

Linked Art

Session 1 0900-1030

Introduction to Linked Data and the Linked Art data model to connect cultural heritage collections data

- Intro to Linked Art and the Linked Art II project
- Linked Art code notebooks
- Linked Art questionnaire
- Questions
- Install software required to work with notebooks in **Session 2**

Session 2 - 1100-1300

Connecting cultural heritage collections data with Linked Art
– transformation, reconciliation and visualisation

- Live run through with code notebooks at front of class
- Participants work through notebooks

Discussion Groups / Questions

Linked Art

https://canvas.ox.ac.uk/courses/158781/discussion_topics/499148

Data Standards in DH - Panel Discussion

https://canvas.ox.ac.uk/courses/158781/discussion_topics/499147

Connecting Cultural Heritage Collections Using Linked Art

Transformation, Reconciliation & Visualisation
with Jupyter Notebooks

Tanya Gray

tanya.gray@humanities.ox.ac.uk



Presentation

1. Linked Art in action

- a. Artist John Ruskin
 - i. timeline
 - ii. Storymap

2. Processes to connect cultural heritage collections

- Transformation
- Reconciliation
- Visualisation

3. Linked Art notebooks

- Exemplar notebooks
- Workshop notebooks

4. Software

- JSON-LD, OpenRefine, Python, Jupyter Notebook

5. Linked Art questionnaire

6. Next steps

1. Linked Art In Action

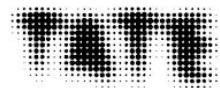


- Data visualisation
 - Collection data for artist John Ruskin from multiple sources
 - Reconciled with additional data from authoritative source using OpenRefine
 - Geocoords
 - Place name identifiers
 - Transform Linked Art JSON-LD to CSV for input to data vis
 - Visualise with Knightlab timeline and storymap

Data Visualisations



ASHMOLEAN
MUSEUM
OXFORD



Harvard
Art Museum

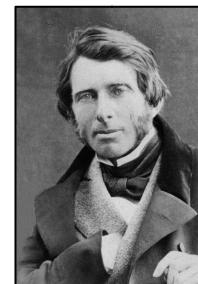
Philadelphia
Museum of
Art

RIJKSMUSEUM
amsterdam

RIJKS MUSEUM

Museum & Galleries
Collection Data

John Ruskin. artist



Unified representation with
Linked Art

Data visualisation service

StoryMap^{JS}

Maps that tell stories.



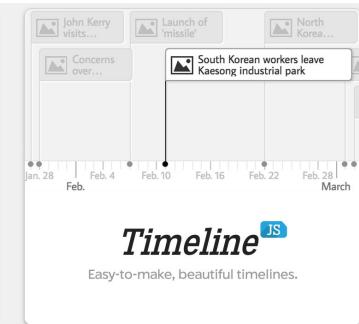
Timeline^{JS}

Easy-to-make, beautiful timelines.

Linked Art:
Sustainable Cultural Knowledge
through Linked Open Usable Data

<https://linked.art/>

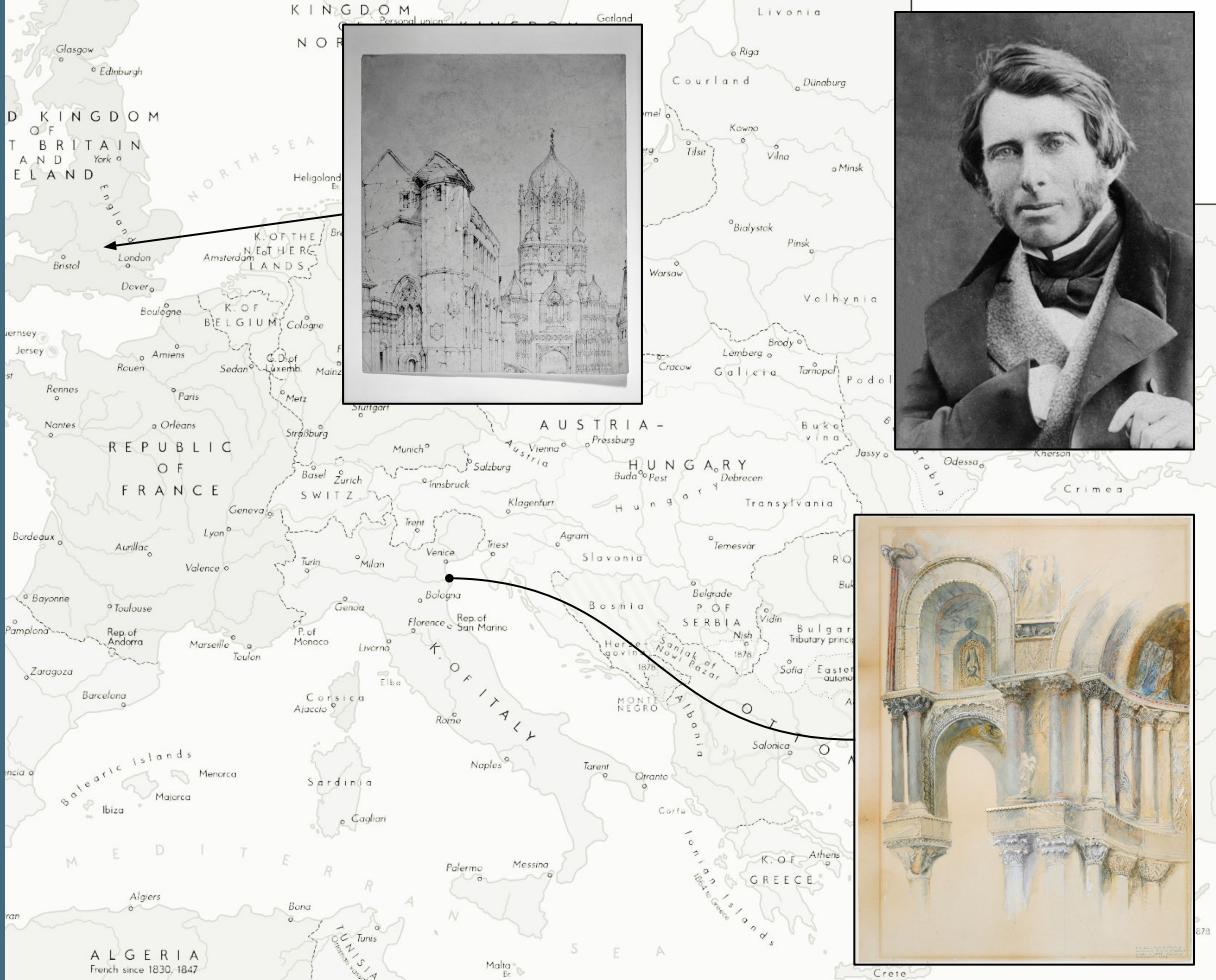
knight lab



John Ruskin

- Prolific artist
- Social commentator
- European travels
- Depicted nature and architecture
- Influential for ideas on society, art, craft, architecture, building preservation

Artworks now in many private and public collections

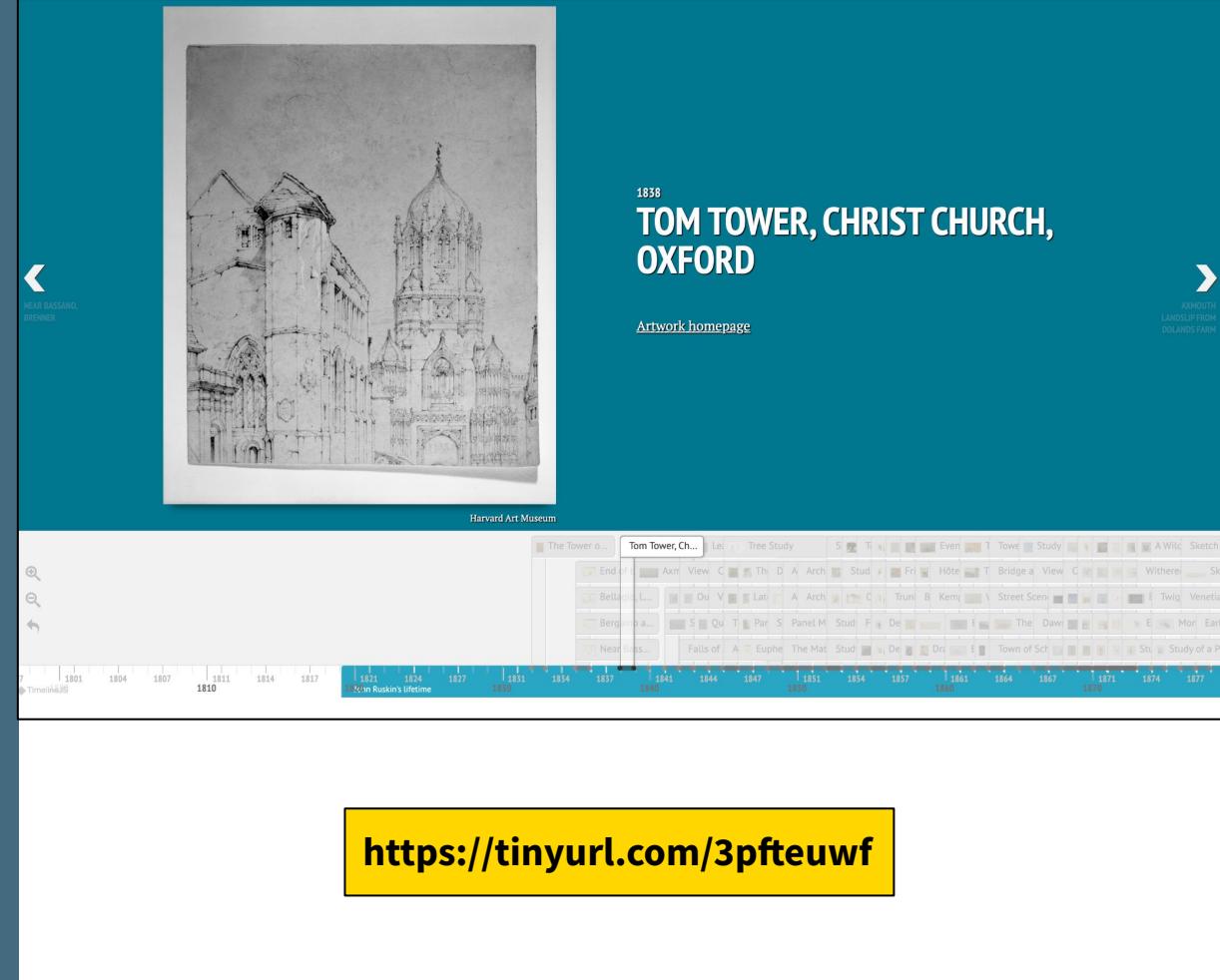


John Ruskin

Timeline visualisation

Uses

- Museum and Gallery collection data
- unified representation with Linked Art
- KnightLab Vis



