DOIs and Research Data in RUcore

For Discussion Today:

- An Update on RUcore

- Object Architecture and DOIs in RUcore

- DataCite and DOI Metadata Indexing

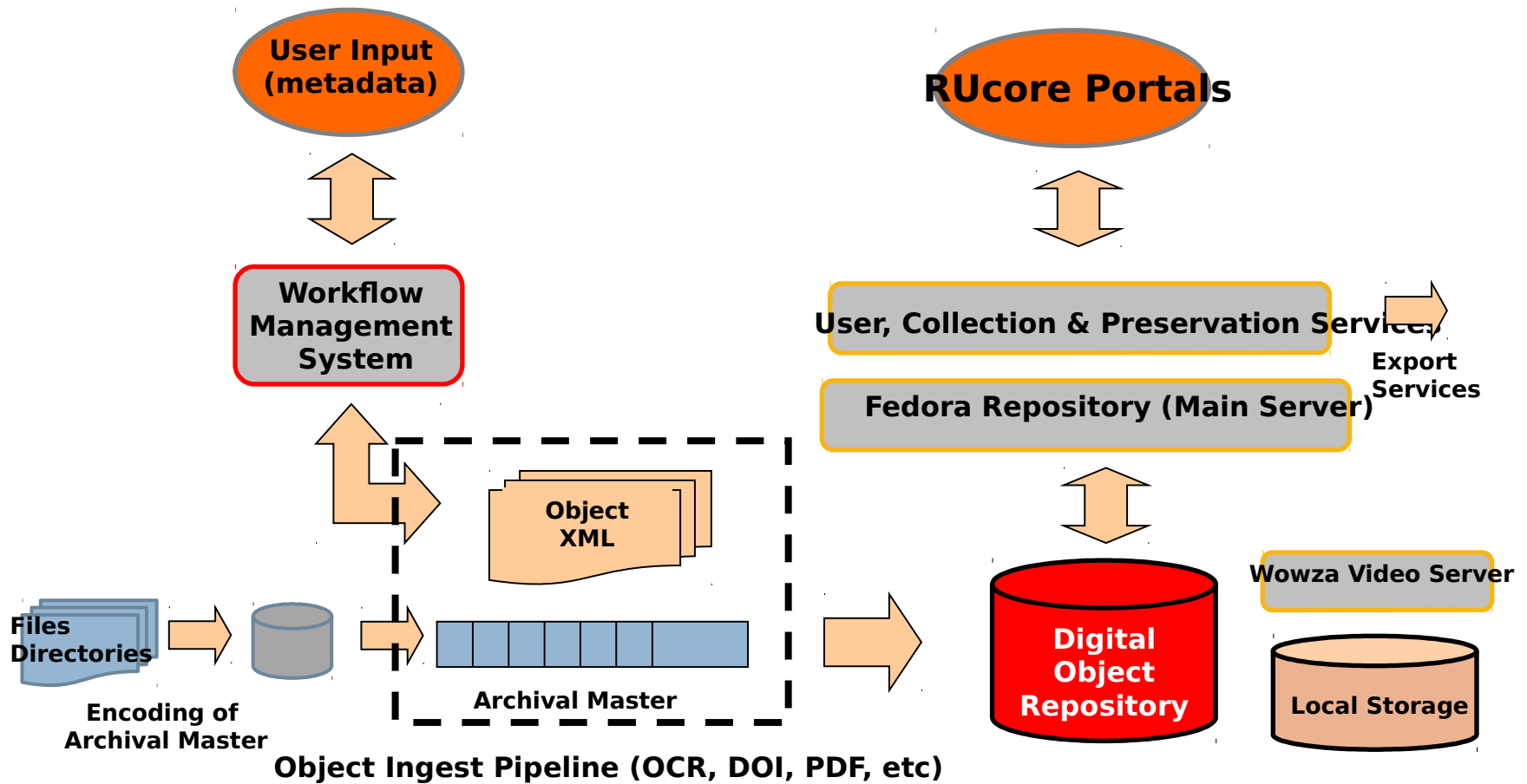- DOIs at the file level

RUcore Data Repository – Progress

- ☐ The Research Data Portal – recent additions
  - Ingest of multi-level directories as submitted by the researcher
  - Use of DOIs as persistent identifiers for all resources (Fedora objects)
  - User accounts for all researchers, capture ORCID id if available

- ☐ Research Data Working Group
  - Established group including library liaisons
  - The group provides support for grants and NSF Data Management Plan

RUcore Architecture Overview

User Input (metadata)

RUcore Portals

Workflow Management System

User, Collection & Preservation Services

Export Services

Fedora Repository (Main Server)

Object XML

Files Directories

Encoding of Archival Master

Archival Master

Wowza Video Server

Digital Object Repository

Local Storage

Object Ingest Pipeline (OCR, DOI, PDF, etc)

