



# Challenges in building an institutional research data catalogue

DataFinder at the University of Oxford

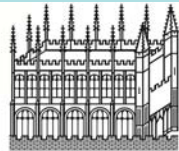
Sally Rumsey & Neil Jefferies  
The Bodleian Libraries

IDCC, Amsterdam, January 2013



# Damaro project

Data management rollout for Oxford



Bodleian Libraries  
UNIVERSITY OF OXFORD

Research Services



DataFinder as the  
key service for  
RDM at Oxford

JISC



UNIVERSITY OF  
OXFORD

# The challenges

## Main areas

1. Metadata
2. User interface design
3. Harvesting and interoperability

## A solution that serves:

1. A wide variety of data sources
2. A variety of data consumers





Absolute perfection?



By [mharrsch](http://www.flickr.com/photos/mharrsch/8331229082/) <http://www.flickr.com/photos/mharrsch/8331229082/>

# 1. Metadata

## Aspect

### Mandatory fields

Minimum core set

Context specific

Tuned to circumstance

### Discipline specificity

### Subject headings

Ease of record creation

### Additional details

## Pragmatic solution

### Manual deposit only

+ DataCite core

### Discipline agnostic

XML files

Common standards

### FAST + Keywords

Auto complete

Harvest

Sell the benefits



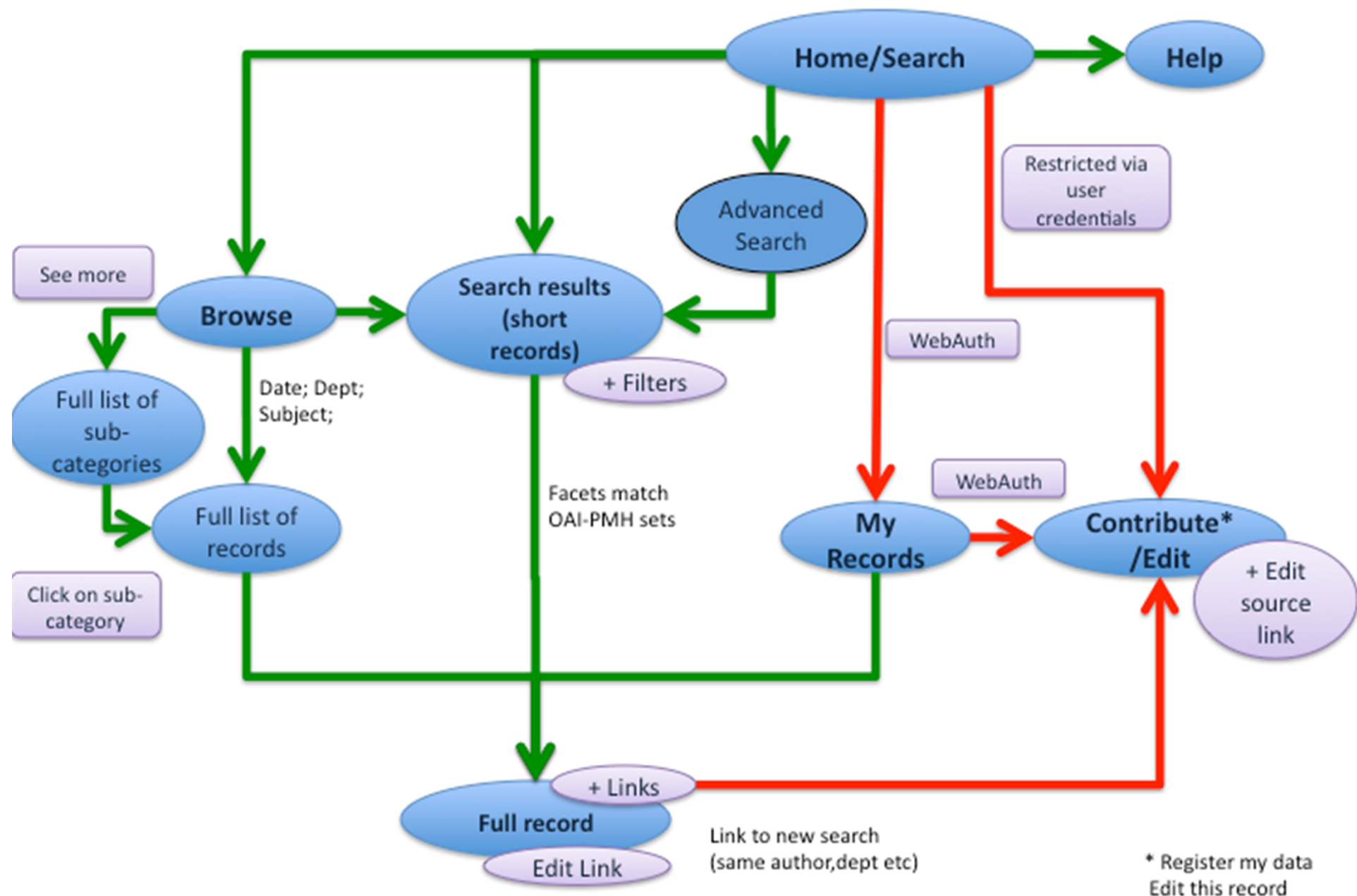
# Minimum mandatory metadata (manual)

