Collaborative Markup of Library and Research Data

Examples from Ontario Council of University Libraries (OCUL)

OCUL and Scholars Portal

- Consortium of 21 Universities in Ontario

- Scholars Portal Project

- Shared infrastructure and technical resources

- 9 services and digital collections (journals, books, e-theses, and more)

Data Services and Collections

<odesi>

- Comprehensive metadata and data repository
- Customized user interface, search tool
- Visualize data using Nesstar interface
- Download data for autheticated users

Scholars GeoPortal

- Geospatial data repository
- Customized map view
- Comprehensive search
- Download spatial data

Scholars Portal Dataverse Network

- Research Data Repository Service
- Built on a self-service model
- Assign persistent identifiers and authoritative metadata
- View to long-term preservation (potencial to connect to other software)

http://odesi.ca

ODESI: Data Service Model

- Developed between 2007-2009

- Jointly funded by OCUL and Ontario Buys

- Provides access to data at OCUL schools (members of the Data Liberation Initiative (DLI))

- Build on collections collaboratively

Over 3,300 datasets in ODESI

Socio-economic data

Additional metadata collections from other repositories

Currently, 5 schools contributing to building the collection

Public Opinion Polls

ODESI System Architecture

Detailed Version - Internal
February 2013

End User

odesi Web App
odesi.ca

Load Balancer

MarkLogic

ODESI Admin
Web App

search1.odesi.ca
MarkLogic Server

MarkLogic
DDI XML DB

search2.odesi.ca
MarkLogic Server

MySQL
Entitlements DB

Nesstar

odesi1.scholarsportal.info
Nesstar Server

MySQL
Nesstar DB

odesi2.scholarsportal.info
Nesstar Server

Nesstar Publisher
Desktop App

OCUL/SP
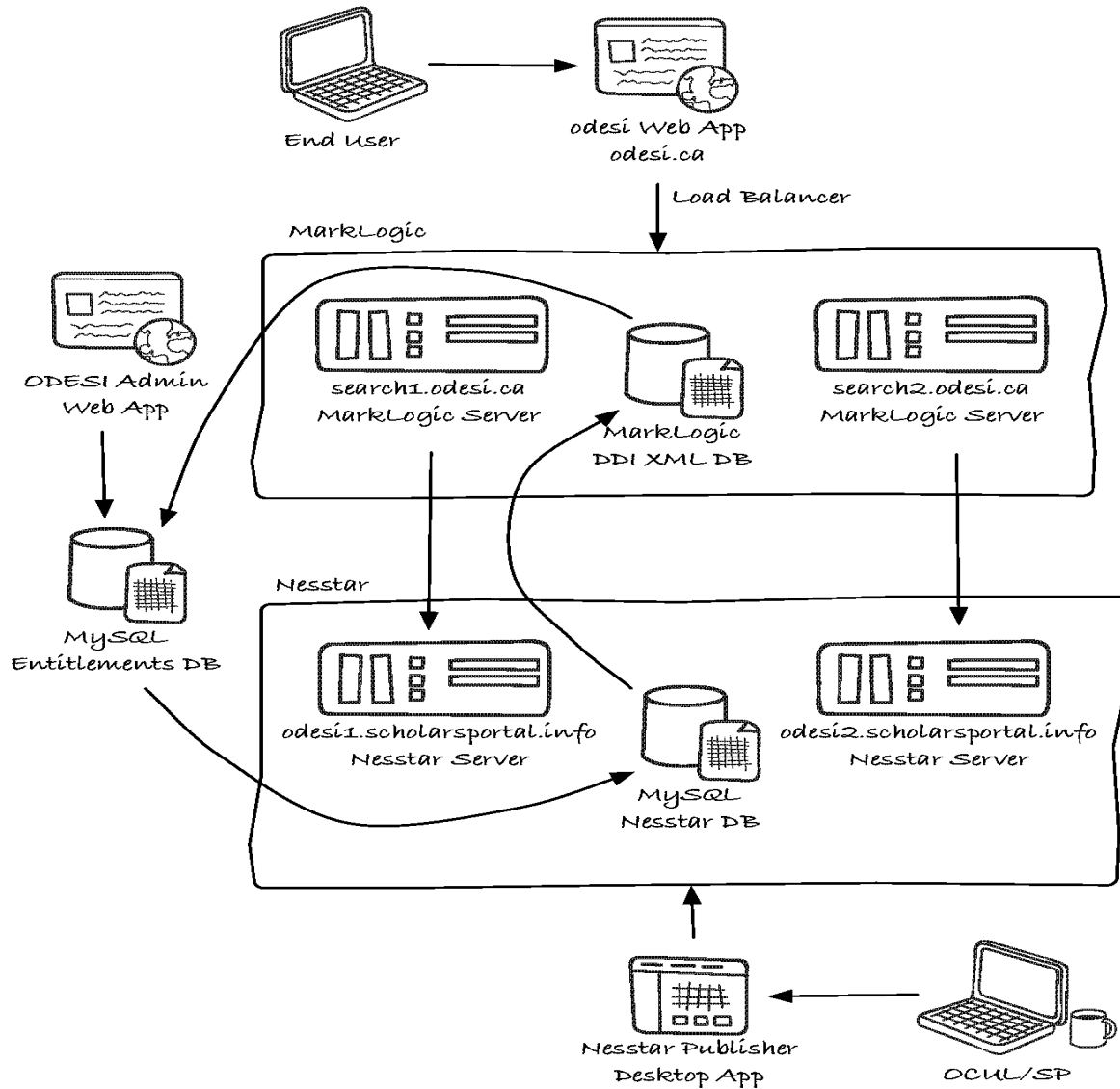
ODESI Metadata

- Nesstar Publisher 4

- Used by many European data archives & UN

- Creates DDI-Codebook metadata

- Allows for distributive access to central server

Nesstar Publisher 4
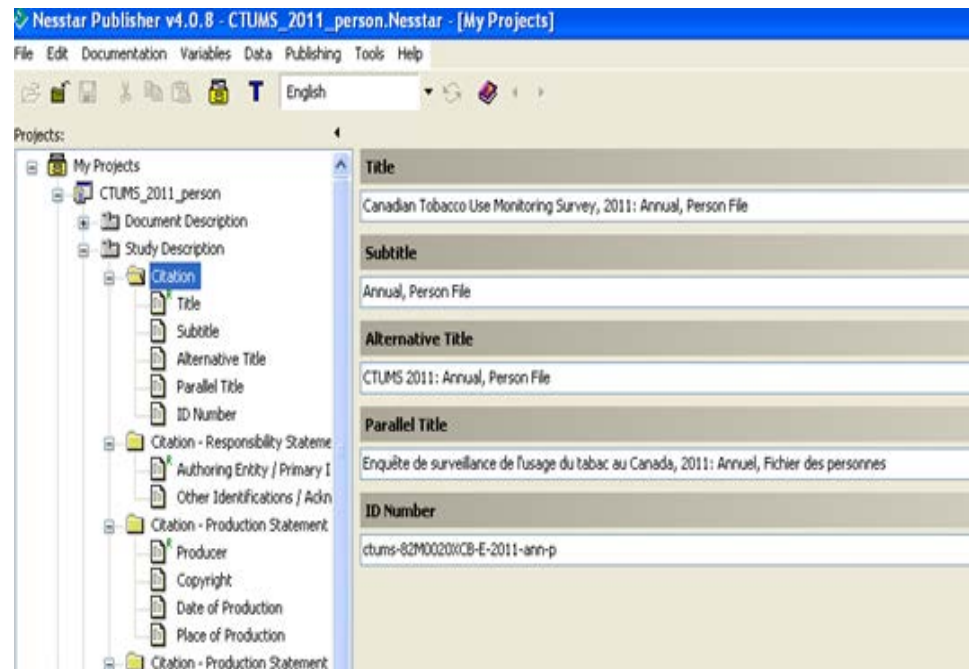
Distributive Markup Model