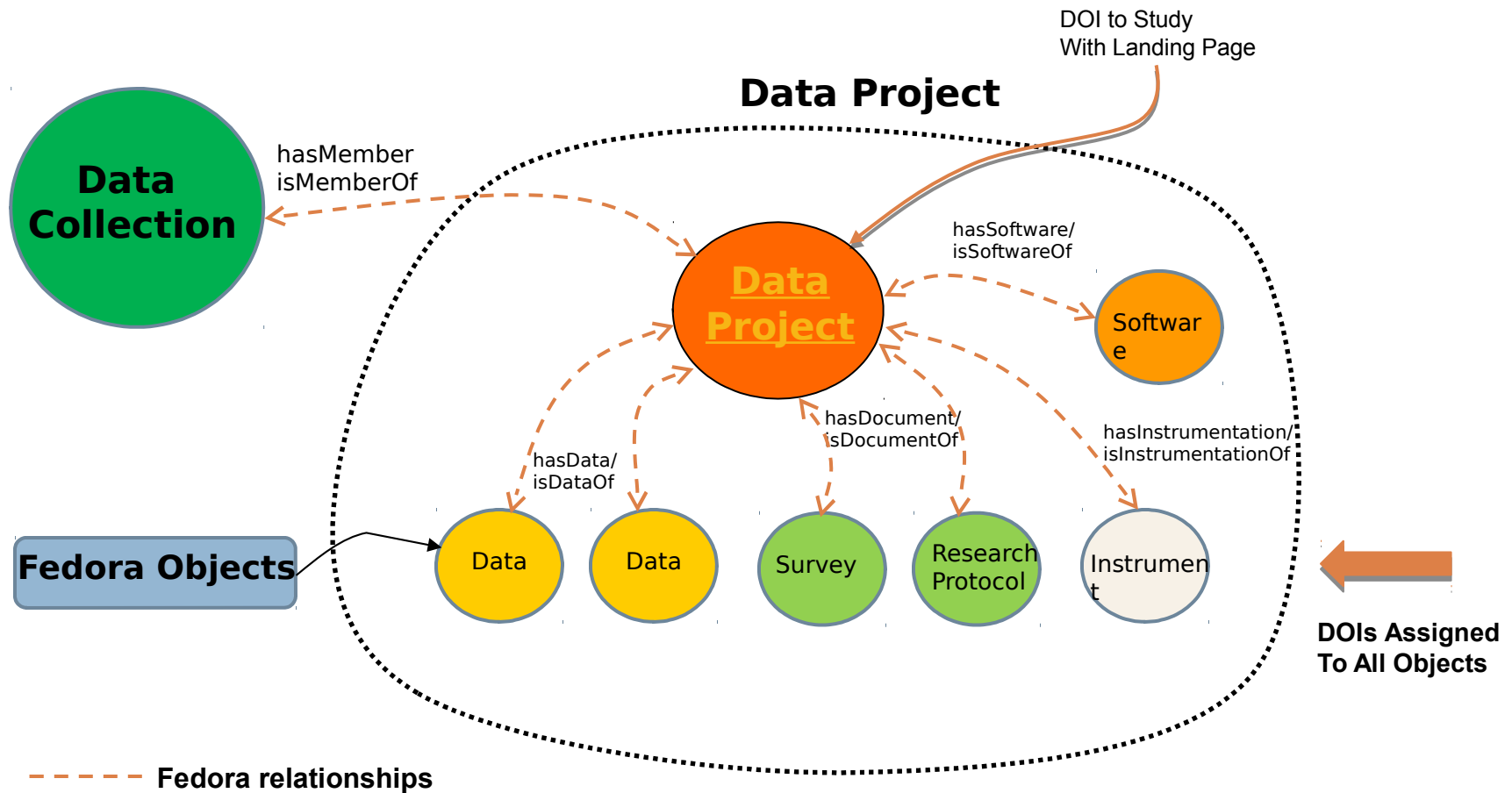# Background on DOIs

- ## DOIs for RUcore

  - The DOI is a persistent digital identifier for any object of intellectual property.

  - The DOI format - a name prefix obtained from a Registration Agency (RA) and a suffix from the publishing organization
    - DataCite/EZID DOIs for data, e.g. 10.4231/D3QJ77X57
    - CrossRef DOIs for journals, e.g. 10.1016/j.acalib.2010.08.004

- ## EZID

  - A service from California Digital Library to enable one to create DOIs and ARKs

  - EZID can be used for all types of objects.

  - EZID is available through a web form and an API with three basic operations: put (create), get (view), and post (modify)

RUcore Data Project and DOI Assignment

DOI to Study
With Landing Page

Data Project

Data Collection

hasMember
isMemberOf

Data Project

hasSoftware/
isSoftwareOf

Softwar
e

hasData/
isDataOf

hasDocument/
isDocumentOf

hasInstrumentation/
isInstrumentationOf

Fedora Objects

Data

Data

Survey

Research
Protocol

Instrumen
t

DOIs Assigned
To All Objects

- - - - - **Fedora relationships**

Microcitation1 and DOIs at the File Level?

- ☐ **RUcore Conventions**
  - The persistent identifier (DOI) is assigned at the Fedora object level
  - An object can be a single dataset or can include many related datasets or other types of files

- ☐ **Why DOIs at the file level?**
  - Researchers have a more fine-grained citation capability
  - Precision in citing specific resources such as custom software
  - More accuracy and elimination of errors in citing
  - Encourage data deposit and nanopublication2
  - Provide citation credit

1 Also micro-attribution
2 Smallest unit of publishable information

Micro-citation and Implementation

Questions

- A compound object has many files. Do all files receive a DOI?
  - Filter by Content Model (dataset, document, article, etc.)
  - See, for example, the Roman Coins Portal

- Do all file-level DOIs have a landing page?

- Can we do even finer-grained linking, e.g. a subset of a dataset (only a few variables), a paragraph in a book, a video clip, etc?

Thanks for Listening

Questions?