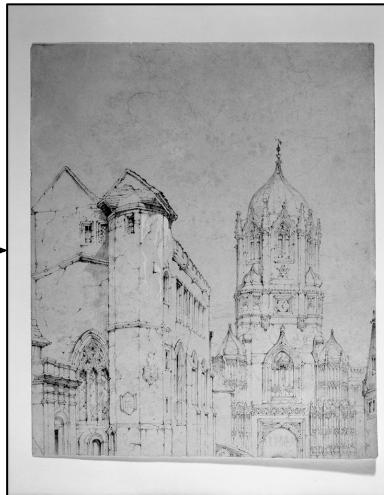
John Ruskin Timeline Visualisation

The image shows a digital exhibition interface. At the top, there is a painting titled "VIEW OF CHAMONIX" by J.M.W. Turner. Below the painting is a color calibration strip. To the left of the painting is a left arrow icon, and to the right is a right arrow icon. To the right of the painting, the title "VIEW OF CHAMONIX" is displayed in large white letters, with "J.M.W. TURNER" above it. Below the title is the caption "An early view of the valley". At the bottom of the interface is a horizontal timeline. The timeline has numerical markers at 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, and 1900. A blue bar highlights the year 1866. A tooltip or dropdown menu is visible over the timeline, containing the text "At the end of 1866, Ruskin had just completed his first major work on architecture, 'The Seven Lamps of Architecture'". On the far right of the timeline, there is a small logo for "BRITISH LIBRARY".

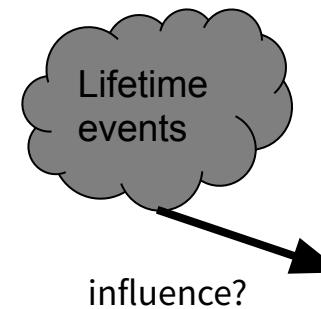
Timeline - Future Scholarship



Early influence



Early artwork



Later artwork

Samuel Prout, Artist

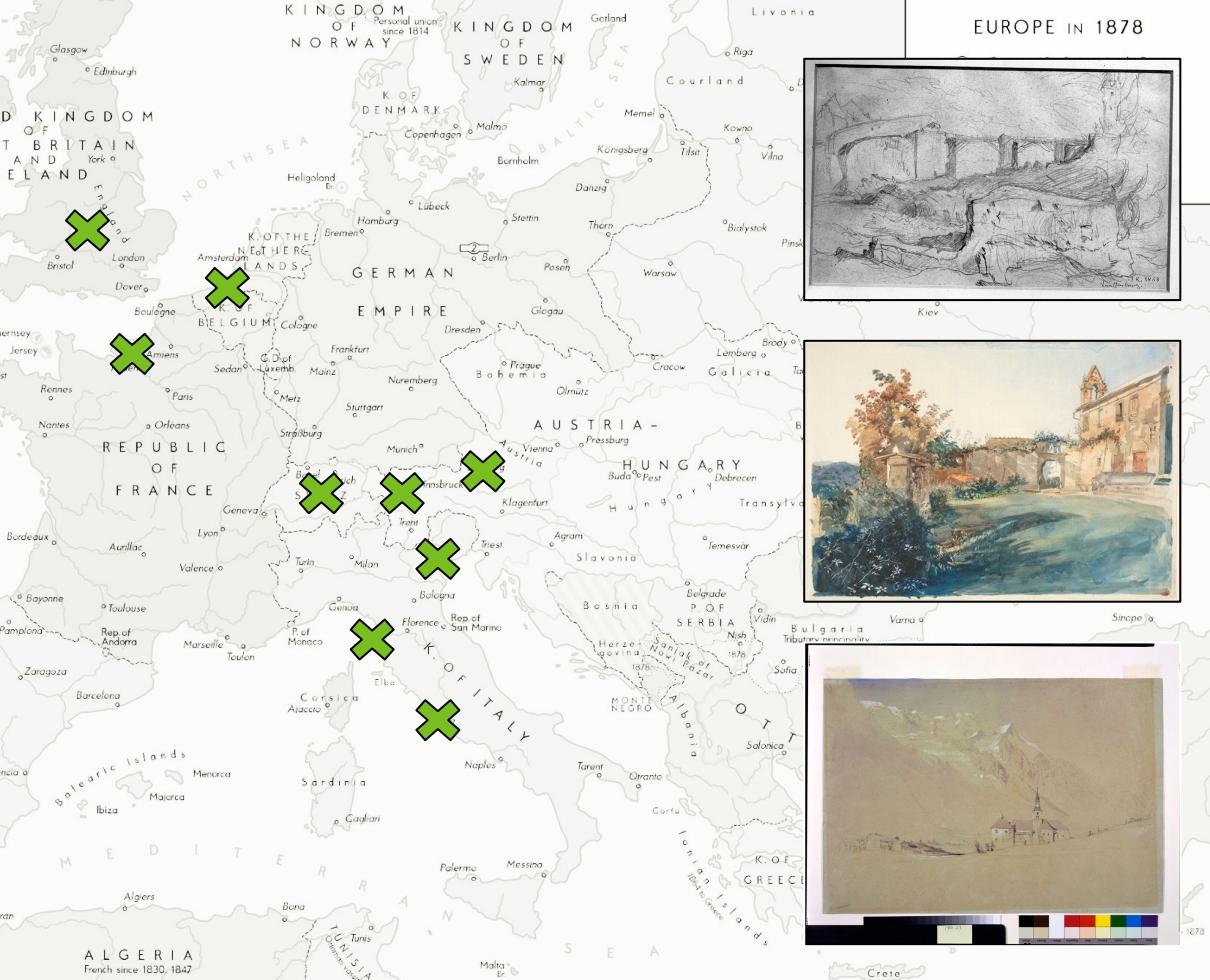
A change in characteristics of Ruskin's artwork through time?

John Ruskin

Travel

Extensive travel in Europe over his lifetime, often to Italy

Recorded travels with drawings and paintings of natural scenery and buildings



John Ruskin

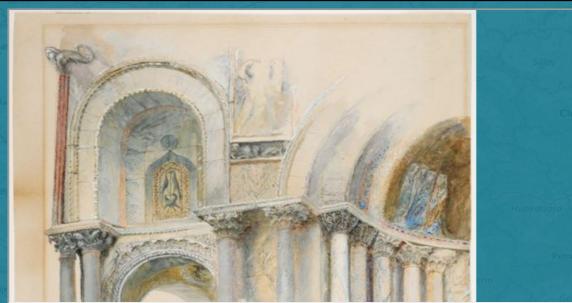
Place name in title



1855

Ashmolean Museum

BELLAGIO, LAGO DI COMO



1879

Harvard Art Museum

PART OF A SKETCH OF THE
NORTHWEST PORCH OF ST. MARK'S



October 1874

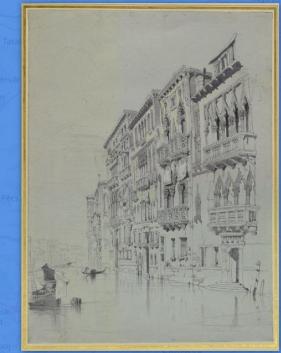
Ashmolean Museum

MONT BLANC FROM SAINT-MARTIN-SUR-ARVE



1855

Ashmolean Museum



6-16 May 1841

Ashmolean Museum

BERGAMO ROAD TO B
THE PALAZZO CONTARINI-FASAN,
VENICE

John Ruskin StoryMap

Uses:

- Collection data
- Unified with Linked Art
- Reconciled with Getty Thesaurus of Geographic Names to extract geocoordinates
- KnightLab vis

<https://tinyurl.com/mrxwv3um>

Map Overview Back To Beginning

United Kingdom Éire / Ireland Danmark Hamburg Berlin Poland Czech Republic Slovakia Hungary Croatia Italy Spain Portugal Algeria Tunisia Libya