- Mark-It! Program and funding

- Nesstar Publisher (at each school)
  - ➢ schools ingest and markup data files
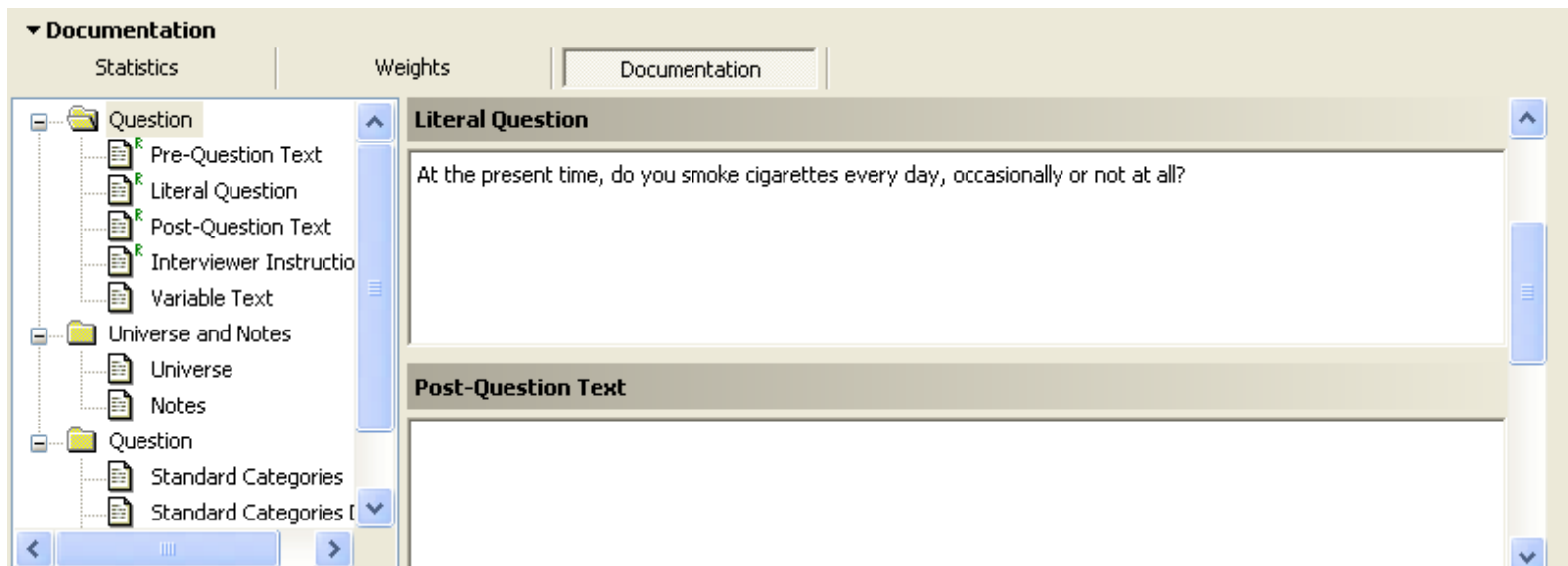
Markup Contribution (%), by school

