now available for download.

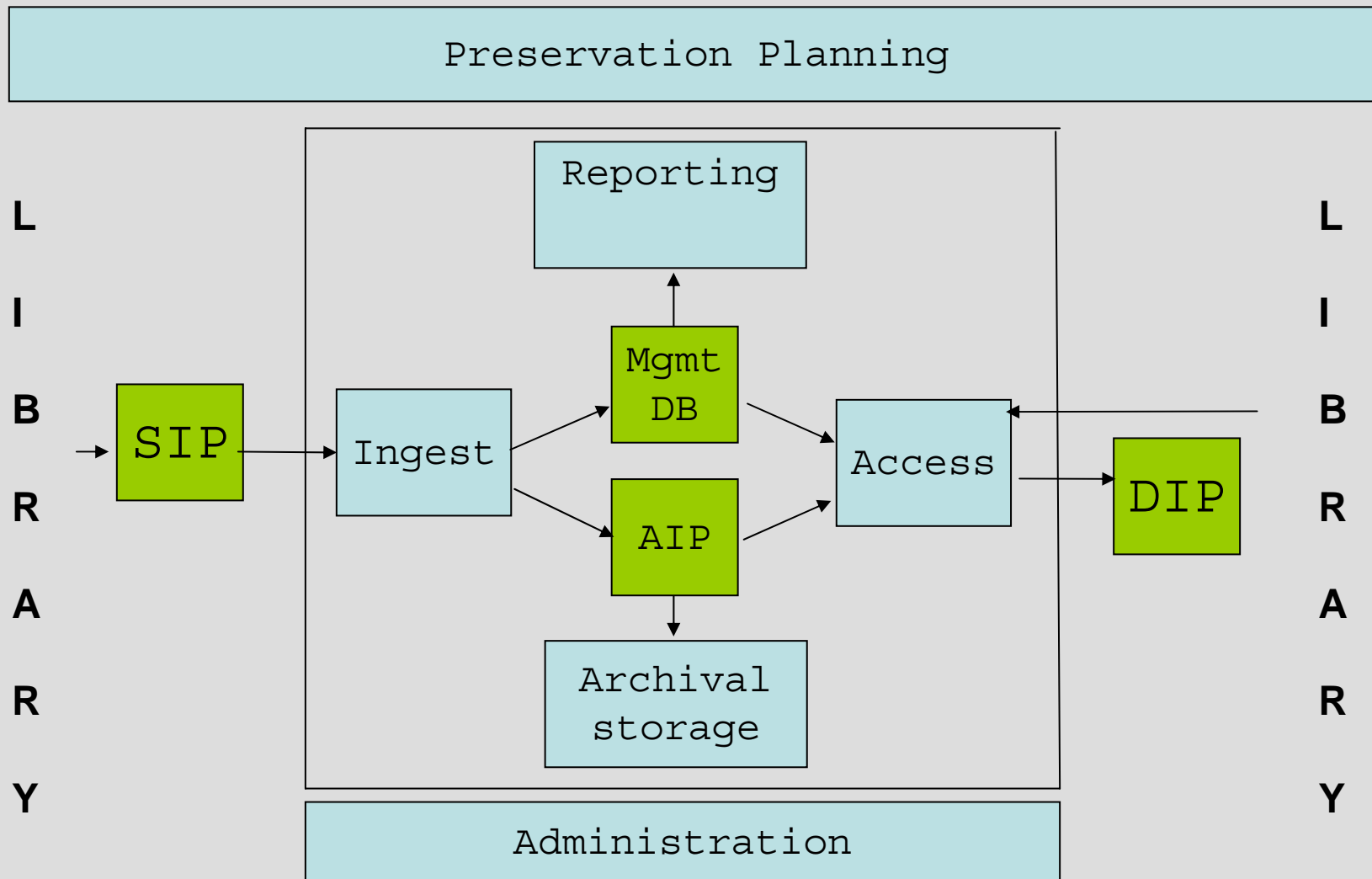
- [Release Notes](#)

