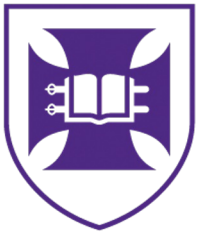
Towards a Generic Research Data Commons: A highly scalable standard-based repository framework for Language and other Humanities data

Peter Sefton[1], Simon Musgrave[2], Nick Thieberger[3]
[1]University of Queensland, Australia; [1,2]Monash University, Australia; [3]University of Melbourne, Australia

The Language Data Commons of Australia (LDaCA) and Australian Text Analytics Platform (ATAP) projects received investment (https://doi.org/10.47486/DP768 and https://doi.org/10.47486/PL074) from the Australian Research Data Commons (ARDC). The ARDC is funded by the National Collaborative Research Infrastructure Strategy (NCRIS).

ARC LIEF LE210100013 (2021-2024) Nyingarn: a platform for primary sources in Australian Indigenous languages



THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

CREATE CHANGE

atap
australian text analytics platform
atap.edu.au

Partner Institutions:

aarnet
Australia's Academic and Research Network

THE UNIVERSITY OF MELBOURNE

MONASH University

Australian National University

THE UNIVERSITY OF SYDNEY

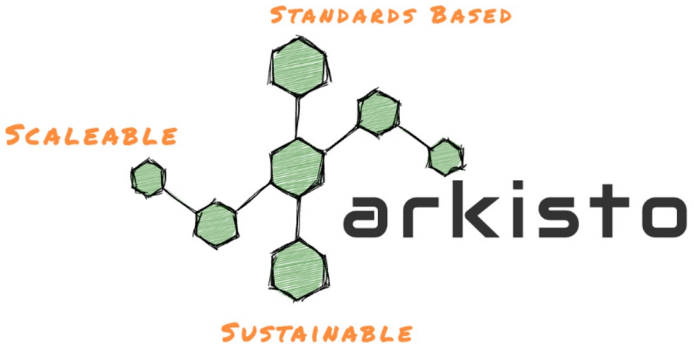
With thanks for their contribution: AIATSIS

DCA

Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC)

- Established 2003
- Researchers concerned to digitise, preserve, and make accessible recordings in the many languages of the region around Australia
- No other agency taking responsibility for these recordings so they were at risk of loss
- Catalog exposes the existence of these recordings, 38,000 items in 690 collections
- Currently represent 1,350 languages, in 205 terabytes, with over 16,000 hours of audio recordings, 3,000 hours of video

arkisto



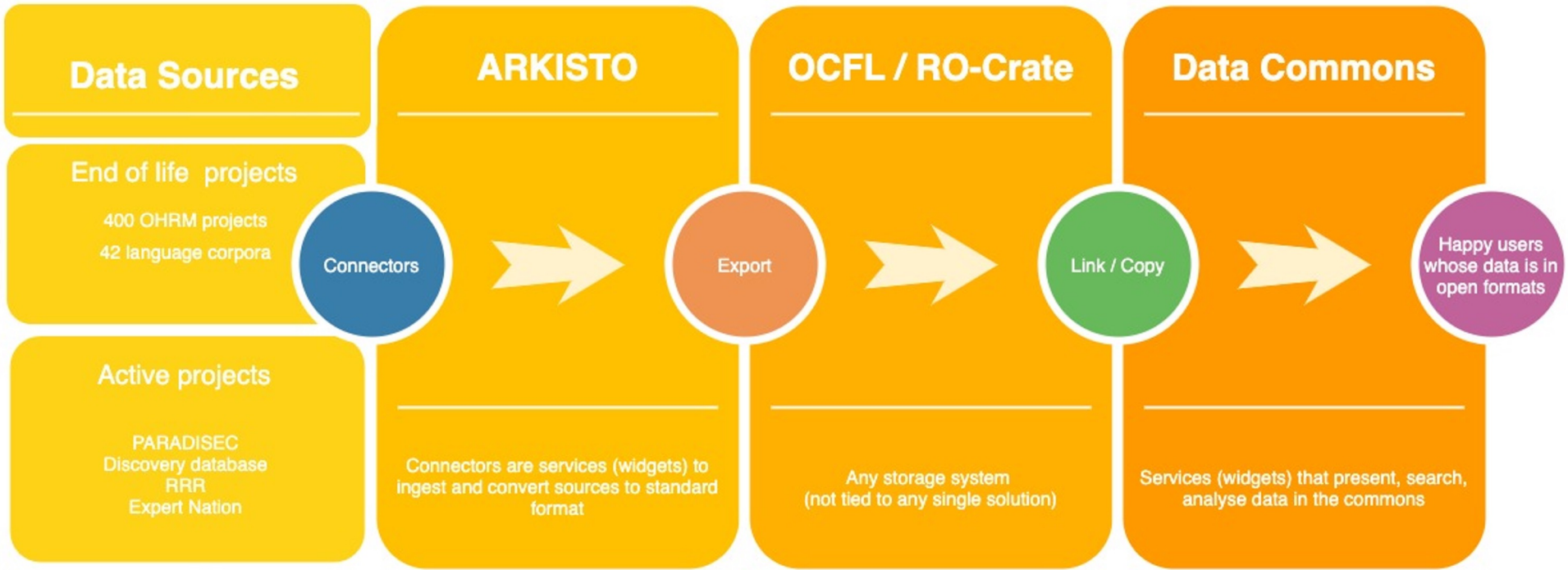
A scaleable, standards based platform for sustainable data.

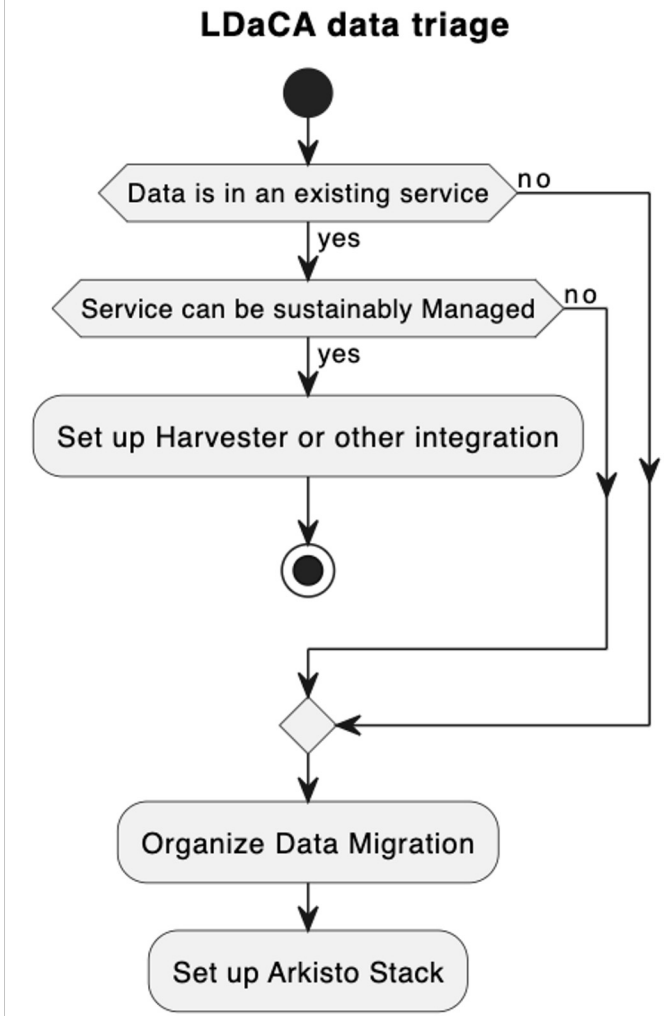
The basis of Arkisto is that the long-term preservability of well-described data is *always* the first consideration.

Data on an Arkisto deployment is alway available on disc (or object storage) with a complete description *independently* of any services such as websites or APIs. Once the data is safe and well described, Arkisto has a flexible model for how data can be accessed using a variety of services.

Arkisto is built on top of Research Object Crate (RO-Crate) and the Oxford Common File System Layout (OCFL).

With Arkisto there is no messy data migration.


Data Sources
End of life projects
400 OHRM projects
42 language corpora
Active projects
PARADISEC
Discovery database
RRR
Expert Nation
Connectors
ARKISTO
Connectors are services (widgets) to ingest and convert sources to standard format
Export
OCFL / RO-Crate
Any storage system
(not tied to any single solution)
Link / Copy
Data Commons
Services (widgets) that present, search, analyse data in the commons
Happy users whose data is in open formats


LDaCA data triage
Data is in an existing service
no
yes
Service can be sustainably Managed
no
yes
Set up Harvester or other integration
Organize Data Migration
Set up Arkisto Stack

Language

- English 509
- Waanyi 396
- Kuku Buyunji 96
- Wakka Wakka^ 4
- Gubbi Gubbi 3
- Bundjalung^ 2
- KUKU YALANJI 2
- Kuku Nyungkul 2
- NOONGAR / NYOONGAR 2
- Worrorra 2
- Yugambeh 2
- Yugarabul 2
- Bibulman 1
- Biri^ 1
- Duungidjawu 1
- GUUGU YIMITHIRR 1
- Gamilaraay / Gamilaroi / Kamilaroi 1
- Giya 1
- Gunggari 1

Yintyingka
Yuru

Fryer Library, The University of Queensland

Contains: Dataset RepositoryCollection

Member Of: UQ Indigenous Language Collection

Language

Biri^
English
Giya
Gubbi Gubbi
Guwamu
Wakka Wakka^
Yuru

Data licenses for access

Default LDaCA No License

Objects: 1

More

Caroline Kelly Papers

Contains: RepositoryObject

Languages: Guwamu English Biri^ Giya Yuru Gubbi Gubbi Wakka Wakka^

Member Of: Fryer Library, The University of Queensland

Personal and professional papers of Caroline Kelly, including correspondence; financial and legal papers; unpublished poetry and stories; theatre records and publications; anthropology field notes, reports and articles; photographs and newspaper cuttings.