Golfo de Gasconne / Golfo de Vizcaya

1845

Harvard Art Museum

BOAT AND SKETCHES OF TWO FIGURES, VENICE

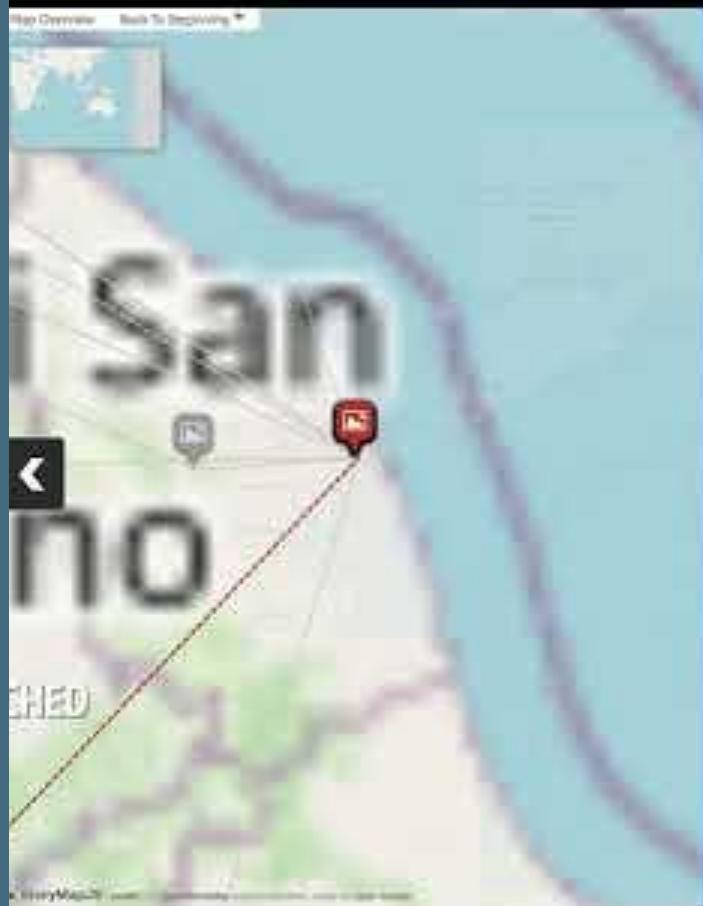
Fine Arts Department, Harvard University, Cambridge, MA, Transferred to the Fogg Art Museum, 1926.

Artwork homepage

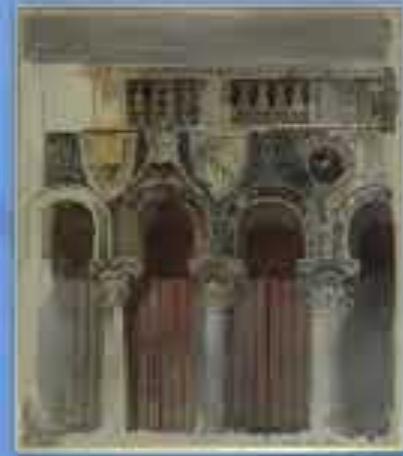
StoryMapJS | Leaflet | © OpenStreetMap and contributors, under an open license.

John Ruskin

StoryMap



STUDY OF THE MARBLE INLAYING
ON THE FRONT OF THE CASA
LOREDAN, VENICE



2. Processes to Connect Cultural Heritage Collections



- 2A. Transformation
- 2B. Reconciliation
- 2C. Visualisation

2A. Transformation



Extract data from source

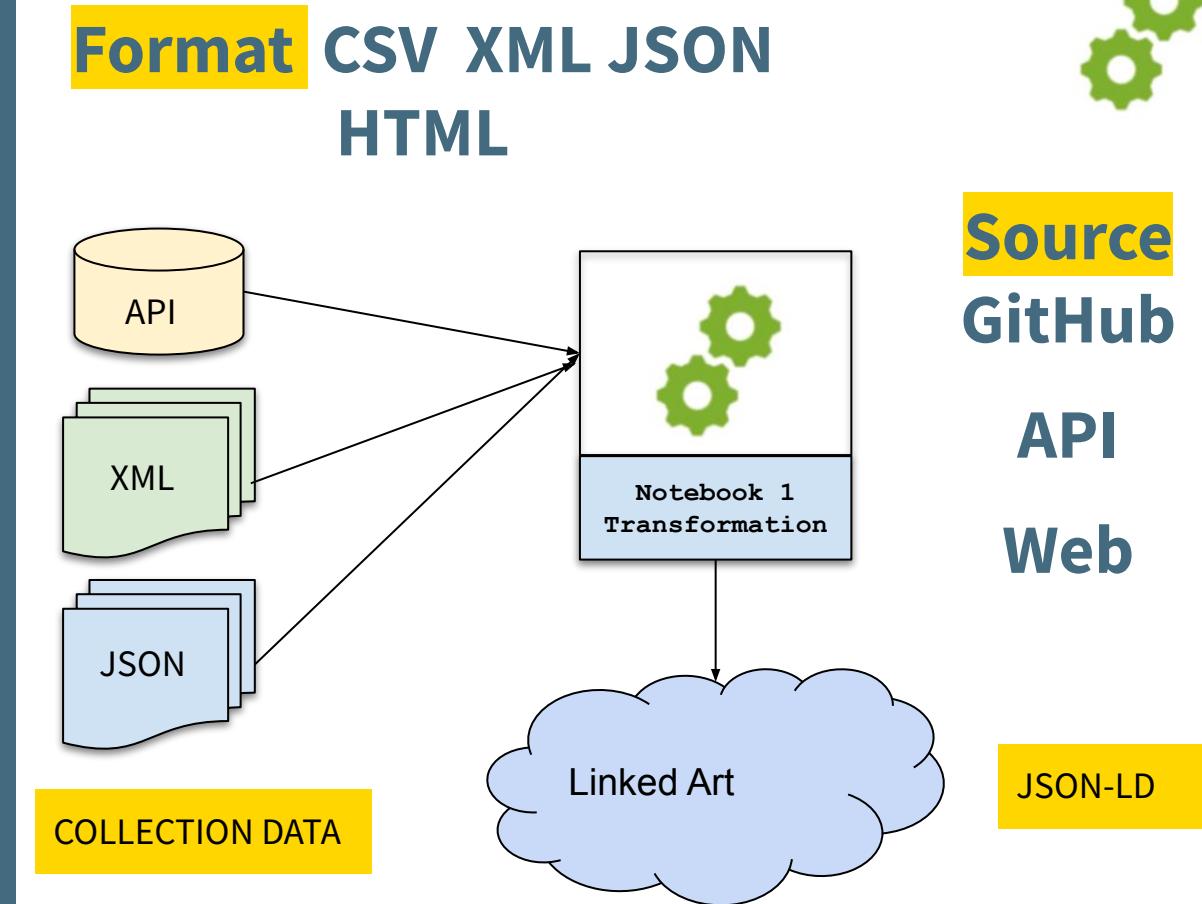
Map data to the Linked Art data model

Transform data to a unified representation using Python code

Publish data as JSON-LD files

Extract Data

- Locate
- Different formats
- Different data models
- Read data into Python dictionary



Transformation Sub-steps: Extract → Map → Transform → Publish

Map Data

Understand

- source data model
- Linked Art data model

Manual process

```
mapp = {
    "id": "id",
    "accession_number": "accession_number",
    "accession_date": "",
    "classification": "type",
    "title": "title",
    "alt_title": "title_in_original_language",
    "notes": "tombstone",
    "date_created": "creation_date",
    "date_created_earliest": "creation_date_earliest",
    "date_created_latest": "creation_date_latest",
    "created_period": "culture",
    "created_dynasty": "",
    "created_inscriptions": "inscriptions",
    "created_notes": "fun_fact",
    "creator": "creator",
    "physical_medium": "Medium",
    "physical_style": "",
    "physical_technique": "technique",
    "physical_description": "",
    "physical_dimensions": "measurements",
    "created_provenance": "provenance",
    "credit_line": "creditline",
}
```

id	id
accession_number	accession_number
accession_date	
classification	type
title	title
alt_title	title_in_original_language
notes	tombstone
date_created	creation_date
date_created_earliest	creation_date_earliest
date_created_latest	creation_date_latest
created_period	culture
created_dynasty	
created_inscriptions	inscriptions
created_notes	fun_fact
creator	NaN
physical_medium	Medium
physical_style	
physical_technique	technique
physical_description	
physical_dimensions	measurements
created_provenance	provenance
credit_line	creditline
collection	department