### **04.12.2007 -- DAITSS 1.2.1 now available**

# OAIS Functional Architecture

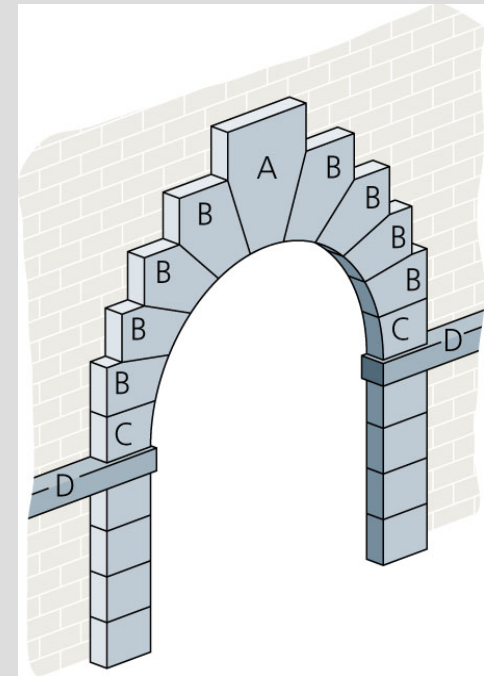


# DAITSS Functional Architecture



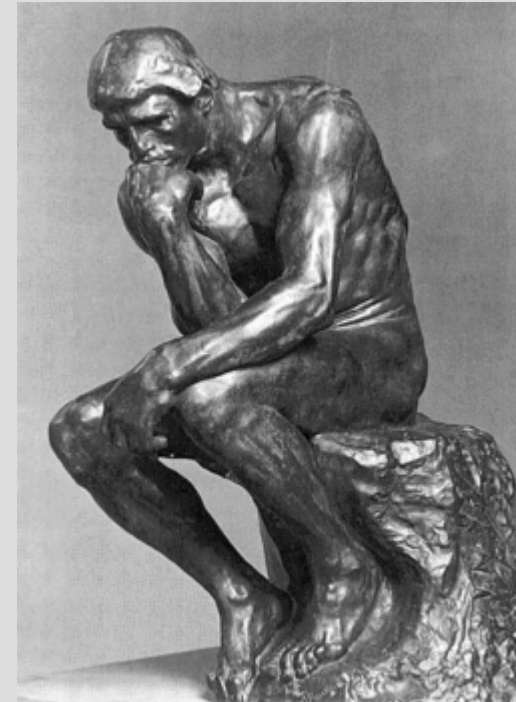
# Key Architectural Decisions

- AIP will be self-defining: database can be rebuilt from data store
  - all metadata stored in external RDBMS and in XML AIP descriptor
- Nothing but Ingest writes to data store
  - audit trail/digital provenance in PREMIS Events records
- Preservation strategies built in from the start
  - preserve original as submitted
  - do format transformations as needed
  - retain original and “last best”