Legend:
- Carleton University — 31%
- Statistics Canada — 19%
- Guelph University — 14%
- Queen's University — 12%
- University of Ottawa — 11%
- Scholars Portal — 9%
- Wilfrid Laurier University — 2%
- University of Toronto — 1%
- McMaster University — 1%
- Laurentian University
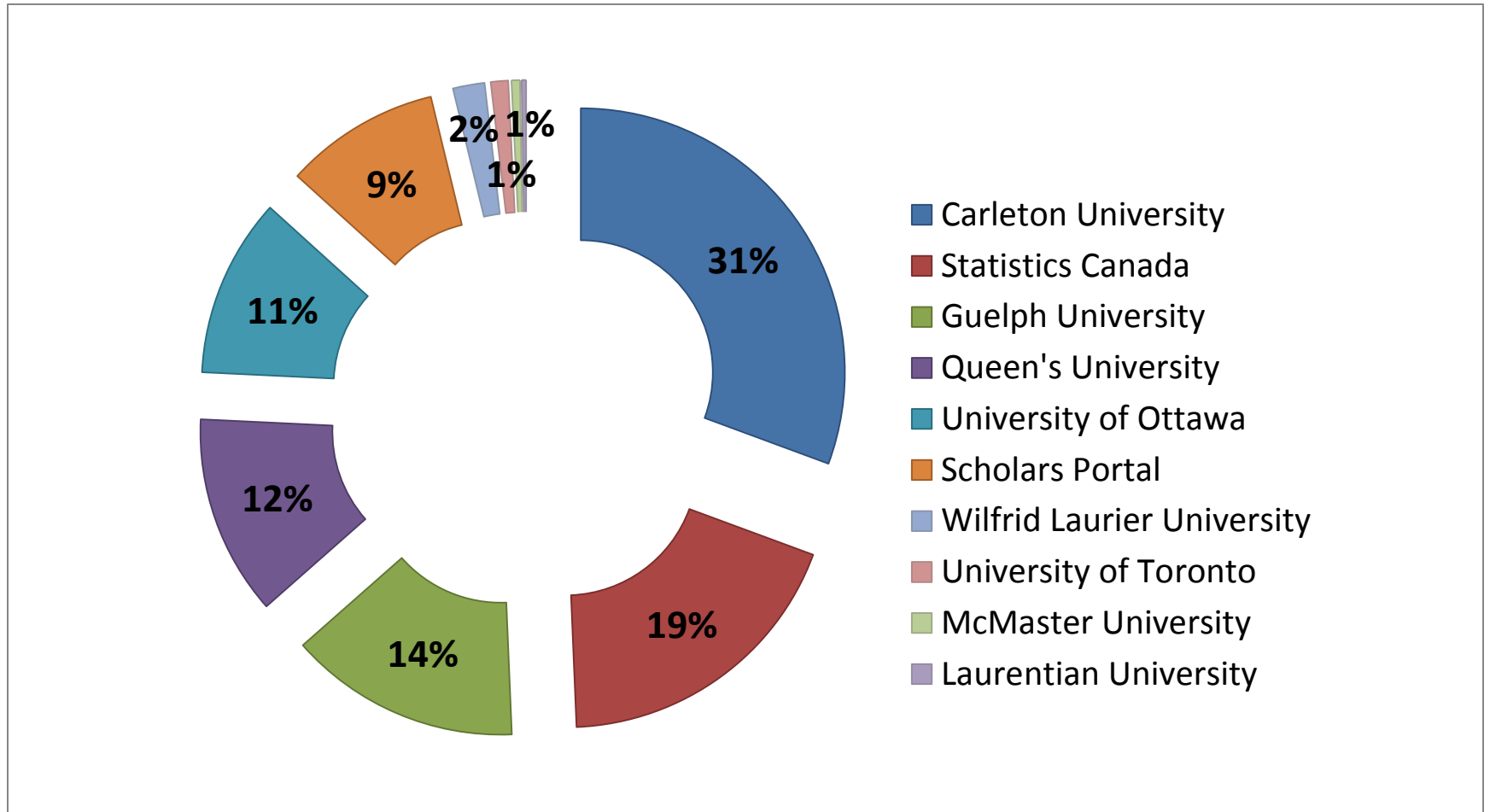
*contributing schools makeup 38% of all OCUL schools