Transform Data

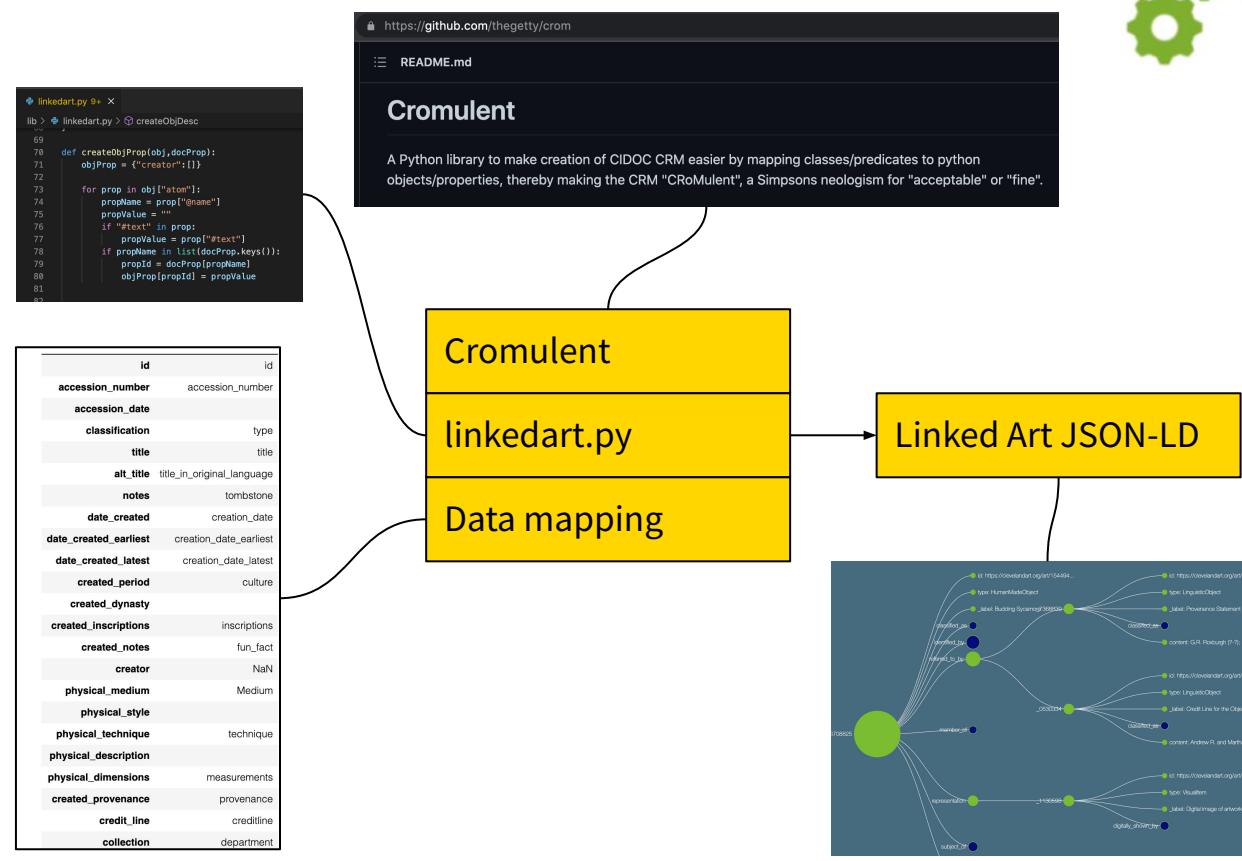


Uses

- Data mapping
- Cromulent
- Custom code linkedart.py

Creates

- Linked Art
- JSON-LD



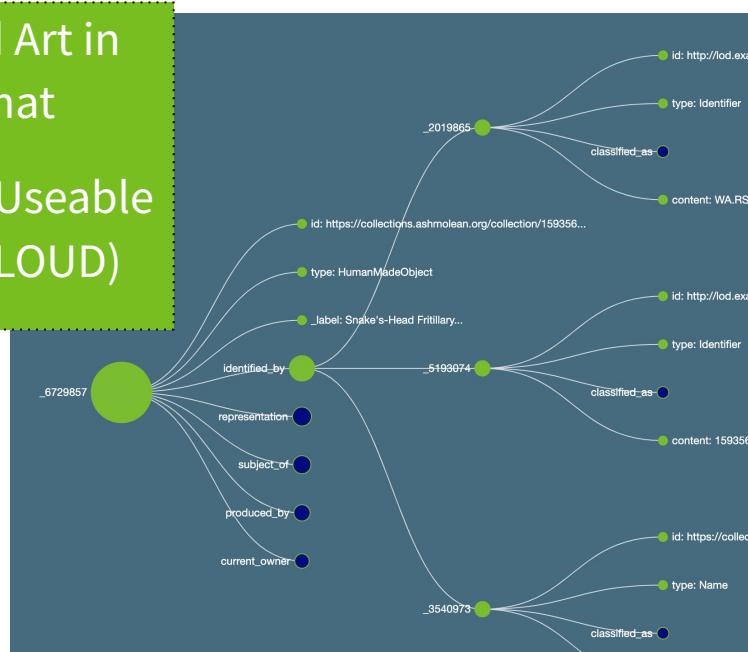
Transformation Sub-steps: Extract → Map → Transform → Publish

Transformation - Publish



Create Linked Art in
JSON-LD format

Linked Open Useable
Data format (LOUD)



```

print(factory.toString(objLA, compact=False))

{
  "@context": "https://linked.art/ns/v1/linked-art.json",
  "id": "https://clevelandart.org/art/74540",
  "type": "HumanMadeObject",
  "label": "Leda and the Swan",
  "classified_as": [
    {
      "id": "http://vocab.getty.edu/aat/300033973",
      "type": "Type",
      "_label": "drawing",
      "classified_as": [
        {
          "id": "http://vocab.getty.edu/aat/300435443",
          "type": "Type",
          "_label": "type of Work"
        }
      ]
    }
  ],
  "identified_by": [
    {
      "id": "http://lod.example.org/museum/Identifier/2015.451",
      "type": "Identifier",
      "classified_as": [
        {
          "id": "http://vocab.getty.edu/aat/300312355",
          "type": "Type",
          "_label": "Accession Number"
        }
      ],
      "content": "2015.451"
    },
    {
      "id": "http://lod.example.org/museum/Identifier/74540",
      "type": "Identifier",
      "classified_as": [
        ...
      ]
    }
  ]
}
  
```

2B. Reconciliation



Enrich data with external data source

- **Identify** place name
 - **Reconcile** place name with name authority
 - Getty Thesaurus of Geographical Names Online (**TGN**)
 - **Extract** authoritative global identifier for place name and geographical coordinates
 - **Add** new data into Linked Art data files
-



Transformation

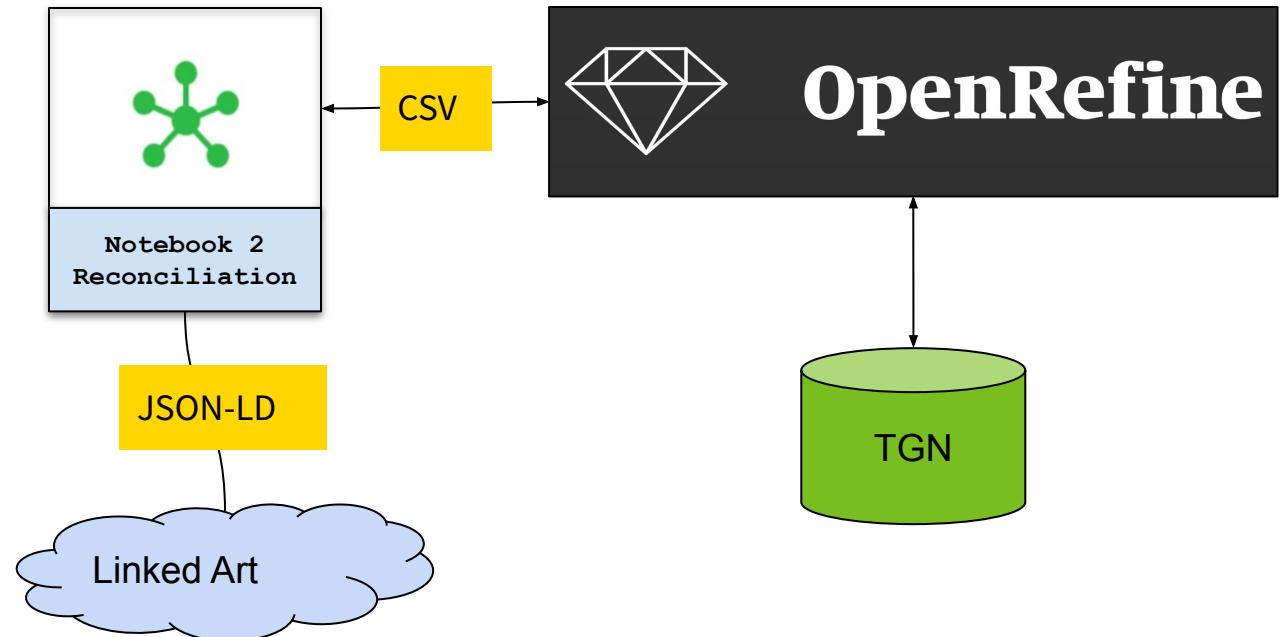


Reconciliation



Visualisation

Reconciliation Code Notebook



Reconciliation Sub-steps: Identify → Reconcile → Extract → Add

2C. Visualisation



1. Transform unified representation of collection data in Linked Art
JSON-LD to CSV
2. Input CSV into KnightLab data visualisation
3. Create visualisation
 - a. Timeline
 - b. Storymap

Advantages with Linked Art

- Unified representation
 - **Reconciled data** from separate sources
 - **Reusable visualisation process** with other Linked Art JSON-LD