UQ Library Collection

Contains: Dataset RepositoryCollection

Member Of: UQ Indigenous Language Collection

Language

English

Data licenses for access

Default LDaCA No License

jupyterhub.atap-binder.cloud.edu.au/user/m.sacalbonequi@uq.edu.au/australis

ATAP Data | cooee - Jupyter Notebook | Home Page - Select or create a notebook | vars.env - Jupyter Text Editor

jupyterhub **cooee** Last Checkpoint: a minute ago (autosaved)

Logout | Control Panel

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3 (ipykernel)

Run | Code | Download

In [9]:
```
# Types of PRIMARY_OBJECTs ie [PRIMARY_OBJECT, X]. What kinds
for entity in ldaca.crate.contextual_entities + ldaca.crate.da
    if 'RepositoryObject' in as_list(entity.type):
        item = ldaca.crate.dereference(entity.id)
        primary_object_types.append(item.as_jsonld())
```

In [10]:
```
import pandas as pd  # this means we will refer to pandas as '

primary_objects_dataframe = pd.json_normalize(primary_object_t
primary_objects_dataframe
```

Out[10]:

	@id	@type	name	dateCreated	
0	arcp://name,cooee-corpus/item/1-001	RepositoryObject	Text 1-001 1788 Phillip, Arthur	1788	'https://data.atap.e
1	arcp://name,cooee-corpus/item/1-002	RepositoryObject	Text 1-002 1788 Phillip, Arthur	1788	'https://data.atap.e
2	arcp://name,cooee-corpus/item/1-003	RepositoryObject	Text 1-003 1788 Phillip, Arthur	1788	'https://data.atap.e
3	arcp://name,cooee-corpus/item/1-004	RepositoryObject	Text 1-004 1788 Phillip, Arthur	1788	'https://data.atap.e
4	arcp://name,cooee-corpus/item/1-005	RepositoryObject	Text 1-005 1788 Phillip, Arthur	1788	'https://data.atap.e
...	...	...	...	...	
1352	arcp://name,cooee-	RepositoryObject	Text 4-421 1897	1897	

jupyterhub.atap-binder.cloud.edu.au/user/m.sacalbonequi@uq.edu.au/australis

ATAP Data | cooee - Jupyter Notebook | Home Page - Select or create a notebo

jupyterhub **cooee** Last Checkpoint: a minute ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help

Run | Code | Download | GitHub | Binder

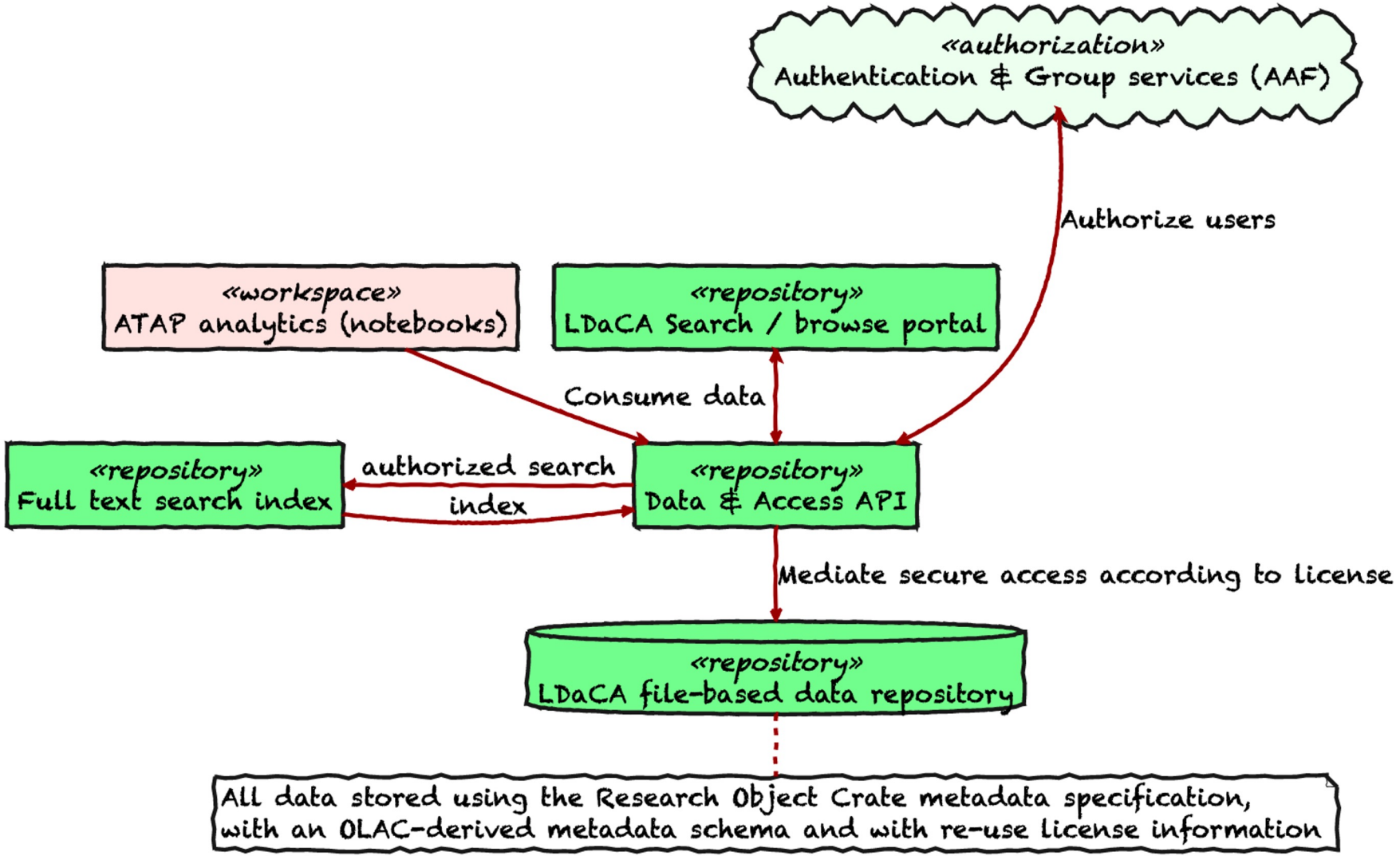
Downloading a file from the ReST API

In [15]:
```
import requests

headers = {"Authorization": "Bearer %s" % API_TOKEN}
response = requests.get(url=url, headers=headers)

print(response.text)
```

<source><g=f><o=b><age=40><status=1><abode=09><p=nsw><r=prw><tt=pc><1-061>
Governor King who has now the command, will make many regulations for the security,
Colony – and likewise some attention to the rising generation, to which hitherto non
if we ever hope for worth or honesty in this settlement, we must look to them for it
e mortals. A school is now establishing on a very extensive plan, for the reception
ren whose parents are not proper for such a charge, under the management of the Govr
en are to be entirely secluded from the other people – and brought up in habits of r
nches of manufactories will be by means of this seminary put on foot particularly ma
the latter to be procured from the Fleece of a remarkable fine breed of Spanish Shee
the former from the Flax which grows spontaneous in the Woods. This with their educa
erent Trades, and the Girls Housewifery and the use of the needle, will be full emplo
s me great satisfaction – as there are now above a thousand children in the place. I
the time when the young Men will become useful members of Society and the Women faith
ryone must hope for our success in so laudable an undertaking – and if no material i
all soon have it on a permanent establishment – I hope when an opportunity offers to
een months since we left England, and I have not heard from any Friend I have. – Col
ly taken up with his two capacities, particularly under the present circumstances, e
the Field with the Men, and I am often lonely enough, and sometimes perhaps fancy th
however with respect to My Dear Sister I am always easy, under your protection I can
only to add Col. P. best respects. [I] f any thing more happens before the sailing o
my sister.
<\1-061><\g=f><\o=b><\age=40><\status=1><\abode=09><\p=nsw><\r=prw><\tt=pc>

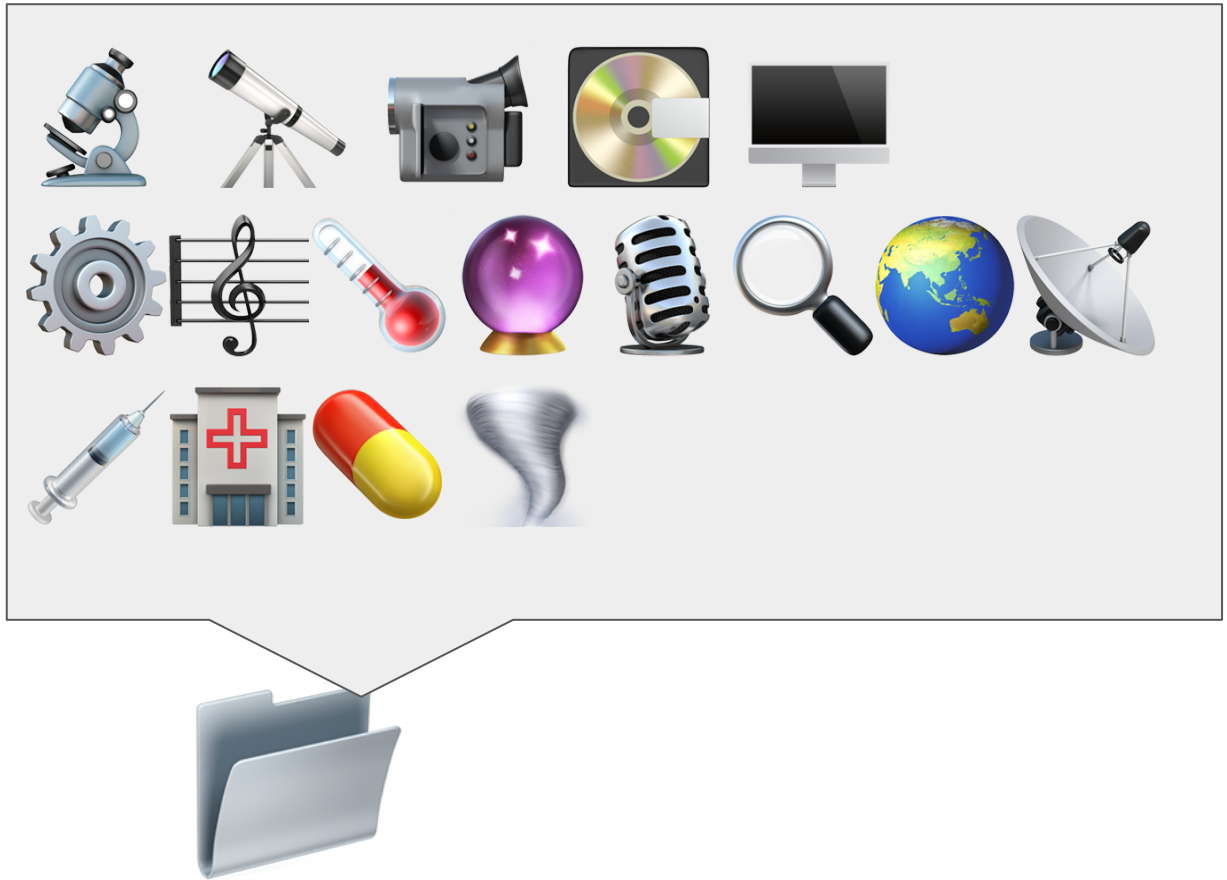

«authorization»
Authentication & Group services (AAF)
Authorize users
«workspace»
ATAP analytics (notebooks)
«repository»
LDaCA Search / browse portal
Consume data
«repository»
Full text search index
authorized search
index
«repository»
Data & Access API
Mediate secure access according to license
«repository»
LDaCA file-based data repository
All data stored using the Research Object Crate metadata specification,
with an OLAC-derived metadata schema and with re-use license information