<b>Field to be completed by depositor</b>	<b>Notes</b>
Data creator(s) name (given; family)	Other roles available
Data creator affiliation(s)	
Data owner	For curation and management
Is the data an output of funded research?	No (no action required)/Yes (requires funder name & grant number)
Title	
Description/Abstract	
Digital/non-digital	Location required (URL or contact details)
Publication year	
Publisher	Default University of Oxford
Earliest date of access	Embargo set here
Terms and conditions	
Subject	Initially FAST headings

## 2. User interface design

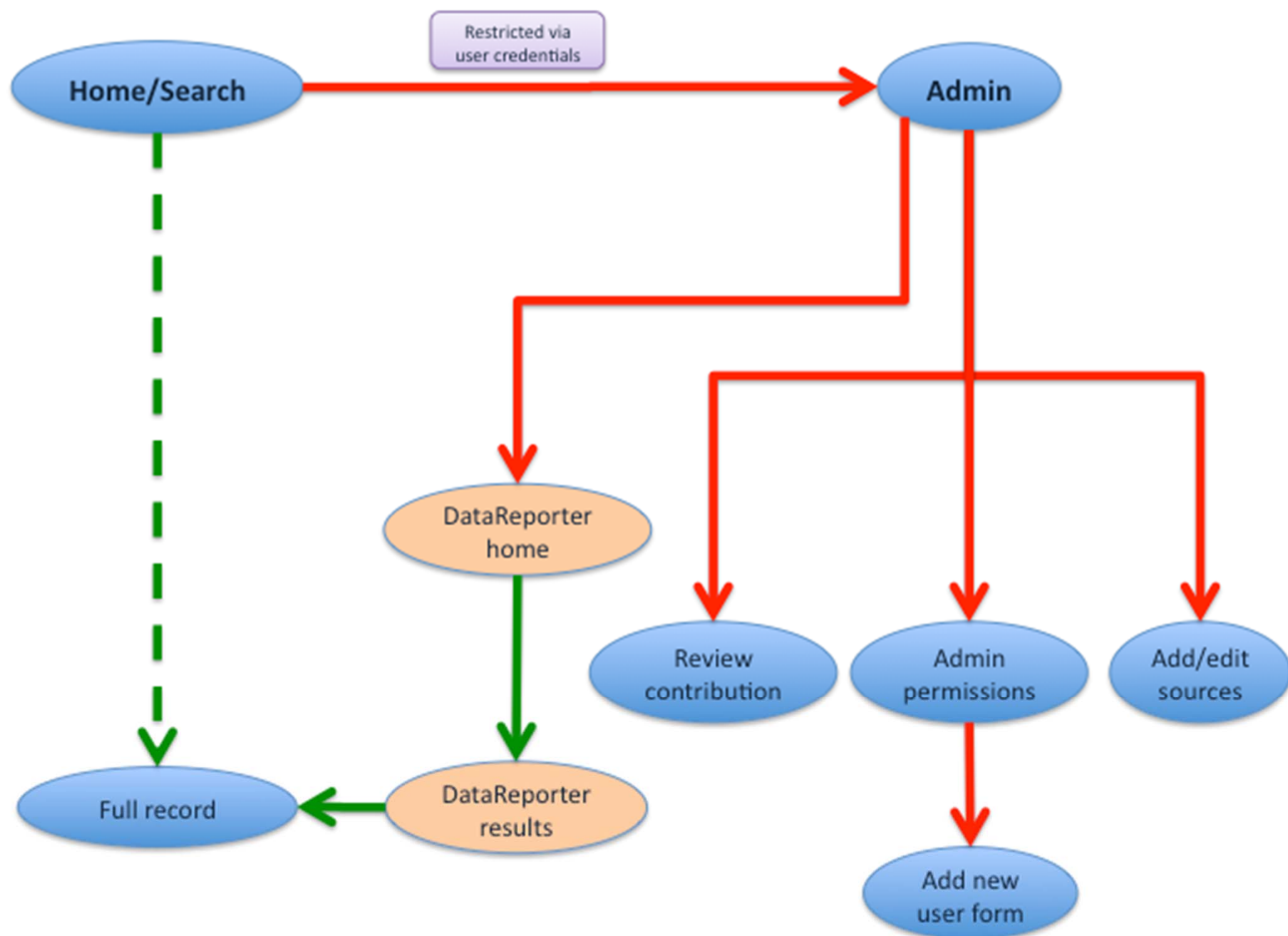
1. Look and feel & branding
2. User journeys
3. Administrative functions