Collaborative Markup Breakdown

35%

65%

■ More than two
■ One
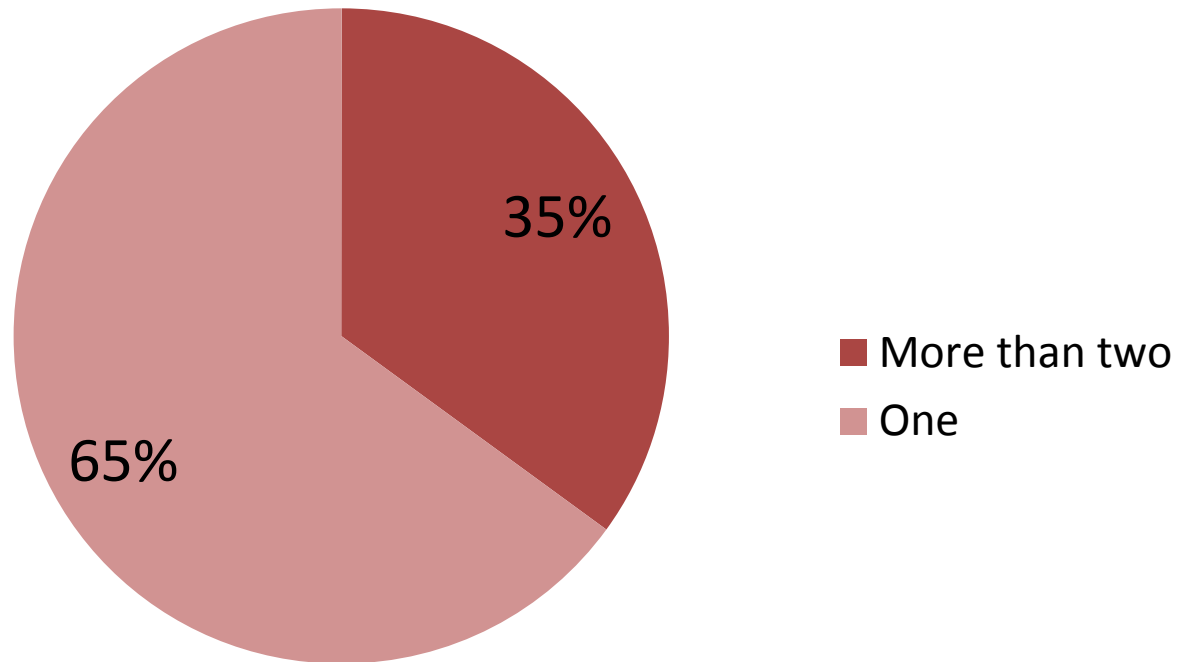
*Colloborative in the sense, 35% of all datasets in ODESI were marked up by 2 or more schools (not necessarily at the exact same time)

Collaborative Markup Strategies 101

- Adhere to established best practises wherever possible; consistency is important!

- Share the workload; mindful of expertise

- Collaborate to improve records; including marking up files that already exist on ODESI (QC Project)

- Be flexible; high priority datasets take precedence

    - New Public Use Microdata Files (PUMFs)

Dataverse @ Scholars Portal

Dataverse Model

- Self-service model; open to researchers
- Ingest data & metadata
  - ➢ Create authoritative metadata (DDI Codebook)
  - ➢ Assign persistent identifiers;
    - ➢ Cite and share their data
  - ➢ Data is stored and repurposed
- Users can search and browse for data
- Data become linkable; there is a home for it

Collaborative Research Support

- Cases are beginning to arrive;
  - one-time deposits; finished studies
  - ongoing file management; projects / studies
  - collaborative research coordination; early stages
- Researchers utilize the 'drop-box' like self-service model to cite and share their data
- Steady growth; capacity to grow beyond the library as a central coordination site

Agri-environmental research data repository Dataverse

Agri-environmental research data repository

Search Studies [Go]

Sort By: [▼]

Studies: **16** |◁ ◁◁ ◁   **1**   2   ▷ ▷▷ ▷|

Ecological interactions in agroforest ecosystems 1988-2008 [Canada]: Agroforest research

by Gordon, Andrew; Thevathasan, Naresh

Abstract: This data is part of a study monitoring the ecological interactions in an agroforest ecosystem. Experimental plots are located at the Guelph Research Station, Victoria Road Agroforestry Land, Guelph, Ontario. The ...

hdl:10864/10167

3 downloads + analyses

Last Released: Mar 18, 2013

Hybrid Maize Trials 2007 to 2008 [Alma, Elora, and Waterloo, Ontario, Canada]: Bioinformatics and Quantitative Genetics

by Lukens, Lewis

Abstract: To produce adapted maize hybrids, inbred lines are often independently selected from distinct populations. We investigate the genetic basis of two elite maize inbreds' adaptation to a short-season environment ...