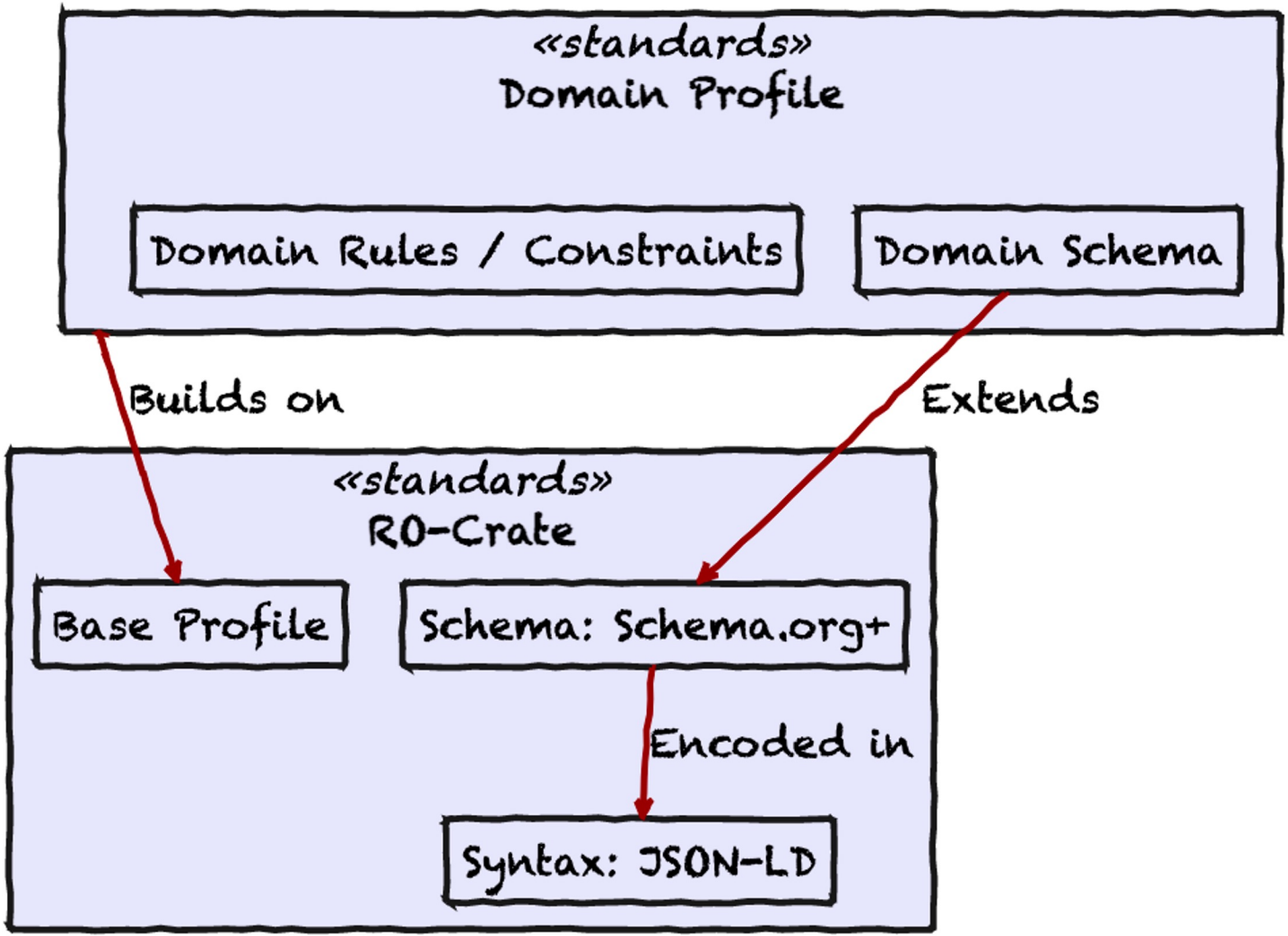
OCFL and “source of truth” — two options

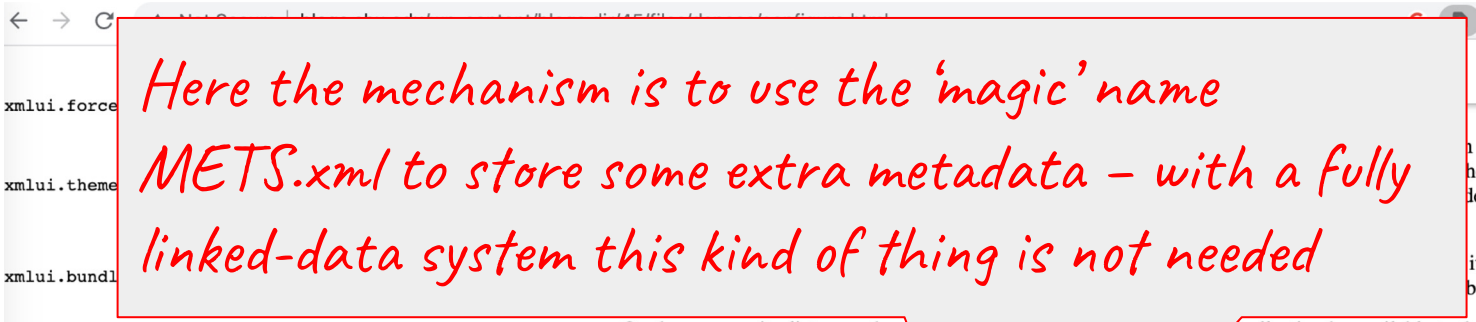
March 21, 2023

Some great things about conferences is how different sessions can play off each other, and how lots of people interested in the same thing are in the same place (virtual or

Bibliographic Wilderness is a blog by Jonathan Rochkind about digital library services, ruby, and web development.
Contact




«standards»
Domain Profile
Domain Rules / Constraints
Domain Schema
Builds on
Extends
«standards»
RO-Crate
Base Profile
Schema: Schema.org+
Encoded in
Syntax: JSON-LD



xmlui.force	
xmlui.theme	
xmlui.bundl	
xmlui.community-list.render.full	True
xmlui.community-list.cache	12 hours
xmlui.bitstream.mods	true
xmlui.bitstream.mets	true

. If set to true, then you need to ensure

submitting a request add the HTTP parameter
he any other configured theme. Note that this is a
development and debugging it should be turned of

item through the administrative interface. If the
be shown to the user as an option.

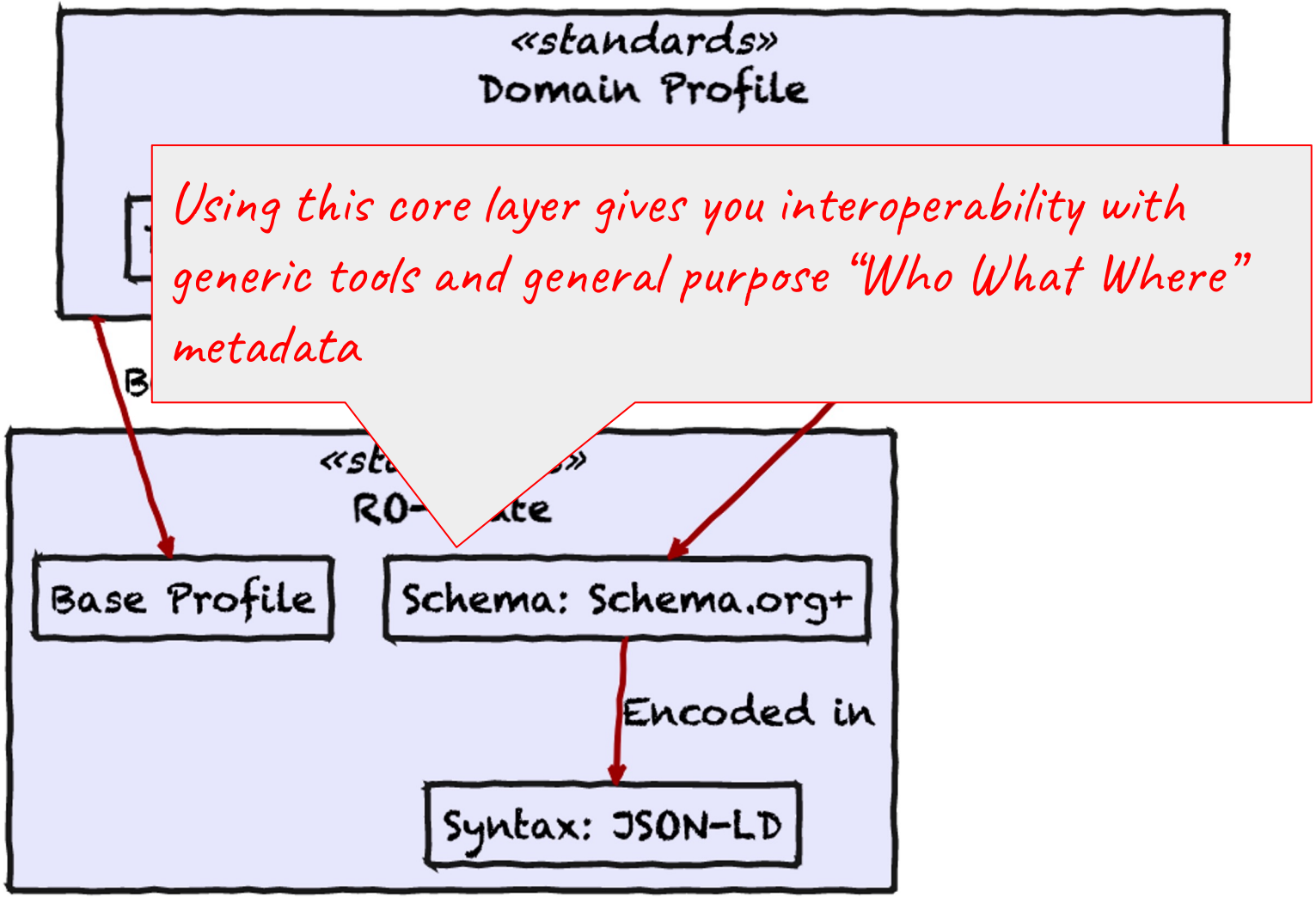
On the community-list page shou /collection be available to the theme. This parameter defaults to true, but if yo
are experiencing performance prob you should experiment with turning this option off.

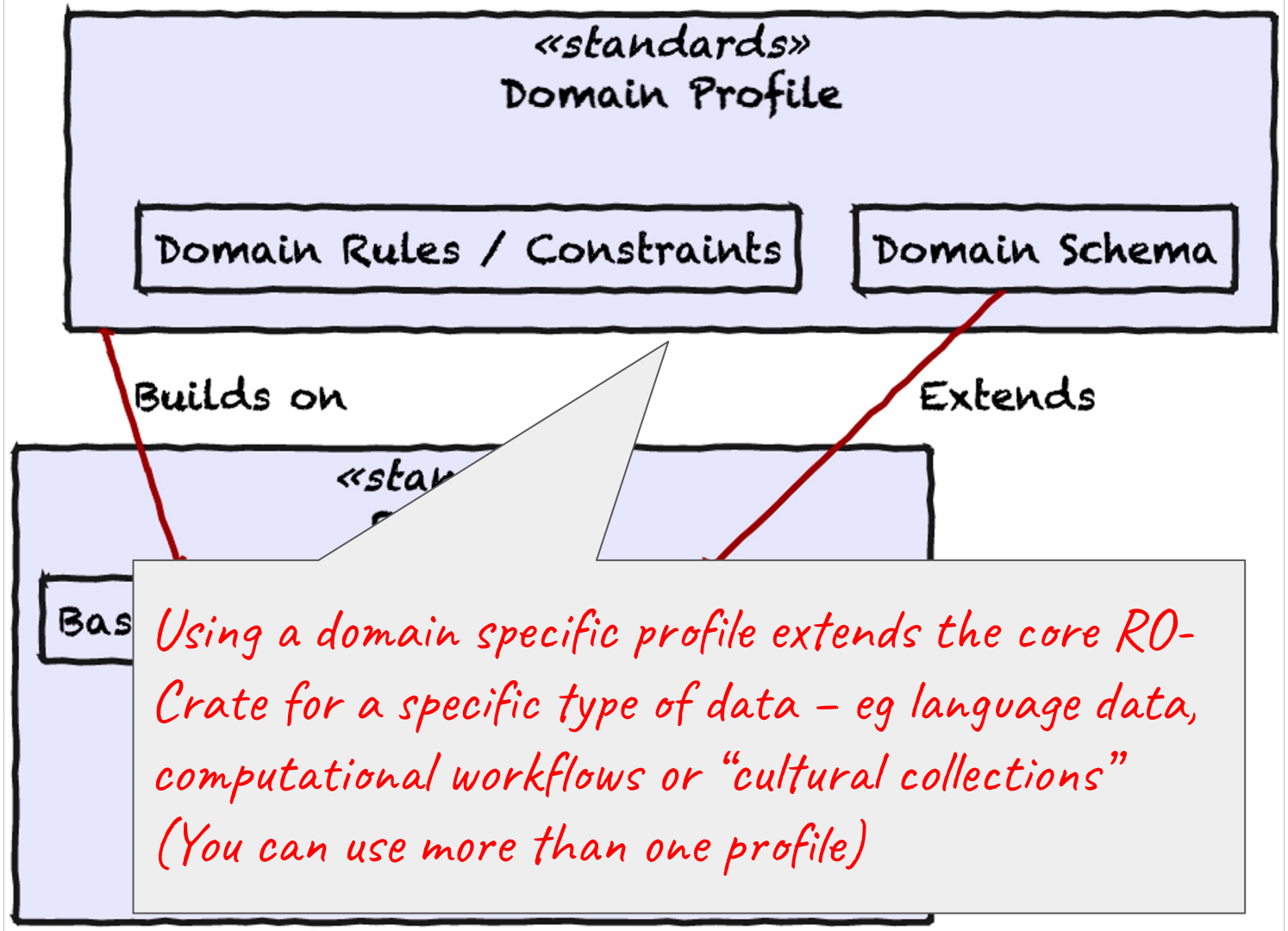
Normally, Manakin will fully verify a cache copy. This means that when the community-list page is viewed the databas
is queried for each community/collec a has been modified. This can be expensive for repositories with a large community
tree. To help solve this problem you ca assumed valued for a specific set of time. The downside of this is that new or editing
communities/collections may not show a period of time.