Visualisation

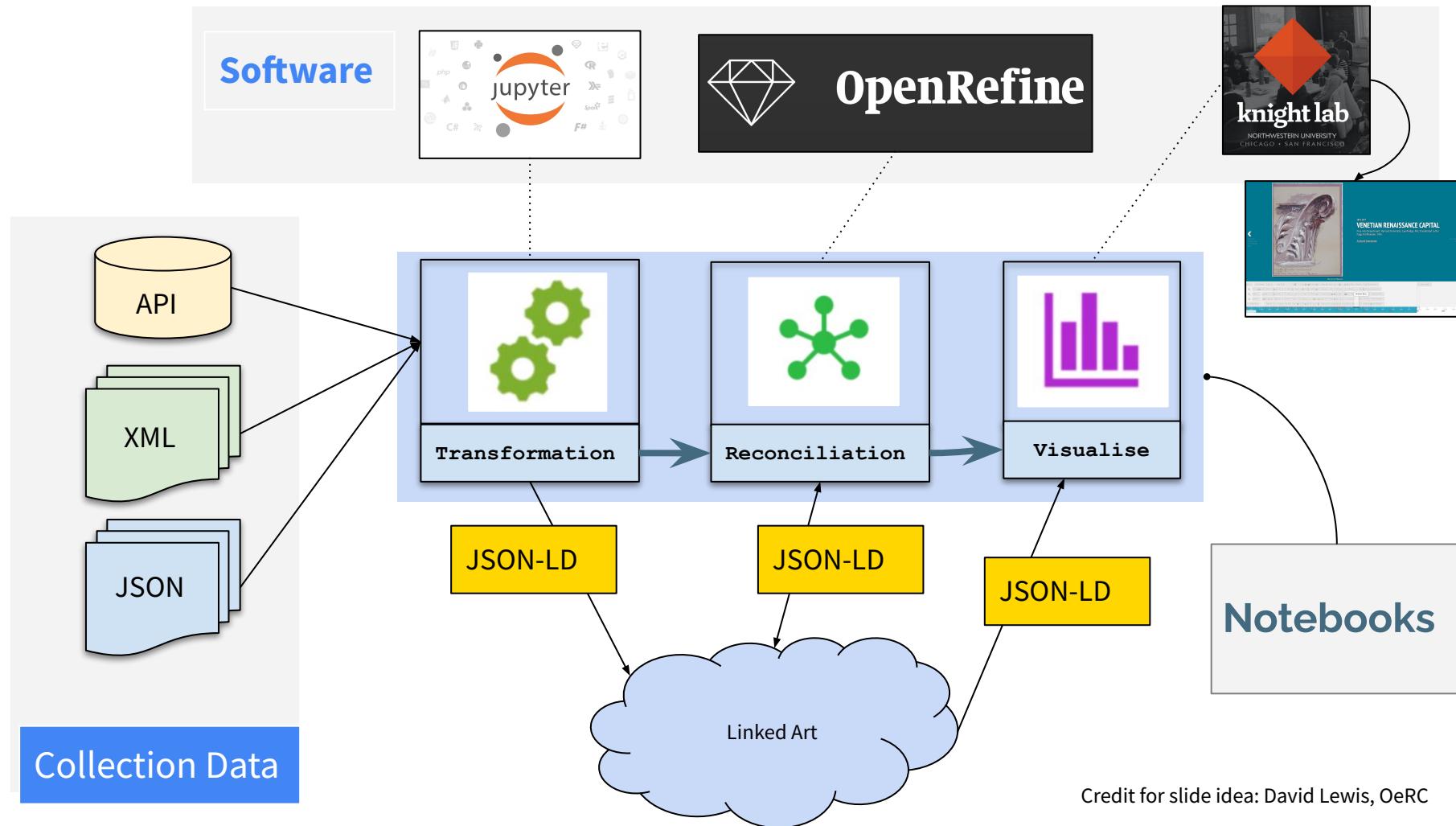
Uses

- KnightLab visualisation
- Linked Art JSON-LD
- Script to transform to CSV/JSON

Creates

- Timeline
- storymap

The screenshot displays two web-based visualization tools developed by Knight Lab. On the left, the **StoryMap JS** interface is shown. It features a map of Eastern Europe with a red marker indicating a location in Donets'k. A dashed red line connects this location to another red marker on a map of the Donbas region, specifically Vuhovakha. The map also shows several other locations with grey markers: Kurakhovo, sinovataya, Nizhnyaya Krynka, Khartsyzsk, Ilovays'k, Mospyne, Norvy Svit, Starobesheve, and Dokuchayevsk. An inset map in the top-left corner shows the location of the conflict area within Europe. Below the map, the text "StoryMap JS" is displayed with the subtitle "Maps that tell stories." On the right, the **Timeline JS** interface is shown. It consists of two panels. The left panel is a horizontal timeline from January 28 to March, marked with vertical grid lines every two days. It contains three event cards: "John Kerry visits..." (with a photo icon), "Concerns over..." (with a photo icon), and "South Korean workers leave Kaesong industrial park" (with a photo icon). The right panel is a larger area where the timeline continues, showing a single event card for "Launch of 'missile'" on February 10, with a photo icon.



3. Linked Art Notebooks

Exemplar notebooks

<https://github.com/tgra/Linked-Art>

Workshop notebooks

[https://github.com/tgra/Linked-Art/
tree/main/workshop](https://github.com/tgra/Linked-Art/tree/main/workshop)

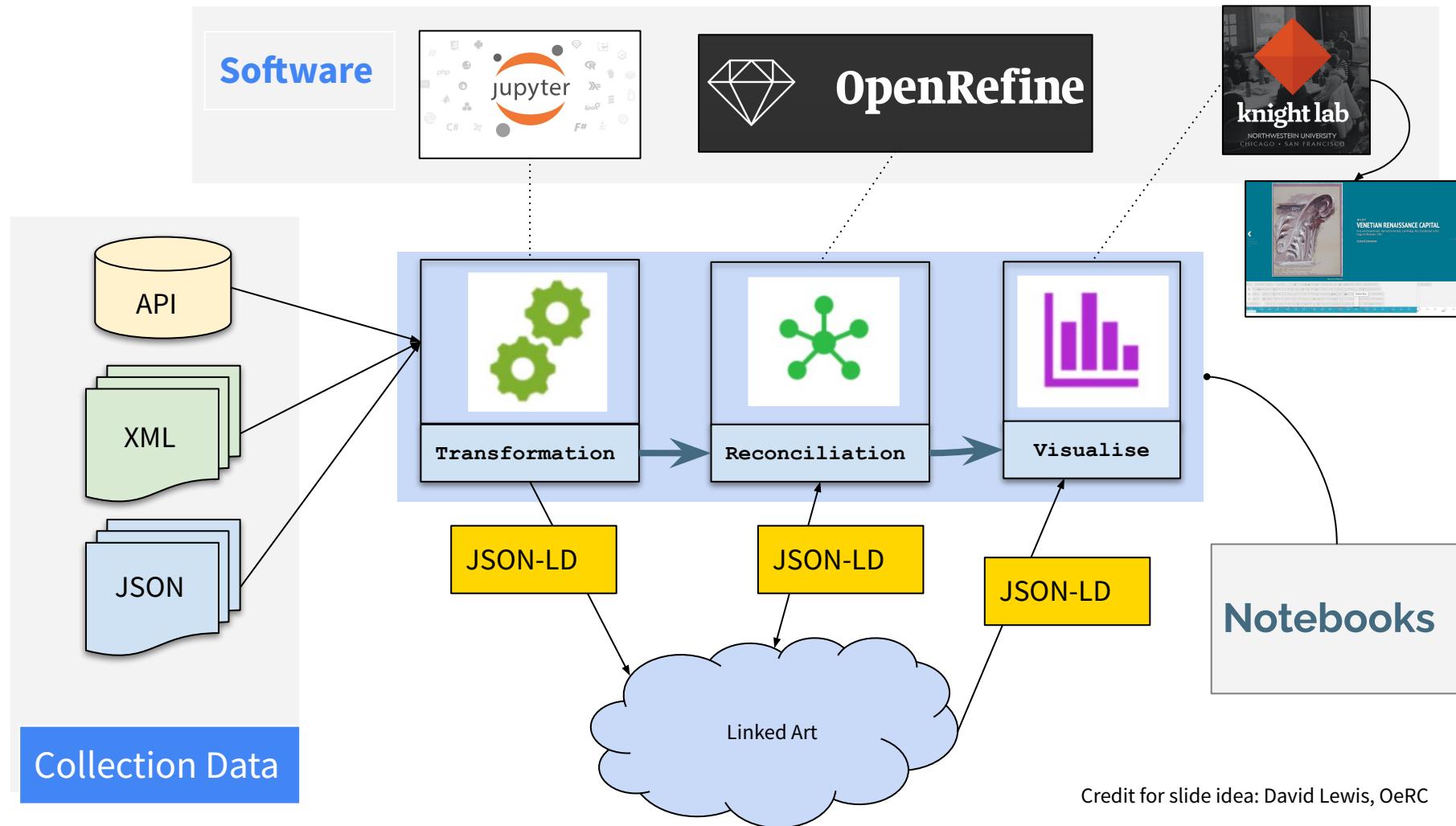
3. Linked Art Notebooks

Workshop notebooks

- Developed for DHOxSS workshop, further to feedback from LAII Questionnaire
- Break down transformation process into simpler steps
- Allow you to **try process with your own data** by uploading file and saving created files in a working directory
- Exemplify
 - Transformation - Extract, Map, Transform, Publish
 - Reconciliation
 - Visualisation

Exemplar notebooks

- Developed as part of Linked Art II project
- Exemplify
 - **transformation processes using different collection data**
 - **reconciliation** using Open Refine
 - **visualisation** using Linked Art and KnightLab visualisation software

