Utrecht University

Cultural Heritage Data Workshop

Centre for Digital Humanities
& Utrecht University Library
May 19, 2025
Utrecht

Stefano Rapisarda,
Chiara Livio, c.livio@uu.nl

Introductions: you

- Name
- Role
- Field of research
- What data are you working with, and how (if any)?

Introductions: us and the library

Make your research privacy proof with the new e-learning Privacy Basics for Researchers

START TODAY! >

Research Data Management Support

Guides Tools Workshops Support Contact

Quick links:

- WE REVIEW YOUR DATA MANAGEMENT PLAN
- HOW TO MANAGE YOUR RESEARCH DATA
- HOW TO PUBLISH YOUR RESEARCH DATA & SOFTWARE AS OPENLY AS POSSIBLE

To support, advise and assist all researchers at Utrecht University with data management, data analysis, research software and code, a multidisciplinary network of data experts, software specialists and research engineers offer their expertise. Through training, tools, infrastructure, consultancy, and custom-made solutions, we guide scientists at every stage of their research.

Introductions: us and the library

Research data trainings

- Quickstart to Research Data Management
- Getting started with Yoda
- Getting started with Virtual Research Environments (VREs)
- Learn to write your DMP (online training, self-paced)
- Privacy basics for researchers (online training, self-paced)

Research software

- Introduction to R and Data
- Introduction to Python
- Writing Reproducible Manuscripts in R and Python
- Best Practices in Writing Reproducible Code

Walk-in events

- Walk-In Hours (weekly)
- Programming Café (monthly)

Tailor-made workshops

Searching for literature

Advice & support to...

- Lecturers
- Researchers
  - Information per discipline
  - Visibility
  - Systematic reviews
  - Reference management
  - Publishing Support
    - Pure research information
    - **Open access**
      - Funding open access
      - Open access in practice
      - Reliable publishers and predatory publishers

Open access

Through open access (OA), research results from around the world are freely accessible to all. For example, academic publications, research data and e-books. Open access is an important part of the Open Science movement.

Open access publishing

The start of your research is a good moment to start thinking about open access publishing. Are you unsure about your options? For example, because you are receiving a grant from an external research funder? Request a consultation with the Publishing Support department.

REQUEST A CONSULTATION WITH PUBLISHING SUPPORT ›

Agenda

1h: **group activity** and discussion
1h: **data toolkit**
1h: data **exploration**
30 min: **reflection** and final considerations
30 min: extra time for **questions**

Cultural Data

How to design knowledge and sculpt your dataset

Why work with data
in the Humanities?

Anything can be interpreted as data

Group activity

Activity description

- Groups
- 10 minutes per group of objects
- 3 types of objects
- Working in groups, catalogue these objects in a tabular database like a spreadsheet
- What information would you think is important to be collected?
- Write the different columns of your dataset

Group presentations

Break

Every act of data-making
(=data modeling)
involves decisions

What is a data model?

A **data model** is a **blueprint for how data is structured, organized, and related**—it defines the logic of your dataset before you even start collecting data, the abstract organisation of the information within a database, where standards and relations are visually explained.

What (usually) drives your data model?

Your research question!