Optionally you may configure Manakin age of metadata stored as a bitstream. The MODS metadata file must be inside the
"METADATA" bundle and named either l. If this option is set to `true` and the bitstream is present then it is made available to the theme fo
display.

Optionally you may configure Manakin to take advantage of metadata stored as a bitstream. The METS metadata file must be inside the
"METADATA" bundle and named either METS.xml. If this option is set to `true` and the bitstream is present then the stored METS file is merged wit
the METS file generated by Manakin for each item. Thus if the bitstream contains a `dmdSec` then there will be two `dmdSec` one from the bitstream and
another generated from the Dublin Core stored inside the database.

Configuring Themes and Aspects


«standards»
Domain Profile
Using this core layer gives you interoperability with generic tools and general purpose "Who What Where" metadata
Base Profile
Schema: Schema.org+
Encoded in
Syntax: JSON-LD

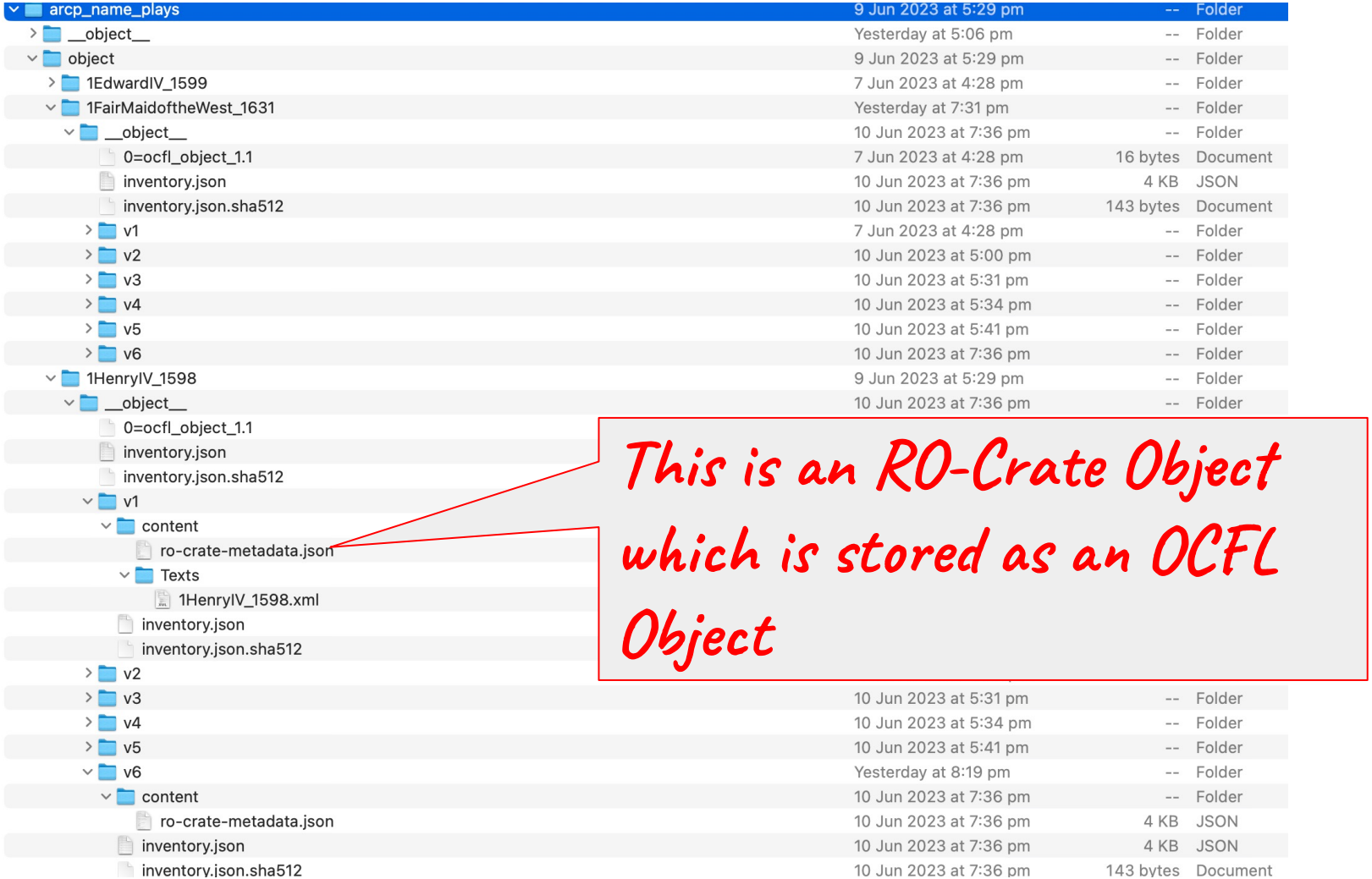

«standards»
Domain Profile
Domain Rules / Constraints
Domain Schema
Builds on
Extends
Using a domain specific profile extends the core RO-Crate for a specific type of data – eg language data, computational workflows or "cultural collections"
(You can use more than one profile)

```
./{collection id}/{item id}
|    ├── {metadata.xml}
|    ├── file1.wav
|    ├── file2.wav
|    ├── file3.mp3
|    ├── ... etc ...
```

The structure an *RO-Crate* MUST follow is:

```
<RO-Crate root directory>/
|   ro-crate-metadata.json    # RO-Crate Metadata File MUST be present
|   ro-crate-preview.html     # RO-Crate Website homepage MAY be present
|   ro-crate-preview_files/   # MAY be present
|    | [other RO-Crate Website files]
|   [payload files and directories]  # 0 or more
```

```
[object root]
    ├── 0=ocfl_object_1.0
    ├── inventory.json
    ├── inventory.json.sha512
    └── v1/
        ├── inventory.json
        ├── inventory.json.sha512
        └── content/
            ├── .fcrepo/
            │   ├── fcr-root.json              <-- Required "header" file holding system metadata about the archival group.
            │   ├── fcr-root~fcr-acl.json      <-- Optional, only present if this Fedora resource has its own ACL.
            │   ├── image.tiff.json            <-- Required "header" file holding system metadata about the binary.
            │   ├── image.tiff~fcr-desc.json   <-- Required "header" file holding system metadata about the binary's descripti
            │   ├── image.tiff~fcr-acl.json    <-- Optional, only present if this Fedora resource has its own ACL.
            │   ├── foo.json                   <-- Required "header" file holding system metadata about the nested container.
            │   ├── foo~fcr-acl.json           <-- Optional, only present if this Fedora resource has its own ACL.
            │   └── foo/                       <-- Required nested structure within .fcrepo/ mirrors content structure
            │       ├── bar.xml.json           <-- Required "header" file holding system metadata about the binary.
            │       ├── bar.xml~fcr-desc.json  <-- Required "header" file holding system metadata about the binary's descripti
            │       └── bar.xml~fcr-acl.json   <-- Optional, only present if this Fedora resource has its own ACL.
            ├── fcr-container.nt               <-- Required file for holding user-properties describing the archival group cor
            ├── fcr-container~fcr-acl.nt       <-- Optional, only present if this Fedora resource has its own ACL.
            ├── image.tiff
            ├── image.tiff~fcr-desc.nt         <-- Required "binary description".
            ├── image.tiff~fcr-acl.nt          <-- Optional, only present if this Fedora resource has its own ACL.
            └── foo/
                ├── fcr-container.nt           <-- Required file for holding user-properties describing the archival part cont
                ├── fcr-container~fcr-acl.nt   <-- Optional, only present if this Fedora resource has its own ACL.
                ├── bar.xml
                ├── bar.xml~fcr-desc.nt        <-- Required "binary description".
                └── bar.xml~fcr-acl.nt         <-- Optional, only present if this Fedora resource has its own ACL.
```


arcp_name_plays 9 Jun 2023 at 5:29 pm -- Folder
__object__ Yesterday at 5:06 pm -- Folder
object 9 Jun 2023 at 5:29 pm -- Folder
1EdwardIV_1599 7 Jun 2023 at 4:28 pm -- Folder
1FairMaidoftheWest_1631 Yesterday at 7:31 pm -- Folder
__object__ 10 Jun 2023 at 7:36 pm -- Folder
0=ocfl_object_1.1 7 Jun 2023 at 4:28 pm 16 bytes Document
inventory.json 10 Jun 2023 at 7:36 pm 4 KB JSON
inventory.json.sha512 10 Jun 2023 at 7:36 pm 143 bytes Document
v1 7 Jun 2023 at 4:28 pm -- Folder
v2 10 Jun 2023 at 5:00 pm -- Folder
v3 10 Jun 2023 at 5:31 pm -- Folder
v4 10 Jun 2023 at 5:34 pm -- Folder
v5 10 Jun 2023 at 5:41 pm -- Folder
v6 10 Jun 2023 at 7:36 pm -- Folder
1HenryIV_1598 9 Jun 2023 at 5:29 pm -- Folder
__object__ 10 Jun 2023 at 7:36 pm -- Folder
0=ocfl_object_1.1
inventory.json
inventory.json.sha512
v1
content
ro-crate-metadata.json
Texts
1HenryIV_1598.xml
inventory.json
inventory.json.sha512
v2
v3 10 Jun 2023 at 5:31 pm -- Folder
v4 10 Jun 2023 at 5:34 pm -- Folder
v5 10 Jun 2023 at 5:41 pm -- Folder
v6 Yesterday at 8:19 pm -- Folder
content 10 Jun 2023 at 7:36 pm -- Folder
ro-crate-metadata.json 10 Jun 2023 at 7:36 pm 4 KB JSON
inventory.json 10 Jun 2023 at 7:36 pm 4 KB JSON
inventory.json.sha512 10 Jun 2023 at 7:36 pm 143 bytes Document
This is an RO-Crate Object which is stored as an OCFL Object

Oxford Common File Layout

OCFL Specifications

This Oxford Common File Layout (OCFL) specification describes an application-independent approach to the storage of digital information in a structured, transparent, and predictable manner. It is designed to promote long-term object management best practices within digital repositories.