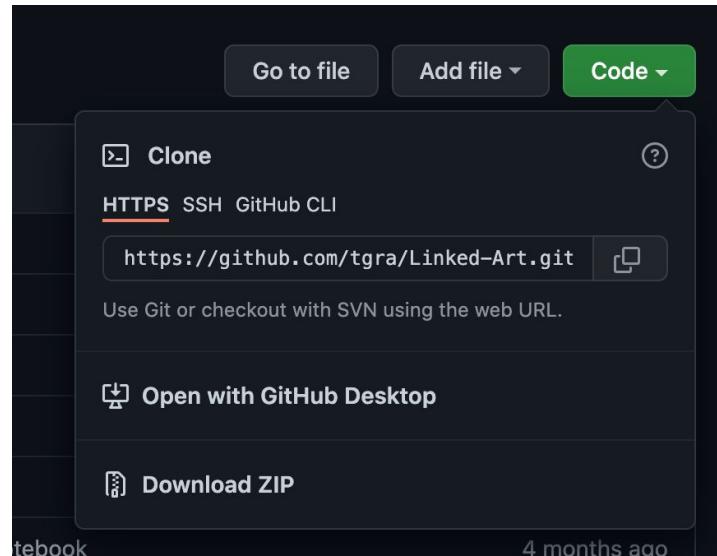


Code Notebooks

GitHub
github.com/tgra/Linked-Art

Install Git

- Check out with ‘git clone’
- Download ZIP



Notebook Tools

- Jupyter command line
- Binder
- Anaconda
- Jupyter Nbviewer
- Visual Studio code
- JupyterHub



Notebook Tools

Jupyter
Notebook via
command line



The screenshot shows a terminal window with several tabs open. The active tab displays the output of a command related to Jupyter Notebook configuration and startup. The text includes:

```
To modify your account to use ssh, please run 'ssh-keygen -t rsa' and then
cat your private keyfile (without the passphrase) into /etc/jupyter/jupyter_notebook_config.py
then restart jupyter notebook or jupyter lab with 'jupyter notebook --allow-root' or 'jupyter lab --allow-root'.
For example while loading config file /usr/share/jupyter/notebook_config.py
File "/root/.jupyter/jupyter_notebook_config.py", line 52, in <module>
    exec(open('jupyter_notebook_config.py').read(), dict(__file__=__file__))
  File "jupyter_notebook_config.py", line 1, in <module>
    from jupyter_core.paths import JupyterConfig, JupyterRuntime
ImportError: cannot import name 'JupyterRuntime'

File "/root/.jupyter/jupyter_lab_config.py", line 52, in <module>
    exec(open('jupyter_lab_config.py').read(), dict(__file__=__file__))
  File "jupyter_lab_config.py", line 1, in <module>
    from jupyter_core.paths import JupyterConfig, JupyterRuntime
ImportError: cannot import name 'JupyterRuntime'

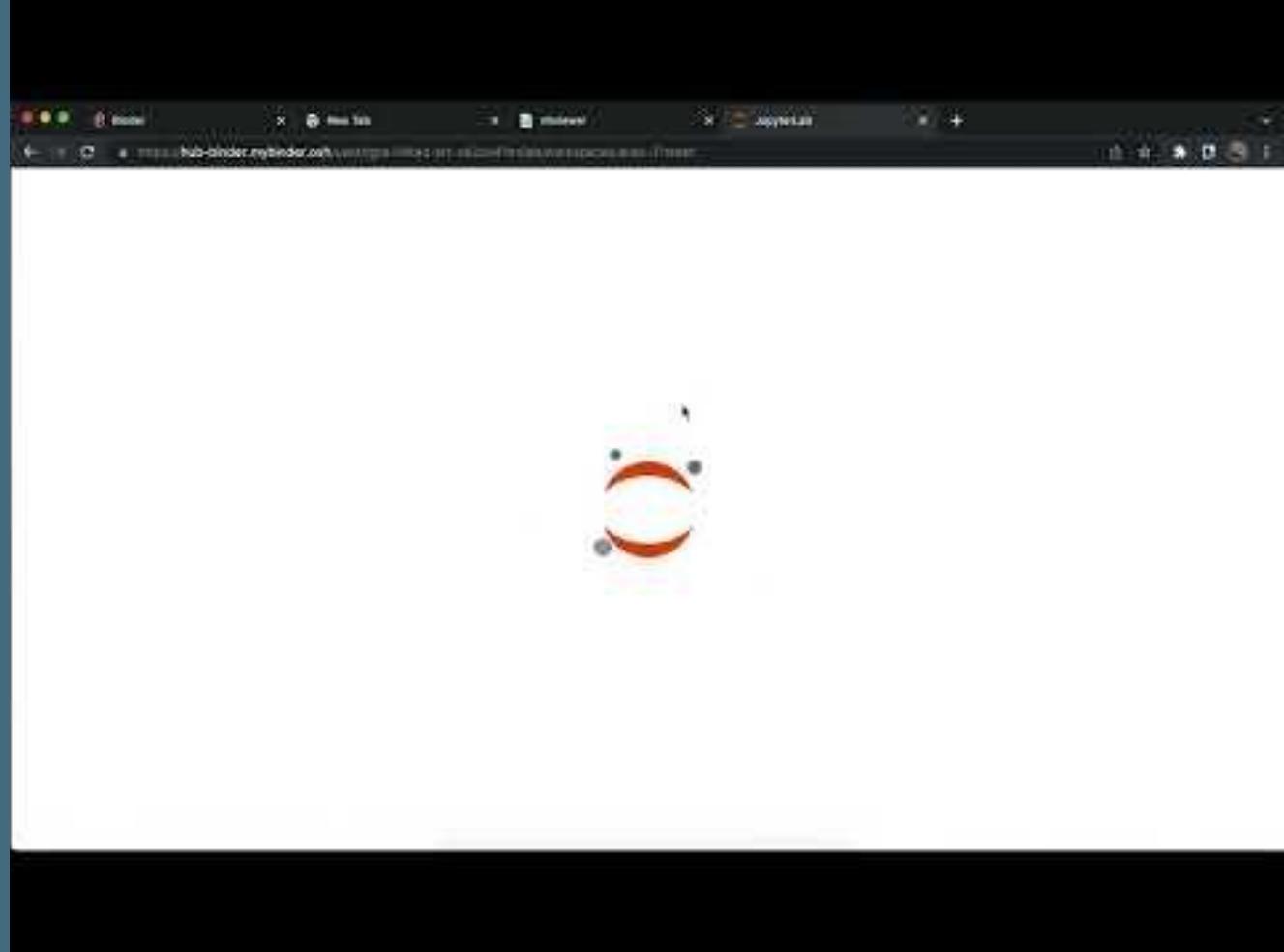
File "/root/.jupyter/jupyter_notebook_config.py", line 52, in <module>
    exec(open('jupyter_notebook_config.py').read(), dict(__file__=__file__))
  File "jupyter_notebook_config.py", line 1, in <module>
    from jupyter_core.paths import JupyterConfig, JupyterRuntime
ImportError: cannot import name 'JupyterRuntime'

[Terminal Error] uncaught exception
2020-05-05 16:11:57,931 [INFO]     Jupyter extension loaded from /opt/miniconda3/lib/python3.8/site-packages/jupyterlab
2020-05-05 16:11:57,931 [INFO]     Jupyter application extensions in /opt/miniconda3/lib/python3.8/site-packages/jupyterlab
2020-05-05 16:11:57,931 [INFO]     Loading notebook from local directory /Users/Hariya/Downloads/Code/JupyterLab/2020-05-05-16-11-57
2020-05-05 16:11:57,932 [INFO]     Jupyter Notebook 6.3.0 is running at
2020-05-05 16:11:57,932 [INFO]     http://127.0.0.1:8888/?token=f7d9880f03f77c139e70d3e9e0f05a3011ff
2020-05-05 16:11:57,932 [INFO]     or http://[::1]:8888/?token=f7d9880f03f77c139e70d3e9e0f05a3011ff
2020-05-05 16:11:57,932 [INFO]     Use Ctrl+C to stop this server and start again with Jupyter Notebook to view content again
```

Below the terminal window, there is some UI for a browser window, showing a URL for a Jupyter Notebook page.

Notebook Tools

Binder via
GitHub



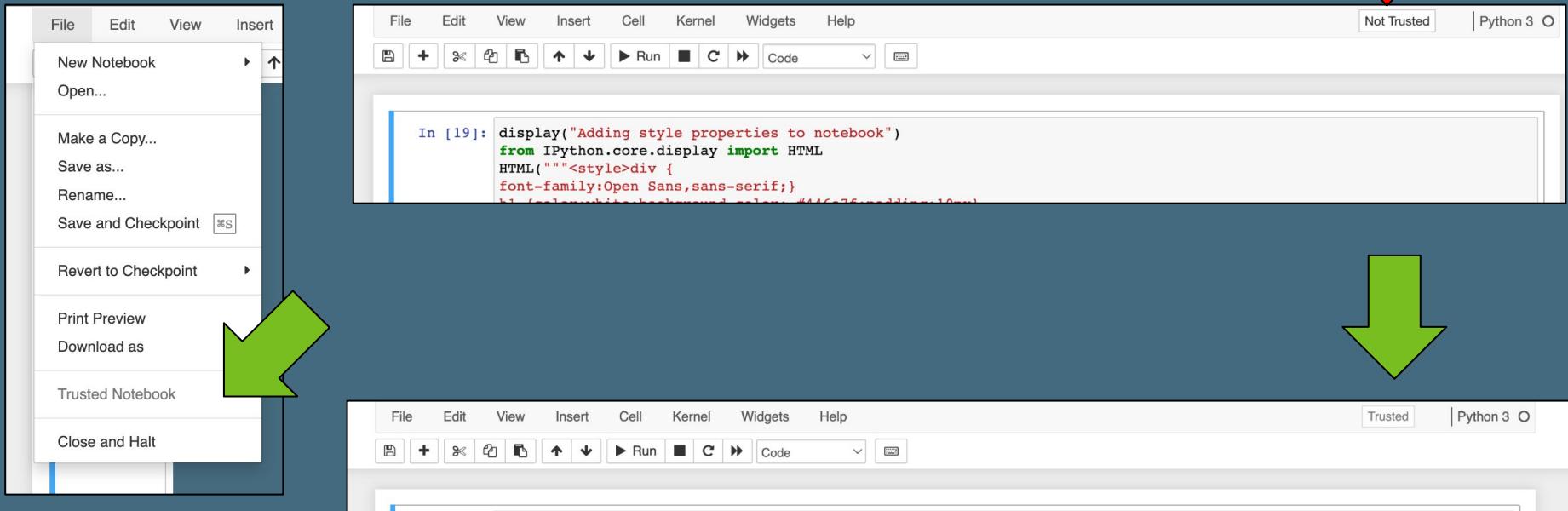
Notebook Tools

Nbviewer

nbviewer.org



Jupyter Notebooks - Trusted vs Not Trusted



How do I trust a notebook?