## DataFinder user journeys: search & access





## DataFinder user journeys: Admin



# Record status

**Draft** – user working on record

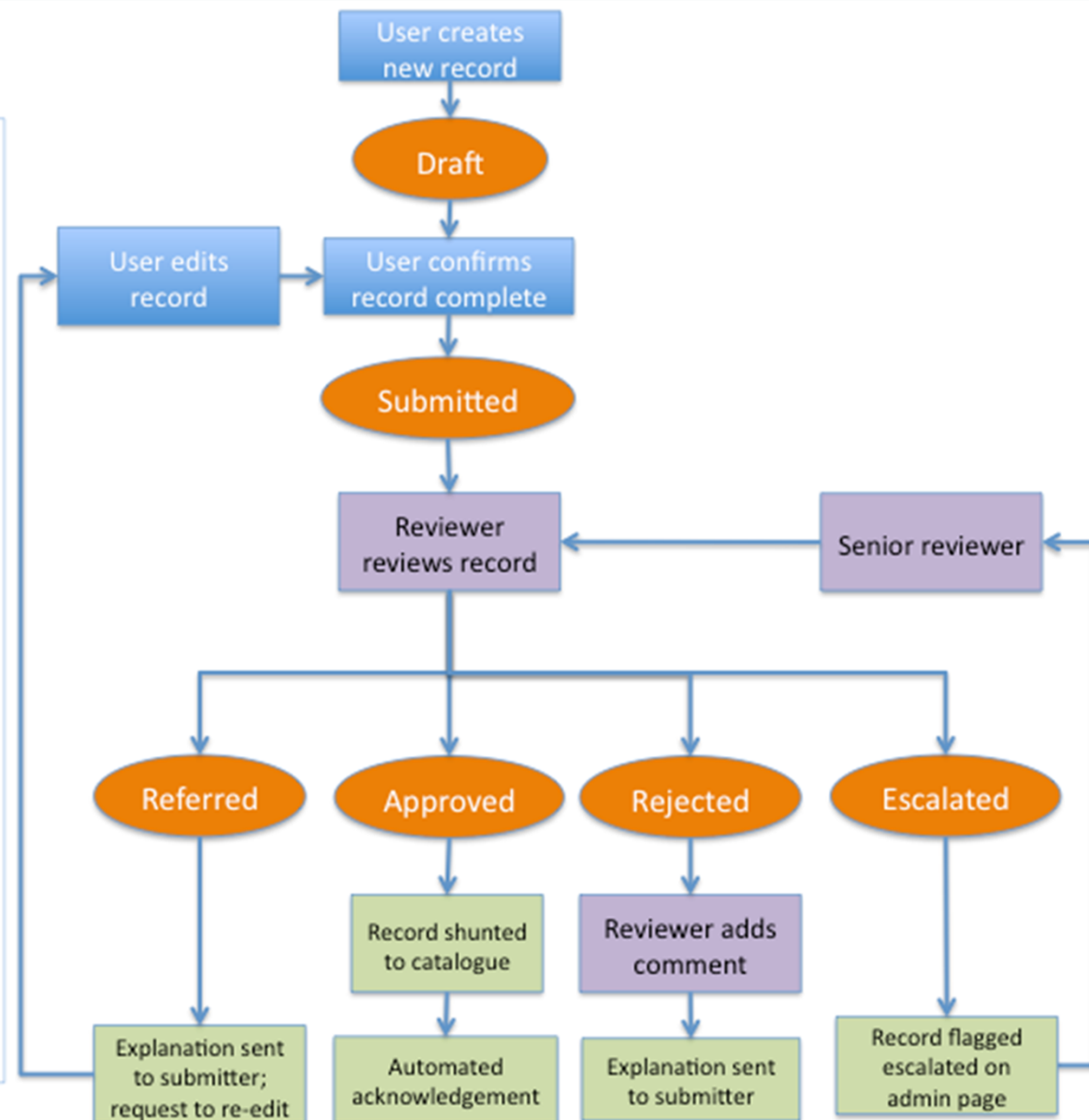
**Submitted** – user has submitted record for review

**Approved** – reviewed submission approved without modification

**Escalated** – reviewed submission to be checked by other staff due to issues such as commercial or legal agreements, ethics. Note of problem added to admin record.