Specifically, the benefits of the OCFL include:

- **Completeness**, so that a repository can be rebuilt from the files it stores
- **Parsability**, both by humans and machines, to ensure content can be understood in the absence of original software
- **Robustness** against errors, corruption, and migration between storage technologies
- **Versioning**, so repositories can make changes to objects allowing their history to persist
- **Storage diversity**, to ensure content can be stored on diverse storage infrastructures including conventional filesystems and cloud object stores

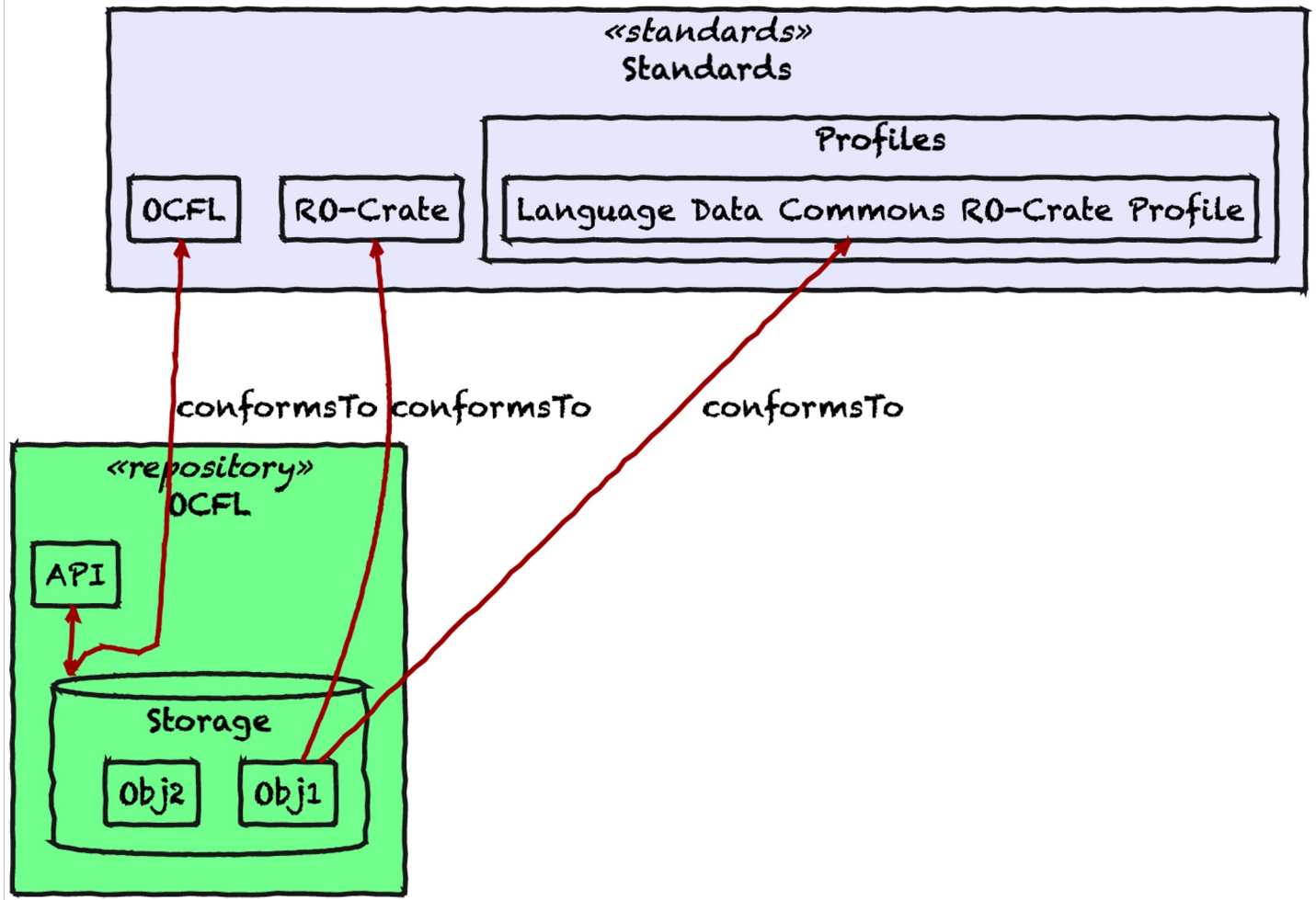
News

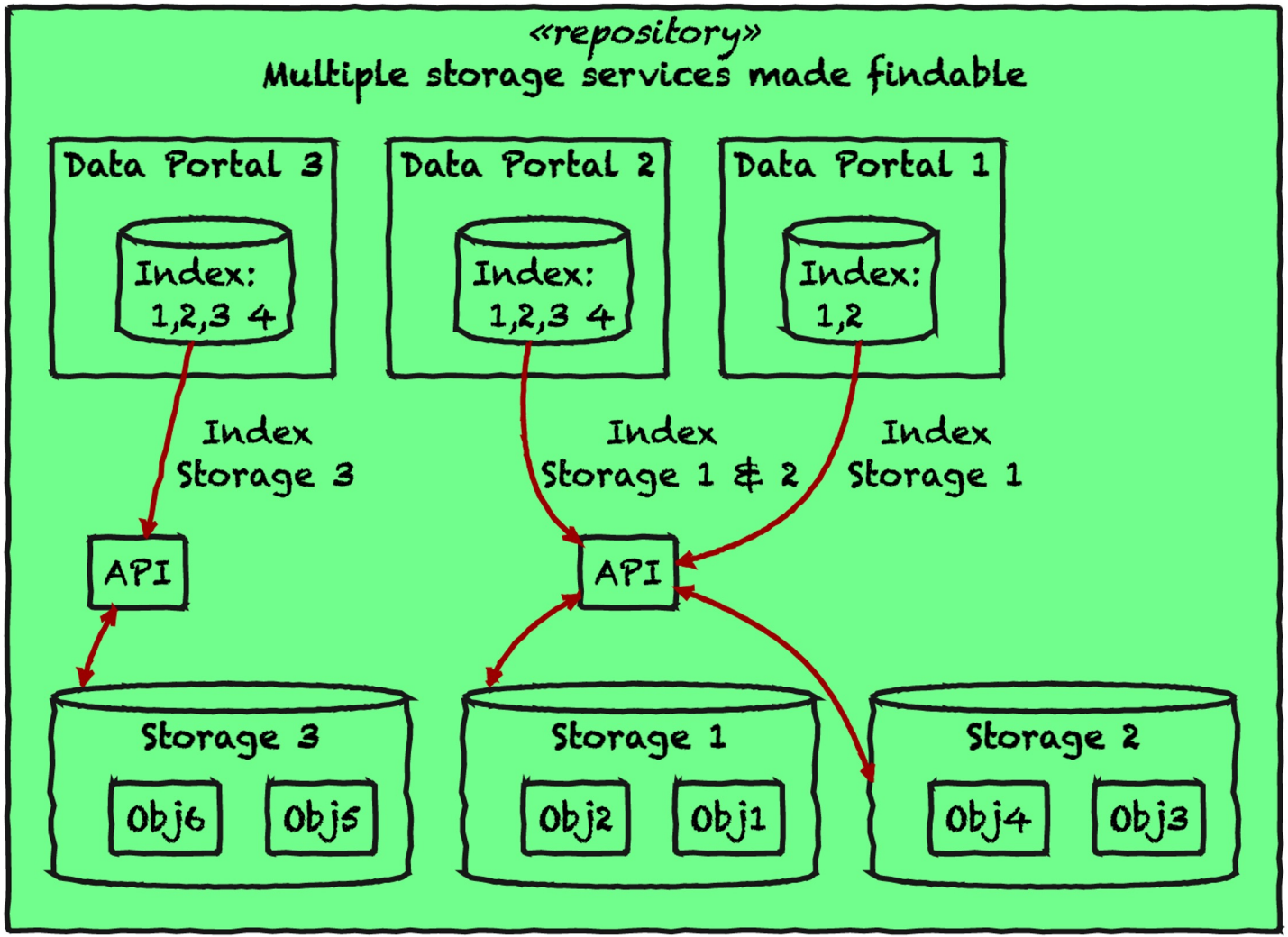
- 2022-10-07: Version 1.1 Release Announcement
- 2020-07-07: Version 1.0 Release Announcement

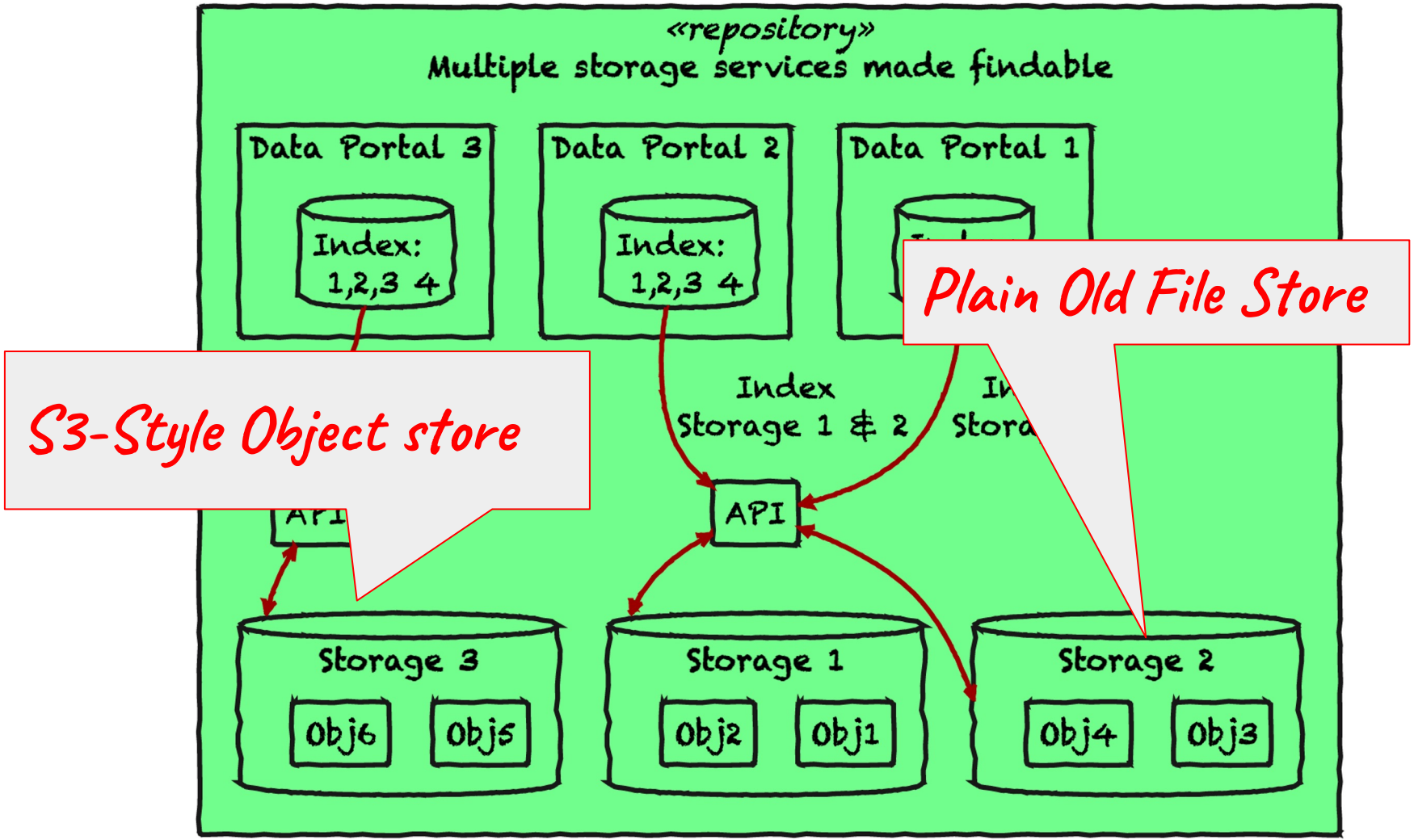
Latest Release (1.1)