Users can explicitly trust a notebook in two ways: 1. **At the command-line**, with: `jupyter trust /path/to/notebook.ipynb`. 2. **After loading** the untrusted notebook, with File / Trust Notebook.

How to Interact with the Notebooks

The notebooks can be approached in various ways

- Read-only
 - Go through example provided
 - focus on the notes to explain the process exemplified
 - Interactive
 - Go through example provided
 - Edit the example code and/or
 - Go to latter part of notebook where you can upload your own data files
-

4. Software

Introducing

- JSON-LD
 - OpenRefine
 - Python
 - Jupyter notebooks
-

1A. JSON-LD

- **Extension to JSON**
 - a human and machine readable data format
 - Simple nested key:value data representation
 - Popular with web developers so more tools and libraries
- **A data format** to embed linked data into JSON
- **Easy to use** - allows for concise JSON with short property names
- **Provides context** to terms and objects
 - **@context** URI that gives information on meaning of terms and constraints on data values
 - **@id** resolvable global URIs as identifiers for entities/objects
 - **Removes ambiguity**
- **Transformable to other Linked Data Serializations** e.g. RDF/XML
- **Linked Open Usable Data**
- Used for **Linked Art Data Model**

Further Information

JSON-LD Homepage

- <https://json-ld.org>

[Manu Sporny Youtube](#)

Linked Art JSON-LD

<https://linked.art/loud/>

1B. OpenRefine

“A free, open source, powerful tool for working with messy data”

- Allows you to query services to match and retrieve authoritative identifiers and associated information, e.g.
- identifiers for
 - terms, people, places, entities, events
- associated information such as
 - geo-coords

Example services

- Wikidata,
- Getty Thesaurus of Geographic Names® Online (TGN)



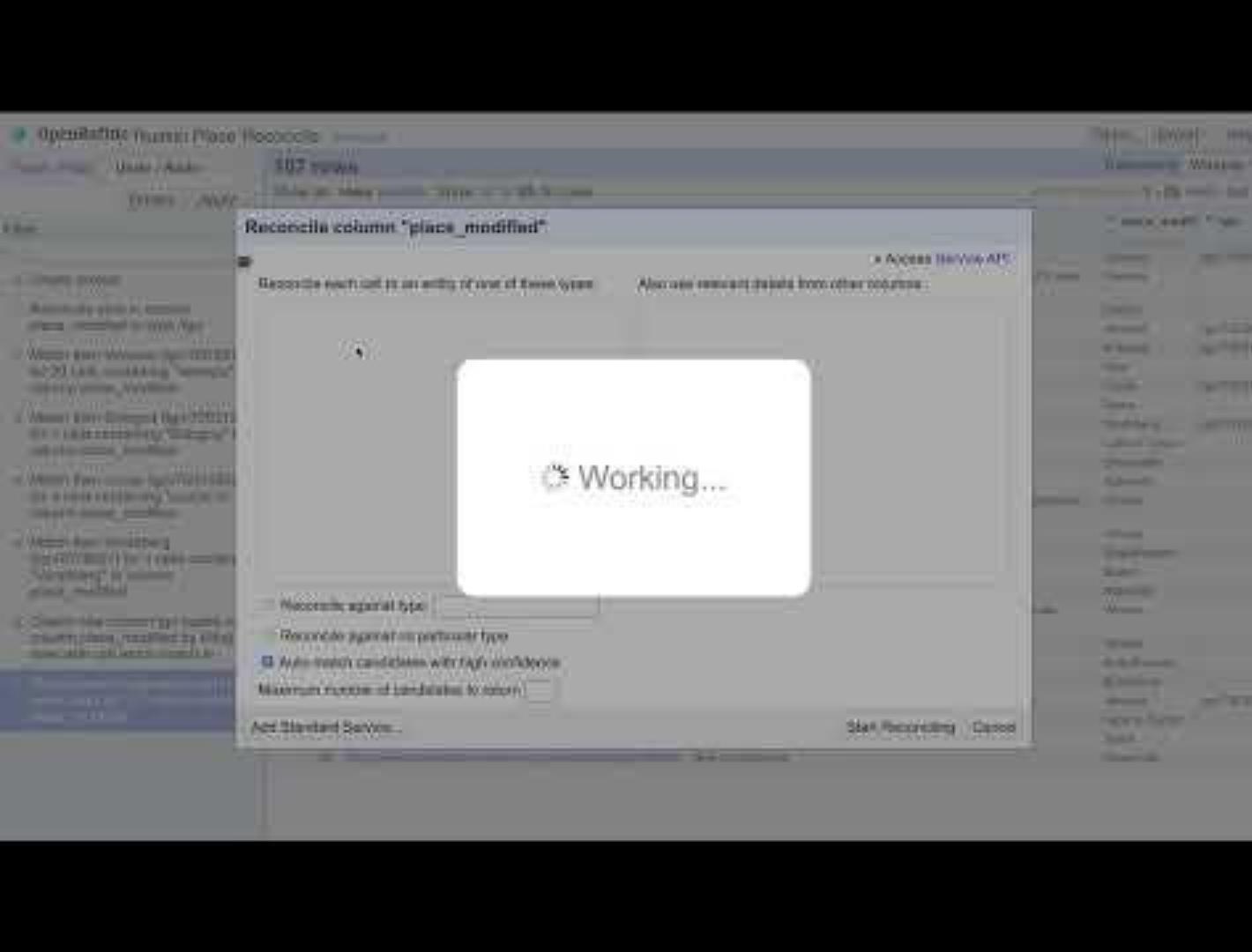
<https://openrefine.org/>

Videos:

- Shows selection of service to query
- Review of reconciliation results

OpenRefine

Select Service to Query



OpenRefine

Review Results

OpenRefine Ruskin Place Reconcile - Remote

Match Item Bologna (sgn/7303127) for 1 cells
containing "Bologna" in column place_modified
Lindo

Open... Export... Help

Extensions Wikidata

107 rows

Show as: rows records Show: 5 10 25 50 rows

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107

place_modified: judgment

2 cells. Sort by: name count

PRIMORDIAL (1)

None (0)

Post by name count

place_modified: best candidate's score

5.00 — 30.00

1 https://www.harvardartmuseum.org/objects/sgn/7303127# Linding down from Pomeria towards Lucca

Match the Cell - Match All Identical Cells

Cancel

inhabited places

Location in Hierarchy:

- World
- Italy
- Sicily
- Agrigento
- Lucca, Italy

2 https://collections.usm.msu.edu/collections/15487 Study of a Lion

Match the Cell - Match All Identical Cells

Cancel

inhabited places

Location in Hierarchy:

- World
- Europe
- Italy
- Lucca
- Lucca, Italy

3 https://www.harvardartmuseum.org/objects/sgn/1298624 Evidence to Faliero, the Voluntary

Match the Cell - Match All Identical Cells

Cancel

inhabited places

Location in Hierarchy:

- World
- Europe
- Italy
- Lucca
- Lucca, Italy

4 https://collections.usm.msu.edu/collections/15488 End of the Lake of Lucca

Match the Cell - Match All Identical Cells

Cancel

inhabited places

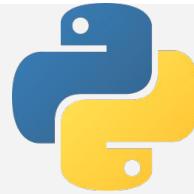
Location in Hierarchy:

- World
- Europe
- Italy
- Lucca
- Lucca, Italy

1C. Python

Python is

- a high-level
- interpreted
- general-purpose
- programming language



Its design philosophy

- emphasizes
 - code readability
 - with the use of
 - significant
 - indentation!

Further Information

Official website

<https://www.python.org/>

1D. Jupyter Notebooks

Code notebooks are a type of tool that provide a

- **web interface to a combination of executable code and associated narrative text**
- can be used for the provision of tools for
 - learning about coded processes
 - other applications e.g. arranging workflows in data science.

Characteristics

- **Free availability**
- **support for Python coding**
- **Ecosystem of supporting tools** and documentation as well as existing notebooks



Further Information

Homepage

<https://jupyter.org/>

Documentation

<https://jupyter-notebook.readthedocs.io/en/latest/notebook.html>

Jupyter Notebook Cheat Sheets

- https://www.edureka.co/blog/wp-content/uploads/2018/10/Jupyter_Notebook_CheatSheet_EduReka.pdf

4. Linked Art II Questionnaire

Published 11 May - 14 June 2022

Questionnaire invited comments on the
Linked Art and the Linked Art notebooks

Review of Linked Art data model
feedback presented here

Respondents

19 respondents

- professional role or interest in collecting institution e.g. museum, library, gallery
 - **14 respondents**
- professional role or interest in academic research using collections
 - **7 respondents**



- Were you aware of Linked Art prior to this questionnaire?