hdl:10864/10277

0 downloads + analyses

Last Released: Mar 18, 2013

The Genetic Architecture of Flowering Time and Related Traits in Two Early Flowering Maize Lines 2007 [Canada]: Bioinformatics and Quantitative Genetics
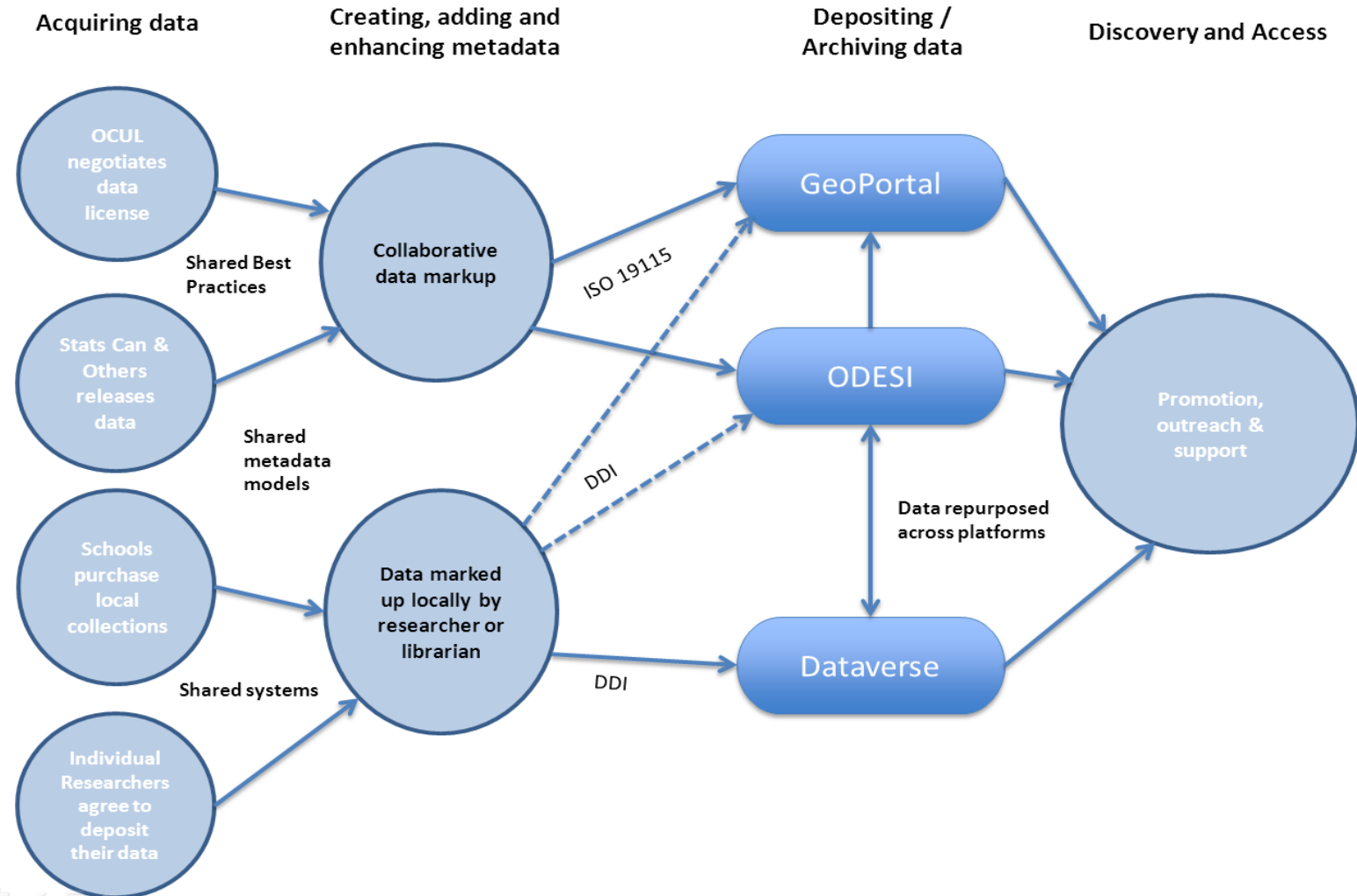
by Lukens, Lewis

hdl:10864/VPCWE

1 download/analysis

Last Released: Jan 23, 2013

OCUL Data Management

Next Steps

- Establish a more robust preservation system for all of our data services

- Utilize / Use our DDI metadata to develop our preservation metadata

- Integrate our data services across domains

- Look to others for expertise and development
  - Other software platforms, tools etc.

The End!

- Thank you and Questions?