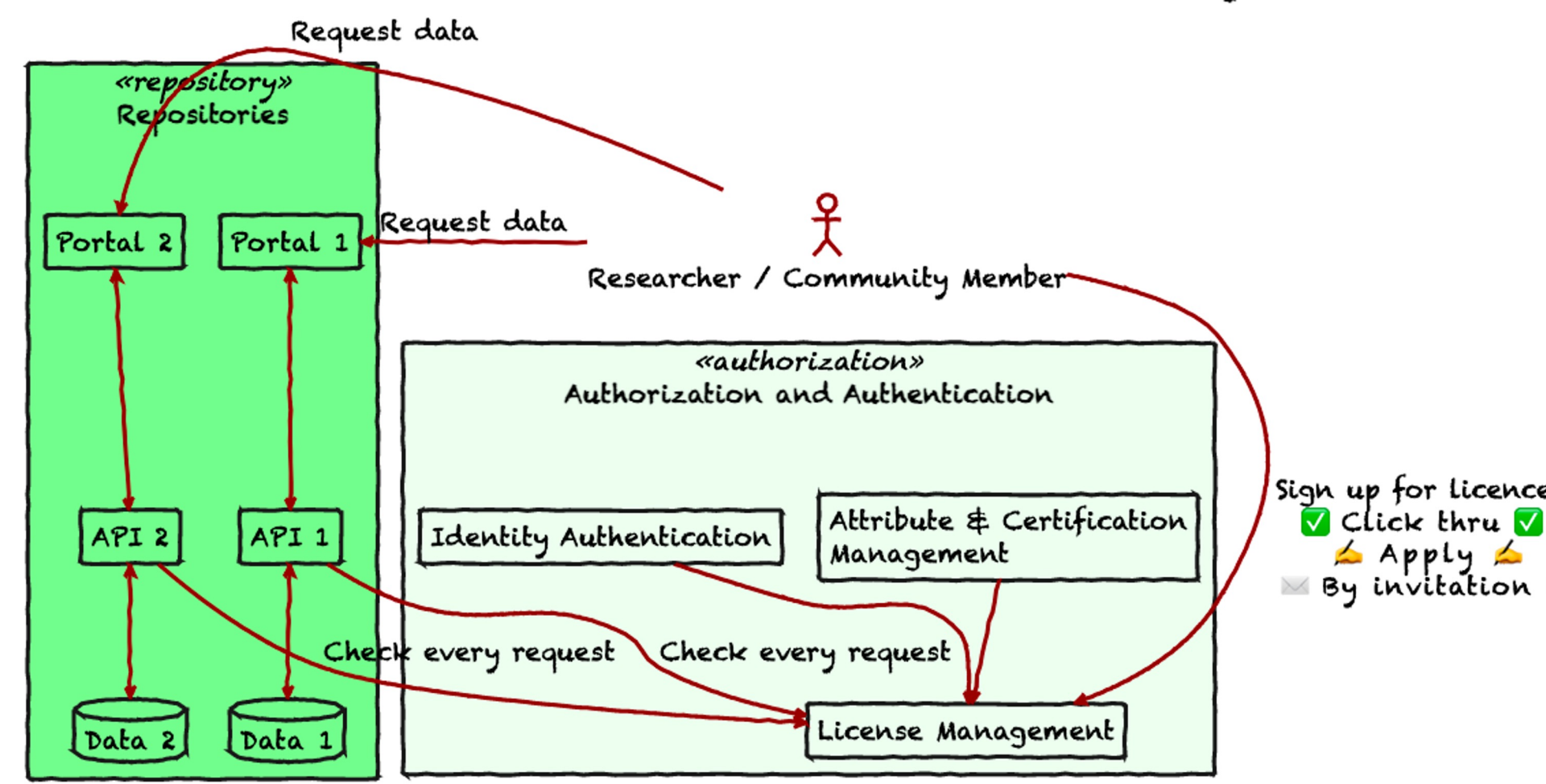
- OCFL Specification v1.1
- OCFL Implementation Notes v1.1
- OCFL Specification v1.1 Change Log
- OCFL Validation Codes v1.1

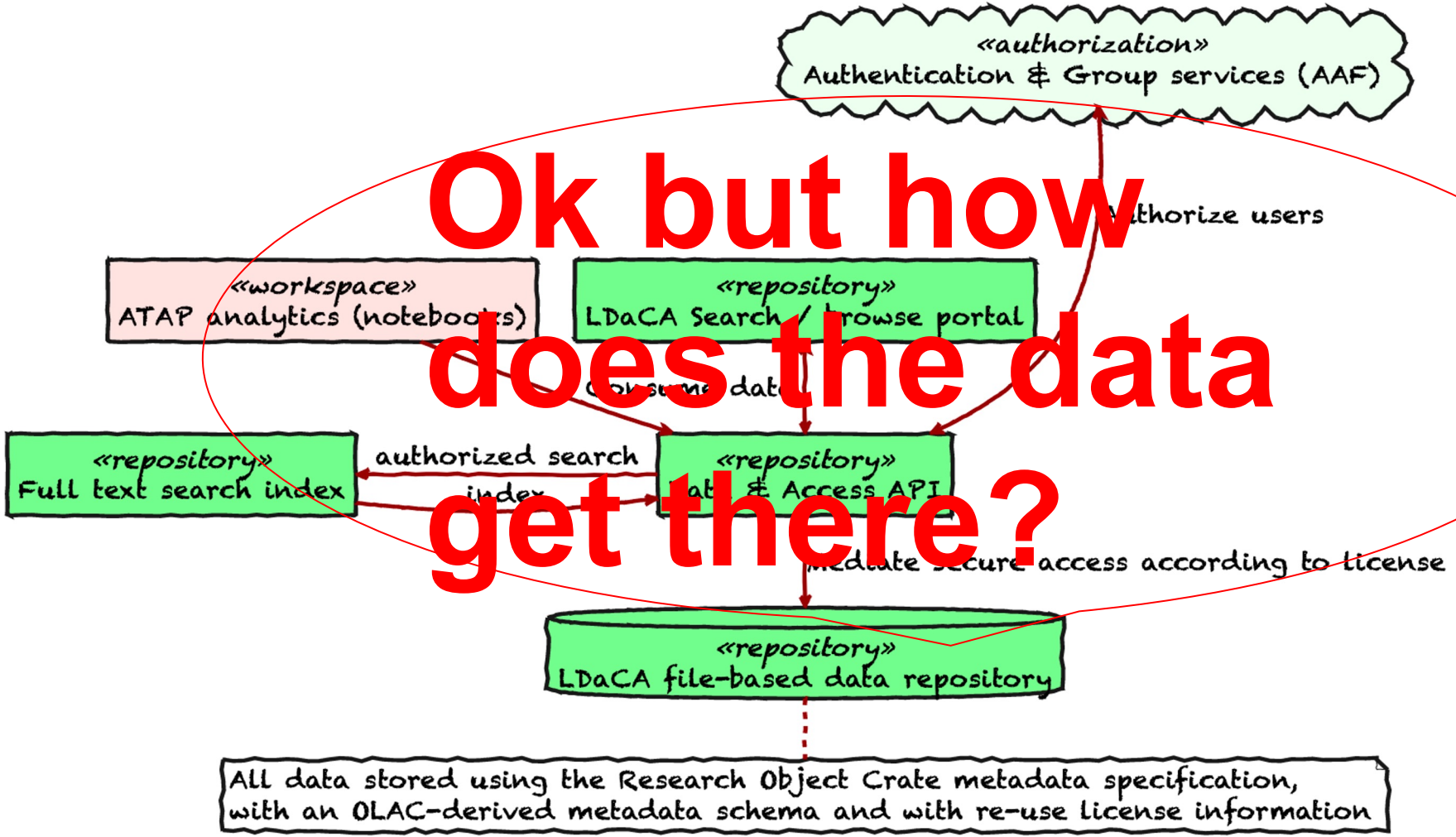
OCFL contains RO-Crate




«repository»
Multiple storage services made findable
Data Portal 3
Index:
1,2,3 4
Data Portal 2
Index:
1,2,3 4
Data Portal 1
Index:
1,2
Index
Storage 3
Index
Storage 1 & 2
Index
Storage 1
API
API
Storage 3
Obj6
Obj5
Storage 1
Obj2
Obj1
Storage 2
Obj4
Obj3

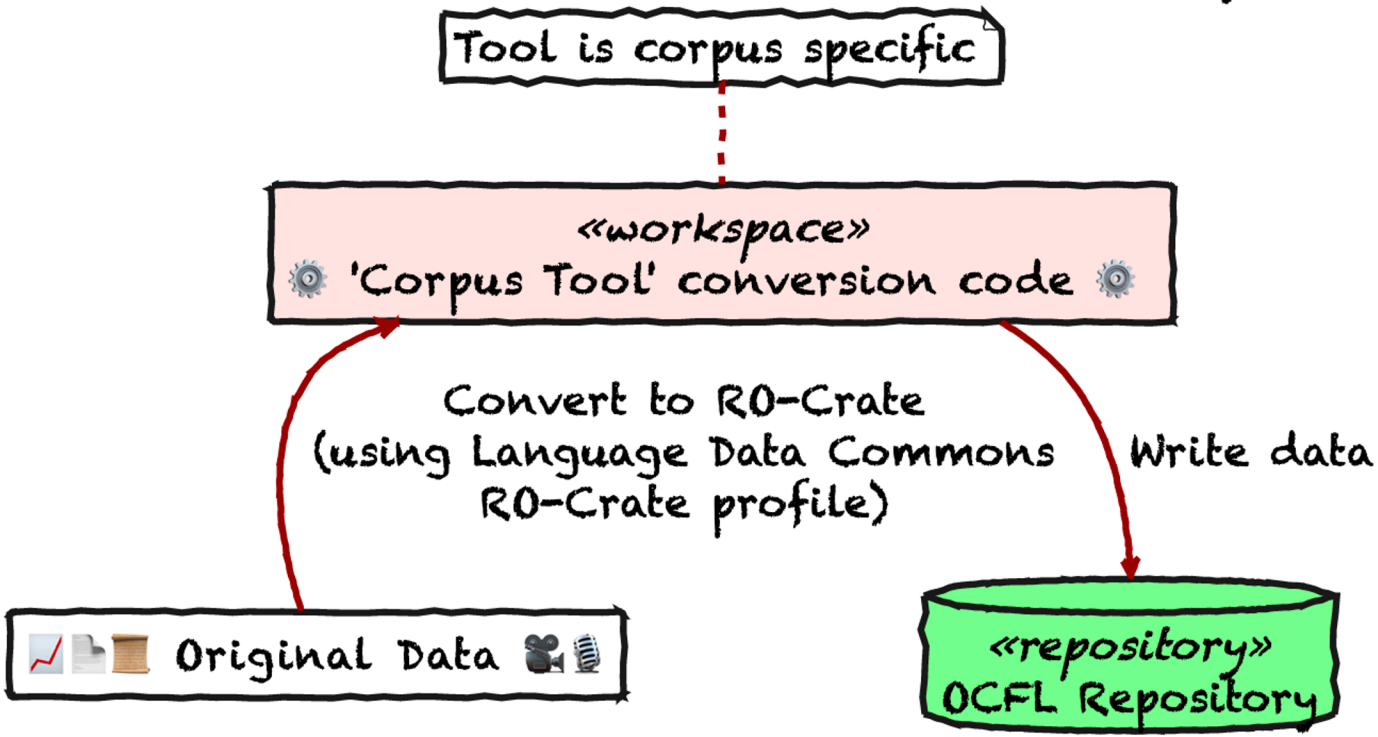

«repository»
Multiple storage services made findable
Data Portal 3
Index:
1,2,3 4
Data Portal 2
Index:
1,2,3 4
Data Portal 1
Plain Old File Store
S3-Style Object store
Index
Storage 1 & 2
API
Storage 3
Obj6
Obj5
Storage 1
Obj2
Obj1
Storage 2
Obj4
Obj3


Request data
«repository»
Repositories
Portal 2
Portal 1
Request data
Researcher / Community Member
«authorization»
Authorization and Authentication
Identity Authentication
Attribute & Certification Management
Sign up for licence
✅ Click thru ✅
✍️ Apply ✍️
✉️ By invitation
API 2
API 1
Check every request
Check every request
Data 2
Data 1
License Management



Ok but how does the data get there?