Prefer not to say | 0

Do you think adding links between cultural heritage collections data from different organisations would improve the quality of the collection data?



- data reconciliation
- improve the data
- more complete or accurate data
- spread accurate or authoritative data
- improve consistency
- make new connections (useful for research)
- improve research
- highlight errors, inconsistencies and out-of-date facts
- Help with transmission of ideas

- It depends
 - what links, whom links are vetted by, for what purpose links are made
- only [with] links to other high quality sources of data
- wouldn't improve data automatically
- will not change the data that already exists
- Possibly add detail
- contingent upon the usefulness of the data

- possibility to reduce the quality of the resource

Do you think adding links between cultural heritage collections data from different organisations would improve the quality of the collection data?



From those answering “no” to the question, a single respondent commented, saying:

It could help people to improve the data (for example using related records to make it more complete or accurate) but it will not change the data that already exists.

Do you think adding links between data from different organisations would improve the utility of the collection data?



- improve quality
- use case of combining data .. different sources .. far more prevalent than [...] single source
- more homogenization of data between collections, collections information would become more trustworthy
- can vastly increase accessibility
 - more pathways and entry points for researchers to follow

- From definitions provided
 - would only improve quality

- raises more problems than solutions

Collections need to be reconstructed in earlier forms, which by definition crosses modern institutions

Do you think adding links between data from different organisations would improve the utility of the collection data?



The person answering “no” to the question commented, saying:

It raises more problems than solutions. Among the definitions proposed, it could only improve quality for real.

Can you identify any challenges when combining cultural heritage collections data?



- Inconsistencies
 - different perspectives
- Vocabulary alignment,
 - if different collections use different terminologies
- Lack of clear licensing
- resources and staff training

- matching and integrating data
 - about same entities
- Inconsistent data fields and sometimes different data formats
 - e.g. can download CSV from place 1
 - place 2 uses API serving JSON, etc.

Can you identify any challenges when combining cultural heritage collections data?

- Reconciling data

- different cataloguing standards;
- agreeing on common authorities for related items (people, places, etc.)
- Data mapping issues.
- contradicting information
- different levels of precision
- Ensuring that linked resources from other data sets are referring to
 - the same concept,
 - the same actor,
 - the same work of art, etc.

- Publishing open data

- Siloed data is created with systems workarounds - need to be fixed before publishing as open linked data

Resourcing

- Standardization can be a long process
 - time, staffing and money.
- not seen as a priority
- little funding is given to it

Data quality concerns

- Reduction of differences, nuance and subjectivity through connections.
- it is important to ensure the quality and certainty of any links being established
- inaccurate links
 - could cloud the meaning of the data and reduce the accuracy and reliability of the data.
- there is no currently standardised method to convey the level of certainty of a link embedded within JSON-LD

Can you identify any beneficial opportunities to combine collections data that you use, with collections data from other institutions?



- **Enrich user experience**
 - links to comparable objects
 - richer authority data drawn from other resources
 - place our objects in broader contexts
 - e.g. other objects from same previous collection
- **Find works**
 - easier to find copies of attested works in digitized form

- **Reconstruct**
 - earlier collections
 - fragments
- **Check data**
 - dates and rights information
 - get overview of contradictory statements

I generally don't use collections data itself.

Is there someone in your organisation, or with whom you collaborate, who has the necessary technical skills to write a script to transform collections data to Linked Art?



- I develop the templates
 - With Software Engineer.. look to automate the creation of JSON-LD
- pass this .. to our middleware vendor
- generally do much of that work myself
 - work with others with a deep enough understanding of the subject and practice
 - can seek advice

- Yes, but it's not part of their job and so it might be difficult to get them to do it.
- I could do but only after a lot of research and study
- number of us who are capable of doing this
 - but it isn't our full time work
 - we don't do it daily
 - there would be a learning curve

Do you think the reconciliation of collections data from different institutions using the Linked Art data model would improve the accuracy of cultural heritage collections data?



- Linking to external authorities will help flush out inconsistencies and omissions in cataloguing
- reconciling free-text to controlled vocabularies
 - brings significant additional value to the data
 - makes it possible to do things reliably and consistently with the data that would not be possible otherwise

- yes and no--depends on how clean or messy the data you have and that you use from other institutions is

- *important that any reconciliation is done with the input of content knowledge experts*
 - to ensure the accuracy of any assertions
- *ideally the source data that led to the reconciliation would be offered within the LOD record*
 - end-users could see the source of the reconciliation

Do you think there would be value in adding globally unique identifiers to cultural heritage collections data?

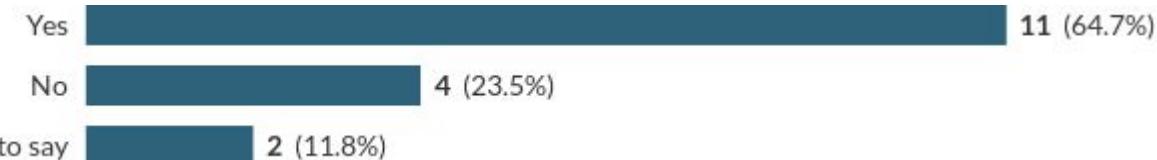


- to distinguish among copies held at different institutions
- for external authority data
- reduce ambiguity if they can be defined at the individual object level
- easy way to connect data together
 - especially in machine readable ways

- resources should be consistently available from a stable URL
 - ideally with a commitment to ensure that such data was available for the long-term.

- *The problem is*
 - which external IDs to use – Wikidata?

If applicable, does your organisation's cultural heritage collections data contain globally unique identifiers?



- in the process of implementing ARKs
- Only a few, for people / organisations; but we plan to add more.
- Library authority files: NACO/Library of Congress, ISNI, etc.
- persistent URI with purl

- we tried, but we cannot continue them because of a change of collections management system

Do you think that using Linked Art to reconcile data from external data providers, and enable a review of artwork descriptions in collections data, would help to improve the quality of the collections data?

Yes  18 (94.7%)

No | 0

Prefer not to say  1 (5.3%)

- will require us to be more precise about the entities we are linking to, e.g. reconcile entities with shared names
- could identify
 - gaps in metadata,
 - lack of standardisation
- very act of attempting reconciliation will create a focus on the underlying data

- Ascertaining the quality of collections data is an interpretive act.
- great deal of further consideration and theoretical understanding is required here.
- concerned whether inherent biases in description (with respect to culture, gender, etc.) could play a negative role here without careful consideration.

Would the application of Linked Art to reconcile data from external data providers to your organisation's collection data, to review, correct or augment the artwork descriptions, help to improve the utility of your collections data, either publicly or privately?



- finding related artworks is always nice for users
- more consistent categorisation and search
 - e.g. facets for places, people, subjects
- Ability to pull in extra details from external authorities
 - e.g. birth and death dates for linked people

- It could, depending on
 - what content was available externally
 - how it had been mapped

- we would like to incorporate getty and wikidata to do this in our data
 - haven't got to the point where we can do it in a way that isn't just manual

Would the reconciliation of collections data from different institutions using the Linked Art data model allow collections data that you use to be augmented, e.g. with the addition of data for empty fields?



- Add authority file ids to the records
- It could
 - depend.. on what content was available externally regarding the work of art and how it had been mapped
- not automatically
 - the curators would want to review everything
- [need to] assure some staff at the museum that the data is accurate or at least comes from a reputable source
 - we run against this in our efforts to include more wikidata

What different sorts of data visualisation could take advantage of collections data connected using Linked Art?

- Geographical and chronological distribution of artwork properties e.g. materials
- Mapping of materials used
 - Show flow of trade based on location of origin
- Artist/creator
 - size of an artist's oeuvre
 - Life event dates
- What questions are researchers asking?
- Experiments with computer vision and image analysis of related images
- Make connections between datasets across institutions
- Link complementary data sets
- Artwork
 - Place of production
 - Provenance
 - Location of depicted people or scenes
- Dates of creators
- Metadata inconsistencies
 - competing or contradictory metadata
- Collections
 - Map where exhibited over time
 - What galleries
 - Next to what other works of art
- Connections between artists and works globally

Acknowledgements



Arts and
Humanities
Research Council

This work was undertaken by the [Linked Art II project](#) at the University of Oxford (Principal Investigator: Dr. Kevin Page, Oxford e-Research Centre) funded by the [UK Arts and Humanities Research Council](#) (AHRC project reference AH/T013117/1). The project's Research Software Engineer was Tanya Gray.

We gratefully acknowledge the participation and contributions of our **project partners** and the wider **Linked Art community**.



Thank you to the museums and galleries that are making their collection data available for re-use, via APIs and data downloads



5. Next Steps

Questions

Linked Art discussion group
<https://tinyurl.com/3rbcrb4v>

Setup for Hands-on Session

Shared folder at
<https://tinyurl.com/2p8d3796>

- Software
- Additional software
- Exercise sheet

Session 2. Hands-on Session with notebooks

- TG walk through notebooks
 - (30 mins)
- Students work through notebooks
 - (90 mins)

Session 2. Hands-on with Notebooks

Workshop notebooks

- TG walk through notebooks (30 mins)
- Students work through notebooks (90 mins)

Resources for Hands-on Session

Shared folder at

<https://tinyurl.com/2p8d3796>

Linked Art discussion group

<https://tinyurl.com/3rbcrb4v>