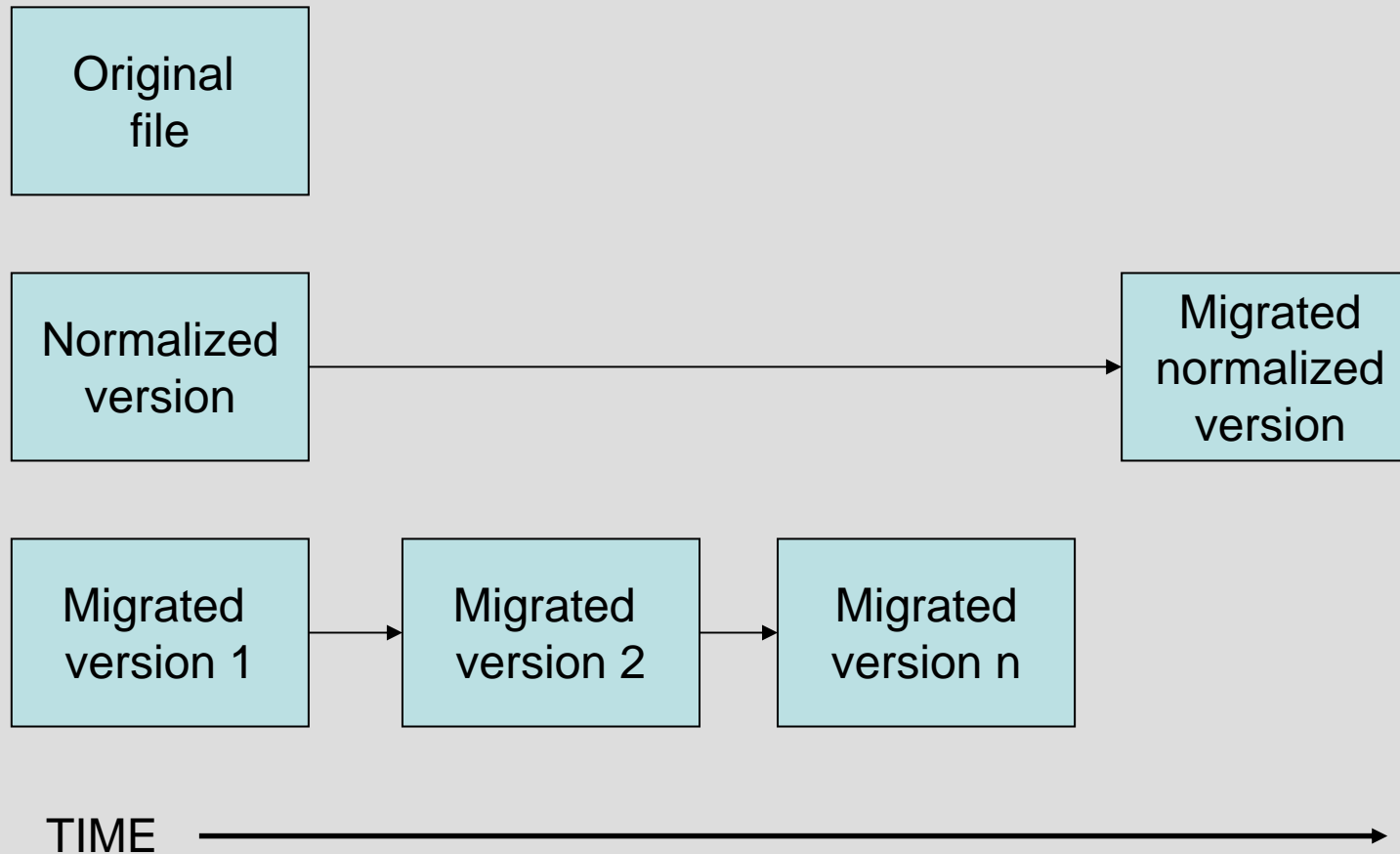


# Decision points:

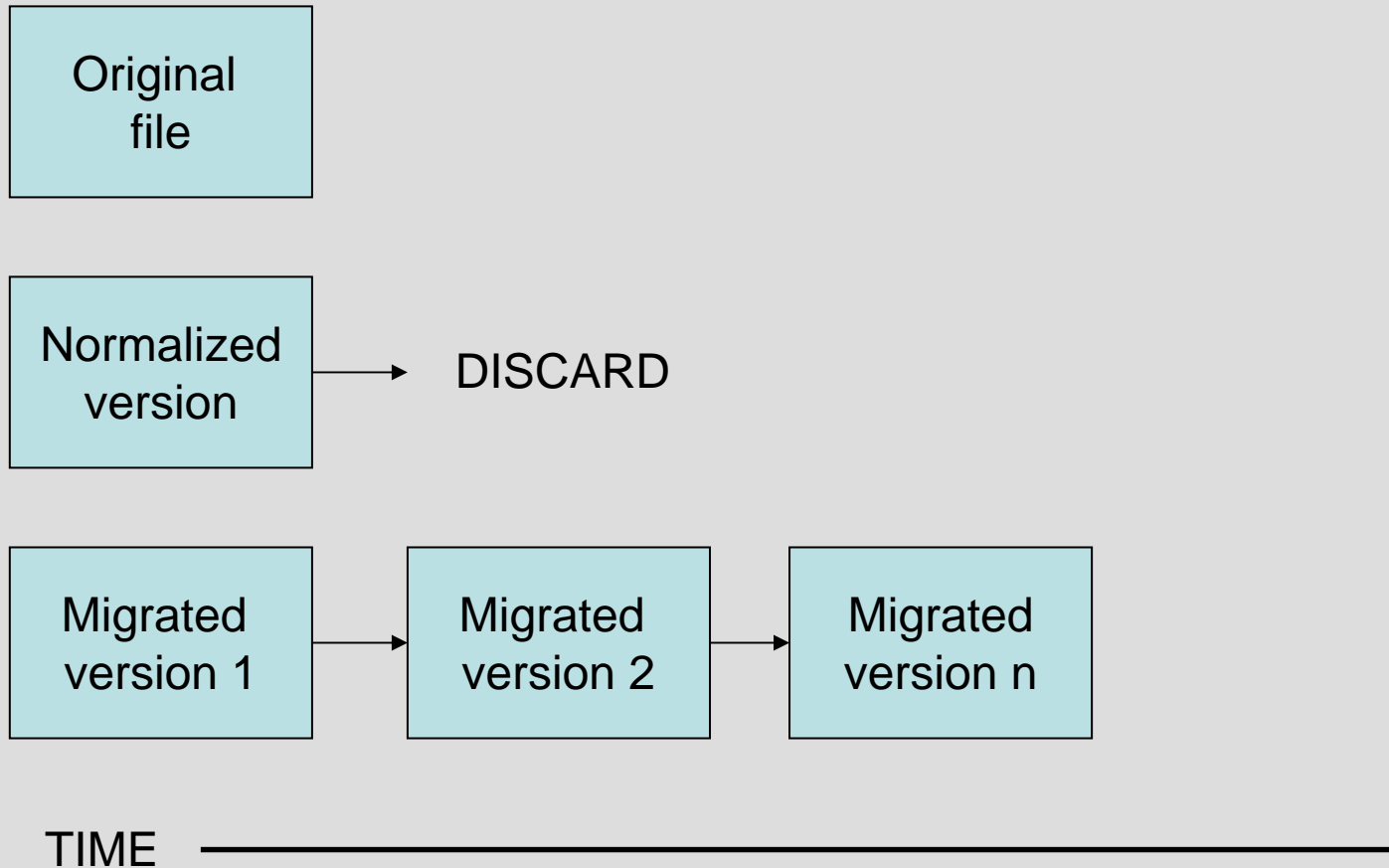
- What types of transformations to do?
  - normalization
  - forward migration
  - localization
- Where to do transformations?
  - not before ingest, because no provenance
  - not after ingest, because only Ingest can write AIP
- What if obsolescence happens after ingest?
  - every dissemination includes a re-ingest step
  - allow mass migration and migration on request
- What if a migration routine is added after ingest?
  - disseminate (mass or on request)