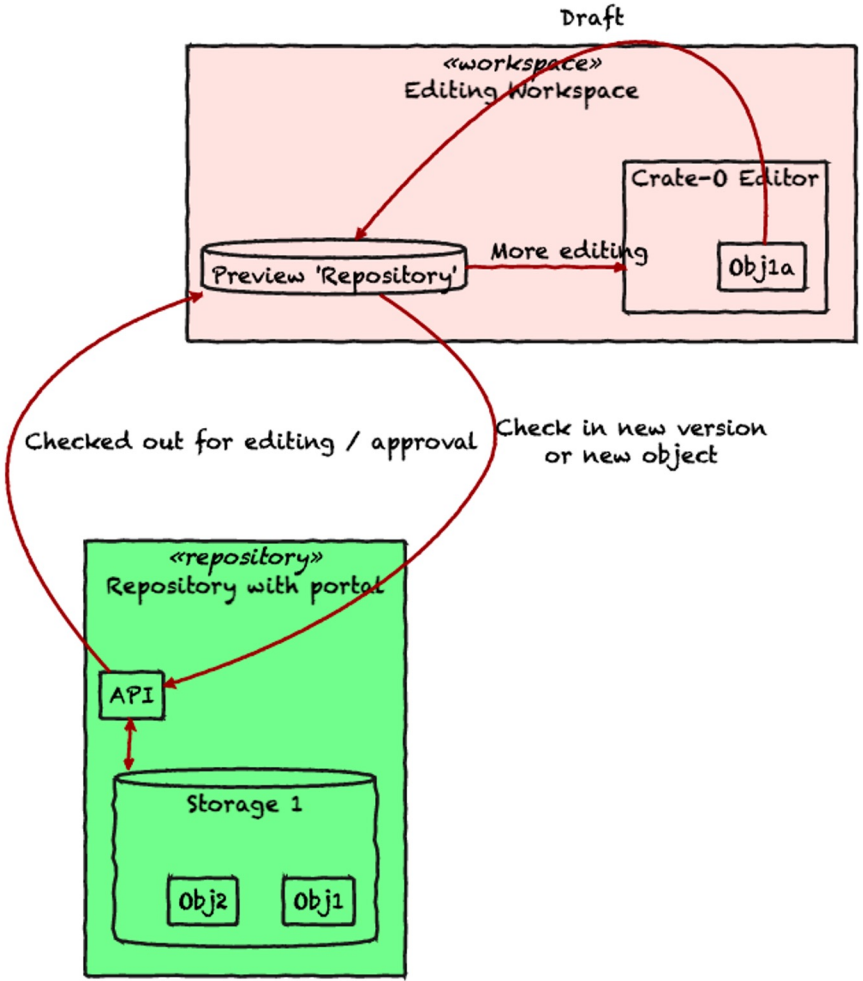
Batch-loading data into LDaCA, simplest view

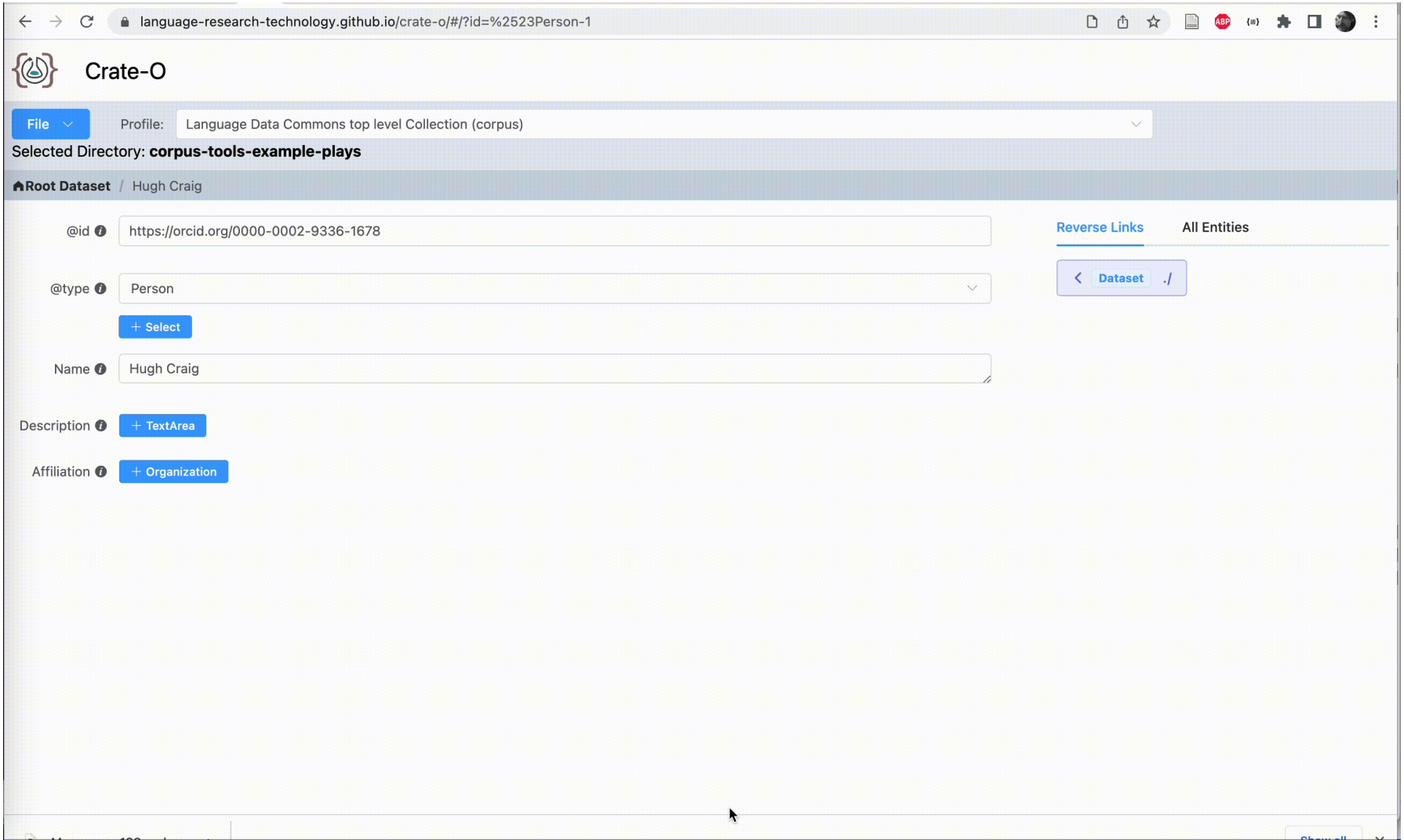


```
{
      "@id": "https://github.com/Lang
      "@type": "SoftwareSourceCode",
      "name": "https://github.com/Lan
      "description": "Converts an RO-                                                    ository Objects and Collections
that are members of a RepositoryColl
      "programmingLanguage": {
        "@id": "https://en.wikipedia.
      }
    },
    {
      "@id": "#provenance",
      "name": "Created RO-Crate using corpus-tools-ro-crate",
      "@type": "CreateAction",
      "instrument": {
        "@id": "https://github.com/Language-Research-Technology/corpus-tools-ro-crate"
      },
      "result": {
        "@id": "ro-crate-metadata.json"
      }
    }
```

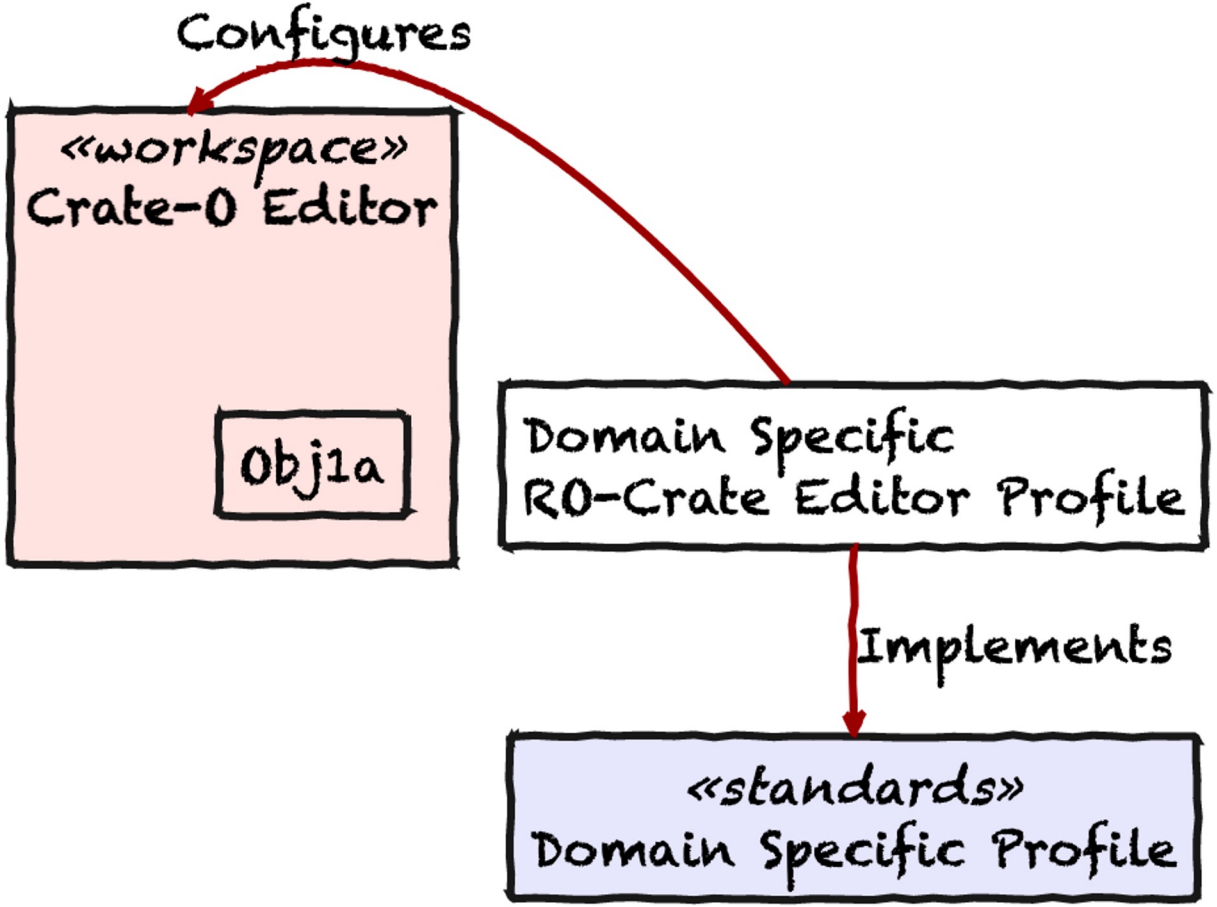

The act of creation of this metadata is documented

Adding Objects using Crate-O (TODO)




language-research-technology.github.io/crate-o/#/?id=%2523Person-1
Crate-O
File
Profile:
Language Data Commons top level Collection (corpus)
Selected Directory: corpus-tools-example-plays
Root Dataset / Hugh Craig
@id
https://orcid.org/0000-0002-9336-1678
@type
Person
+ Select
Name
Hugh Craig
Description
+ TextArea
Affiliation
+ Organization
Reverse Links
All Entities
Dataset ./