**Referred** – more/better information needed before submission can be approved. Submission returned to the submitter with a note of the problem and how to rectify it

**Rejected** – the admin reviewing the record has decided that there is something fundamentally wrong with it – Reasons for rejection sent to submitter



# Populating DataFinder

Ref: institutional publications repositories

Lack of existing sources

Sell the benefits

‘Create once, use many times’

Citation

Visibility

Impact

Mediated service

Questionnaire approach

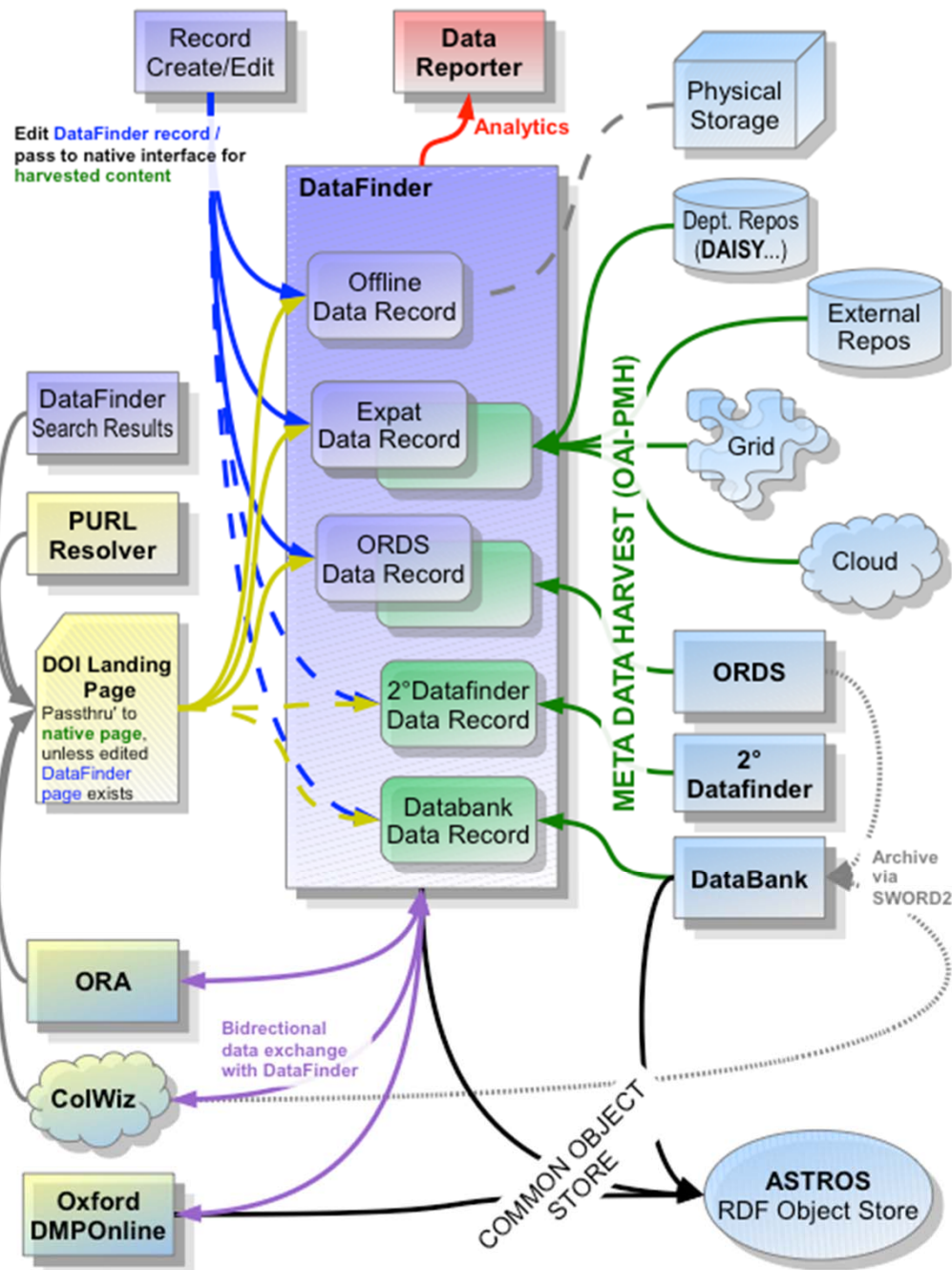
Aim for ‘good enough’ metadata for citation



Bodleian Libraries  
UNIVERSITY OF OXFORD

# Outstanding Challenges

- Obtaining rich, high quality metadata
  - Aim for good enough
  - Encourage rich metadata to generate interest and impact
- Culture change
  - Academics
  - PGRs
- De-duplication
- Sustainability
- Measuring impact
  - Will take time
- Integration with other systems
  - eg Oxford DMPOnline [WIP]



# 3. Harvesting & interoperability

## DataFinder Technical Architecture V1.1

Neil Jefferies,  
The Bodleian Libraries



<http://damaro.oucs.ox.ac.uk/>  
sally.rumsey@bodleian.ox.ac.uk  
neil.jefferies@bodleian.ox.ac.uk