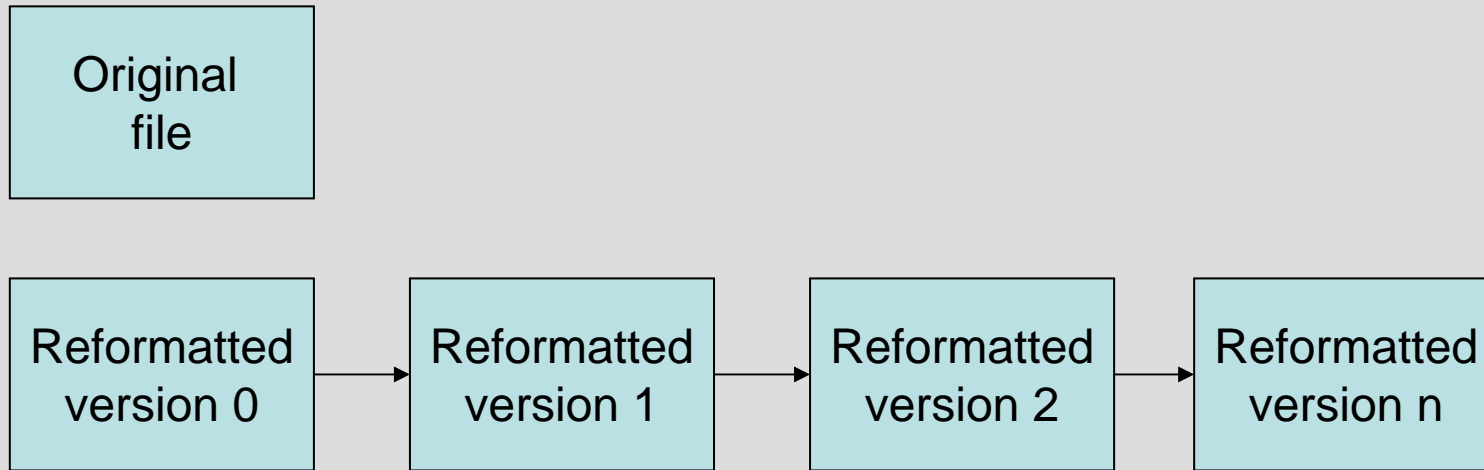
# Original plan




# Current processing

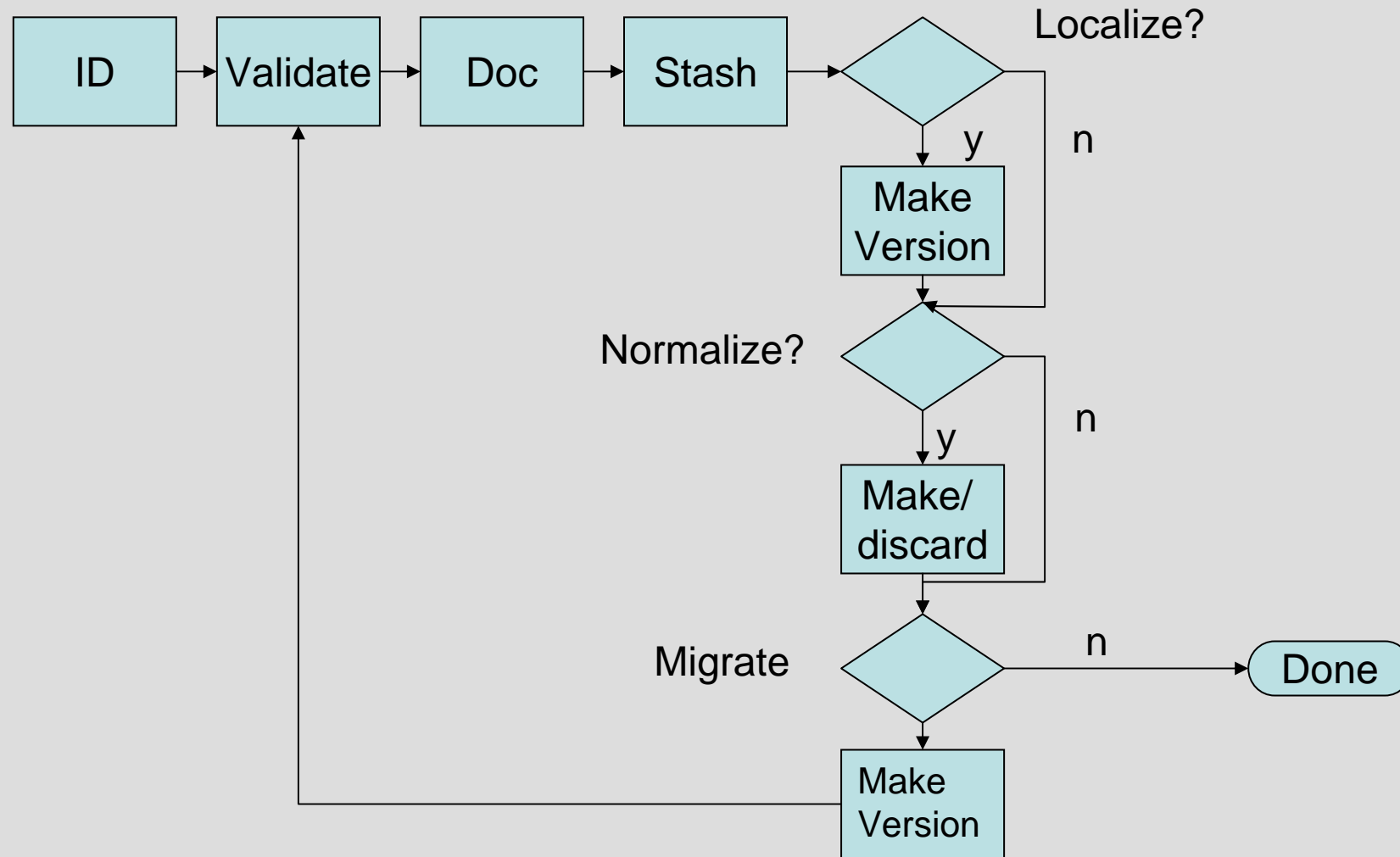


# Next version?

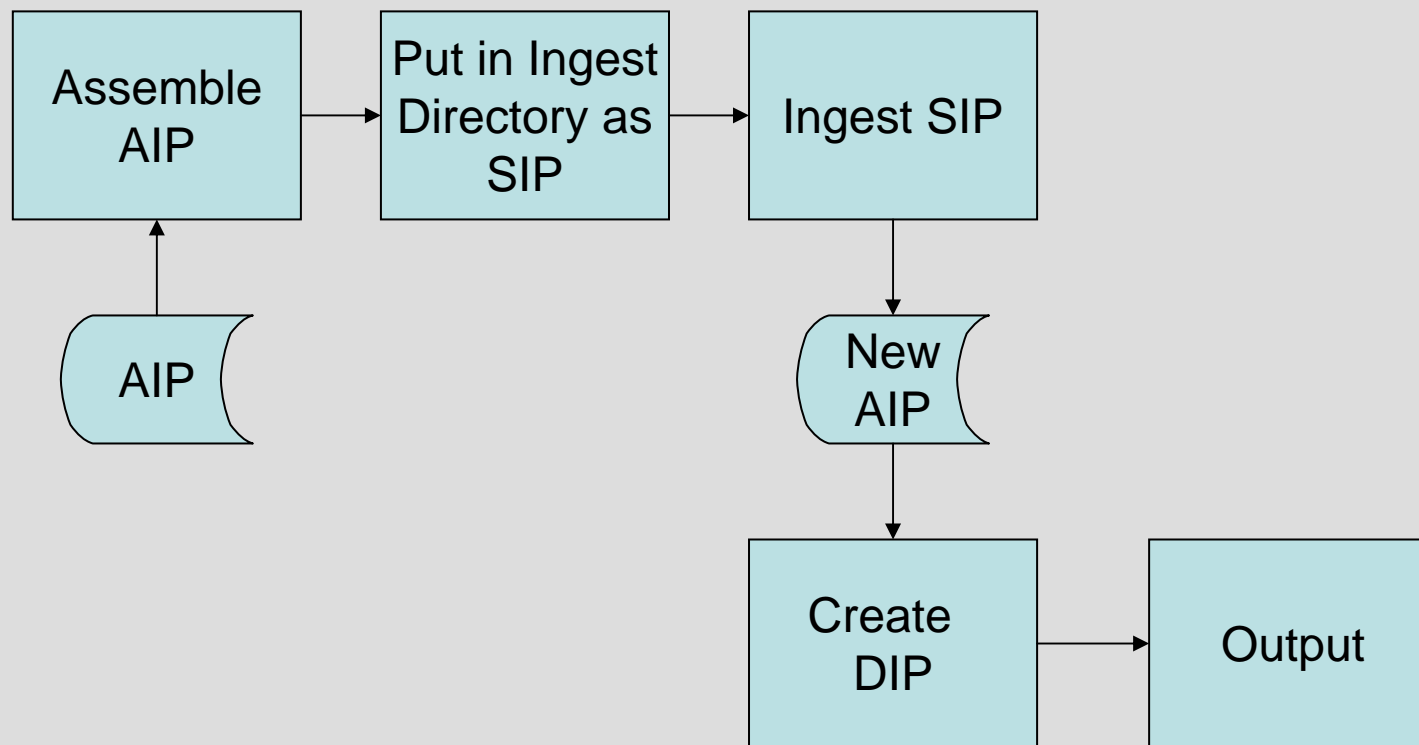


TIME 

# Ingest steps (partial) for each file in the SIP



# Dissemination steps (AIP to DIP)



# Format Transformations

- QuickTime and AVI
  - video streams (cinepak, SMC, AVC1) to Motion JPEG A (MJPA)
  - audio streams to uncompressed PCM (LPCM)
- Comma Separated Values (CSV)
  - normalize to follow RFC 4180



# Plans for DAITSS 2.0

- Move to web service architecture
- Make transformation logic rule (table) driven, not hard coded in data file class
- Use transformation service that reformats based on rules (agnostic to migration, normalization, etc.)
- Record more detailed provenance information by describing software as PREMIS agents