RO-Crate Profiles Driving an Editor



A COrpus of Oz Early English (COOEE)

Name	A COrpus of Oz Early English (COOEE)
Description	Material to be included had to meet with a regional and a temporal criterion. The latter required texts to have been produced between 1788 and 1900 in order to become eligible for COOEE. It was mandatory for a text to have been written in Australia, New Zealand or Norfolk Island. But in a few cases, other localities were allowed. For example, if a person who was a native Australian or who had lived in Australia for a considerable time, wrote a shipboard diary or travelled in other countries. Contains: Letters, published materials in book form, historical texts
Date Published	Not Defined
@id	arcp://name,cooee-corpus/corpus/root
Author	Clemens W. A. Fritz
Citation	From English in Australia to Australian English
Temporal Coverage	1788-1900
Conforms To	https://purl.archive.org/language-data-commons/profile#Collection
Identifier	ATAP

Objects in Collection: 1357

Text 1-001 1788 Phillip, Arthur
Text 1-002 1788 Phillip, Arthur
Text 1-003 1788 Phillip, Arthur
Text 1-004 1788 Phillip, Arthur
Text 1-005 1788 Phillip, Arthur
Text 1-006 1788 Phillip, Arthur
Text 1-007 1788 Phillip, Arthur
Text 1-008 1788 Phillip, Arthur
Text 1-009 1788 Bench of Magistrates
Text 1-010 1788 Fowell, Newton

load more...

Access

Attribution 4.0 International (CC BY 4.0)
Public Metadata Indexed

Content

Language
English: 4071
Linguistic Genre
Private Written: 610
Public Written: 405
Government English: 195
Speech Based: 147
Modality
WrittenLanguage: 4071
File Formats
text/plain: 2714

Retrieve Metadata

Download metadata
Open metadata in a new window

Notebooks

cooee notebook

cooee notebook

Description
A sample notebook for the cooee data

@id	cooee.ipynb
Author	Foley, Ben
Conforms To	https://purl.archive.org/language-data-commons/profile#Notebook
Encoding Format	application/x-ipynb+json
Input	A COrpus of Oz Early English (COOEE)

Access

Git Repository
cooee
Notebook Location
https://github.com/Australian-Text-Analytics-Platform/cooee/blob/main/cooee.ipynb

launch binder

Notebook Viewer

[1]:
```
%%capture
import sys
!{sys.executable} -m pip install -r requirements.txt
```

[2]:
```
import json                      # json library to read json file formats
import requests                  # Uses the requests library for REST apis
import os                        # Loads operating system libraries
from ldaca.ldaca import LDaCA    # Loads the LDaCA ReST api wrapper
from rocrate_lang.utils import as_list # A handy utility for converting to list
```

[3]:
```
# Specify location where collection is
LDACA_API = 'https://data.atap.edu.au/api'
COLLECTION_ID = 'arcp://name,cooee-corpus/corpus/root'
```

UTS Research Data Portal

Home Help About

Search

Author

Ian Burnett (7)
Sipei Zhao (7)
Xiaojun Qiu (6)
Anthony Lele (6)
Malcolm Rigby (6)
All...

Keywords

Welding (6)
acoustics (6)
signal processing (6)
international students (4)
First year postgraduate students (3)
All...

DatePublished

2020 (14)
2019 (9)
2021 (9)
2022 (9)

Ambient Vibration of a Cable-Stayed Bridge

This publication is the dataset component of a data paper. A full-scale short-span cable-stayed bridge, located on the top of a wind-exposed hill in the state of the New South Wales (NSW) in Australia, was instrumented to measure its dynamic response to ambient vibrations. The main purpose of the exercise was to generate sufficient ambient vibration datasets necessary for conducting Operational Modal Analysis (OMA). Wind, passing vehicular and pedestrian traffic over the bridge, as well as the vehicles travelling on the highway underneath the bridge provide adequate sources of ambient vibration excitation for this bridge. A dense array of time-synchronised uni-axial accelerometers was permanently mounted on the deck and on the cables of the bridge. Since the structural modal features vary with temperature, the ambient temperature was also continuously recorded. The shear strain response at one end of the bridge was also measured constantly to identify the volume of passing traffic over the bridge. Data acquisition was conducted non-stop for specific periods and the measured data were transferred over a 4G cellular network to the database. It is the intention of the authors that the datasets can be employed for further development and validation of OMA frameworks and will be of interest to the bridge engineering research community.

Hamed Kalhori | Mehrisadat Makki-Alamdari | Bijan Samali | Chul-Woo Kim | Benjamin Halkon | Ambient Vibration Dataset | Bridge Structural Analysis | Cable-Stayed Bridge | Operational Modal Analysis. | 2020

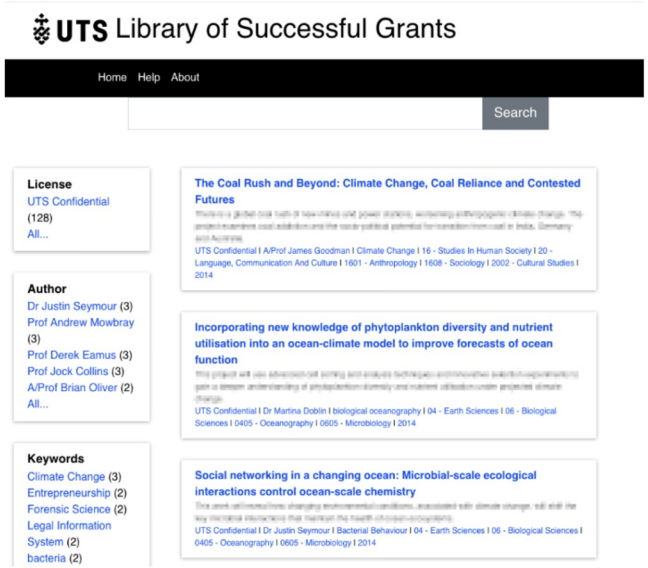
Australian Public Health Orders Issued by Australian State and Territory Governments: Dataset 2004-2017

The powers available to the state in the name of advancing or protecting the public's health or human biosecurity include disease surveillance; the power to compel provision of information; the monitoring, prohibiting or compelling of particular behaviours; involuntary social distancing measures including detention, isolation and quarantine; and, finally involuntary medical testing and treatment. Public health orders are the mechanism used to activate the most coercive aspects of public health and human biosecurity powers in Australia. They exist in some form in each Australian jurisdiction; however, the nomenclature, their availability and associated processes, and the specific ambit of their power differ, at times quite markedly. This dataset relates to a multi-year project that utilised methods of public information audit, administrative engagement and freedom of information processes to collect data on the use of public health and biosecurity powers in Australia. This dataset contains tabular data recording summaries of each reported exercise of a coercive public health power during the period 2004-2017 that were disclosed by each jurisdiction. Each order or action is recorded with textual summary or description of each order or action. This includes date of order, nature and requirements, public health risk addressed, duration of the order, actions/enforcement taken, comments by the researcher on orders and general notes on the data. The data reported here are largely forms of public health order, although warrants for arrest or detention of individuals, alongside other 'enforcement measures', are also included as instances of the use of coercive public health powers. The dataset also includes copies of original documents (often redacted) and correspondence provided by jurisdictions as a result of administrative action or in response to open government/freedom of information processes

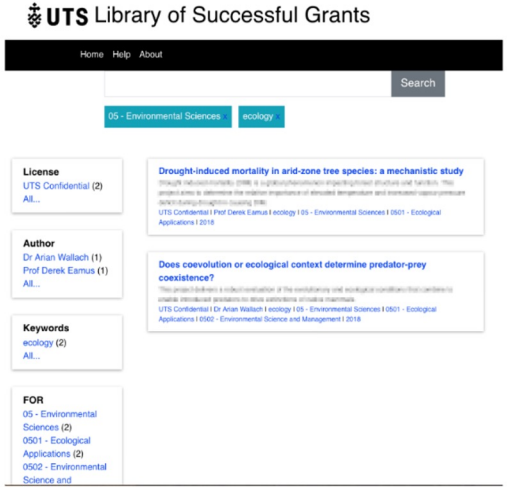
Case Study: UTS Successful Grants Repository

Status: Live Q3 2020 (UTS staff only)

The UTS Successful Grants Repository is searchable repository of successful grant applications to a variety of funding bodies by UTS researchers. These applications are made available by the UTS Research Office to UTS research staff for professional development and provide the basis for research to improve research performance The Successful Grants Repository shows the use of Oni as a platform for indexing document collections with strict access control.



A screenshot of the UTS Library of Successful Grant showing the search



A screenshot of the UTS Library of Successful Grants showing grants related to Environmental Science and Ecology via

UTS Expert Nation

Home Help About

Search

SectorName
Medicine, Dentistry and Health (1385)
Education-Primary & Secondary (500)
Law and Justice (365)
Agricultural,Pastoral,Forestry and Fishery (296)
Architecture, Building and Engineering (293)
All...

MilitaryServicePlace
France (867)
Gallipoli (254)
Egypt (242)
England (139)
Palestinian Territories (40)
All...

UniversityConnections
USyd (2051)
UMelb (1983)
UAdelaide (716)
UQ (345)
Sydney Teachers College (199)
All...

Matters, Reginald Francis
Medicine, Dentistry and Health I France I USyd

Whiteman, Reginald John Nelson
Medicine, Dentistry and Health I France I USyd

Willcocks, George Charles
Medicine, Dentistry and Health I Egypt I USyd

Whiting, James Ernest
Education-Primary & Secondary I France I Sydney Teachers College

White, Roy Sylvester
Education-Primary & Secondary I France I Sydney Teachers College

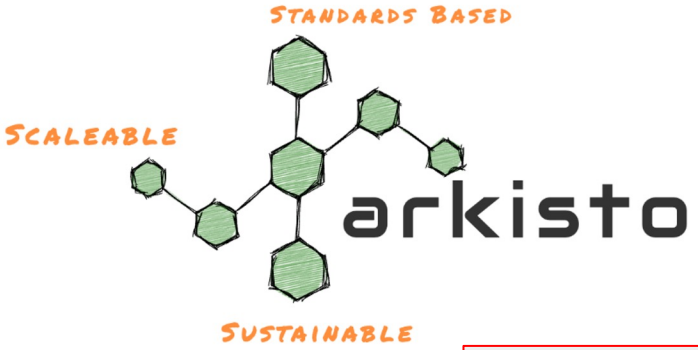
Whiteley, William Reeve
Education-Primary & Secondary I France I USyd

WHITE, EDWARD ROWDEN
Medicine, Dentistry and Health I UMelb

WHITE, NEWPORT BENJAMIN
Agricultural,Pastoral,Forestry and Fishery I UMelb

Sandford, John Lindsay
Banks, Business, Finance and Commerce I London I UAdelaide

arkisto



A scaleable, standards based platform for sustainable data.

Want to join us? Your project here? Get in touch...

The basis of Arkisto is t ed
data is

Data on an Arkisto deployment is n
independently of any services suc to
has a flexible model for how data can be accessed using a variety of services.

Arkisto is built on top of Research Object Crate (RO-Crate) and the Oxford Common File System Layout (OCFL).

With Arkisto there is no messy data migration.