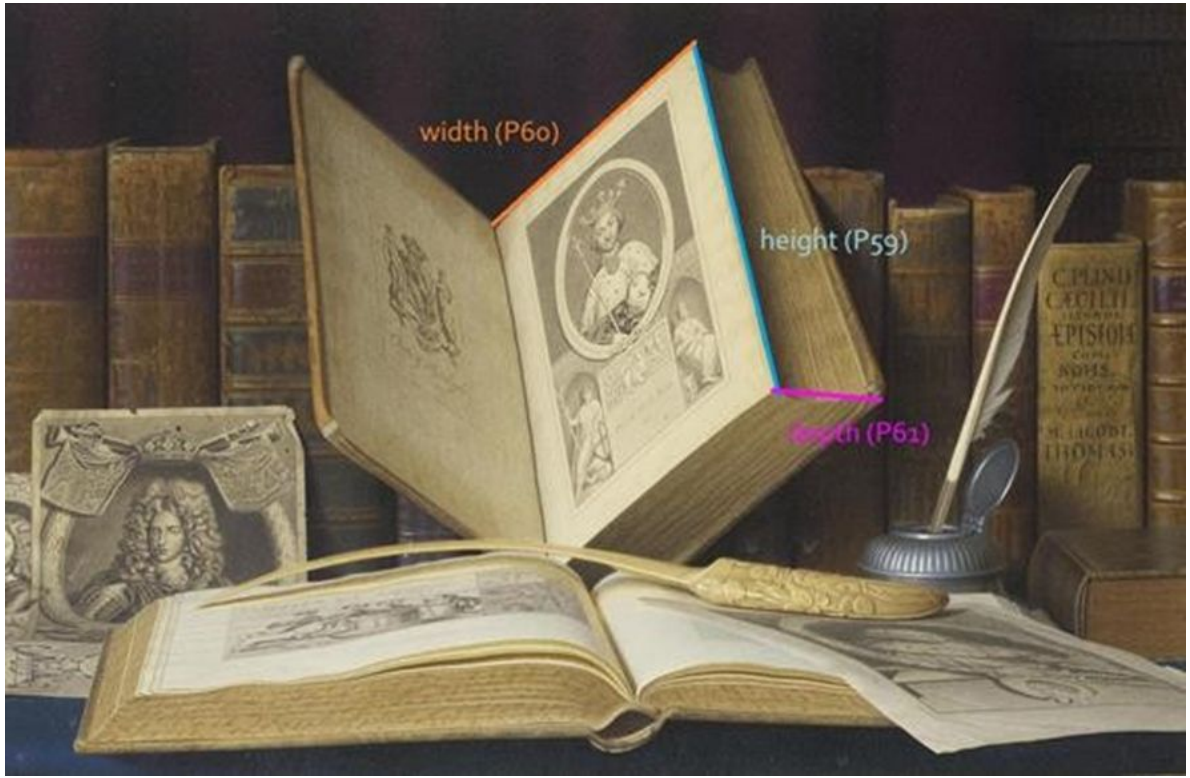
Assume that

Raw data do not exist

Data Toolkit

How can you organize historical/cultural data?

Finding Data Models



https://database.factgrid.de/wiki/FactGrid:Data_model_for_manuscripts

1 manuscript items
- 1.1 Label and Description
  - 1.1.1 Label
  - 1.1.2 Description
  - 1.1.3 Alias
- 1.2 instance of
  - 1.2.1 manuscript type
  - 1.2.2 lost manuscripts
- 1.3 core information (DFG:Schlagzeile)
- 1.4 Project information
- 1.5 Provenance
- 1.6 online resources
- 1.7 Literature

2 special information for codices
- 2.1 instance of
- 2.2 materiality of codices

3 manuscript content
- 3.1 section in a document
  - 3.1.1 necessary information
  - 3.1.2 PhiloBiblon textual witness
  - 3.1.3 manuscript related information
  - 3.1.4 content related information
    - 3.1.4.1 PhiloBiblon: treatment of incipit Property:P70 and
  - 3.1.5 external information
- 3.2 work
  - 3.2.1 PhiloBiblon: Treatment of work
- 3.3 Field of Knowledge and Topic/Object of Discourse

Alexander Severus
A.D. 231-235

Denomination
Denarius (67)

Material
Silver (87)

Iconography
"spear" (62)

Script
Latin (471)

Institution
U.Va. (494)

Department
Roman (448)

Legend
"imperator" (266)

Collection
U.Va. Miscellaneous (48)

Index Terms

Century
A.D. 3rd (363)

Deity
Mars (13)

Dynasty
Severan (14)

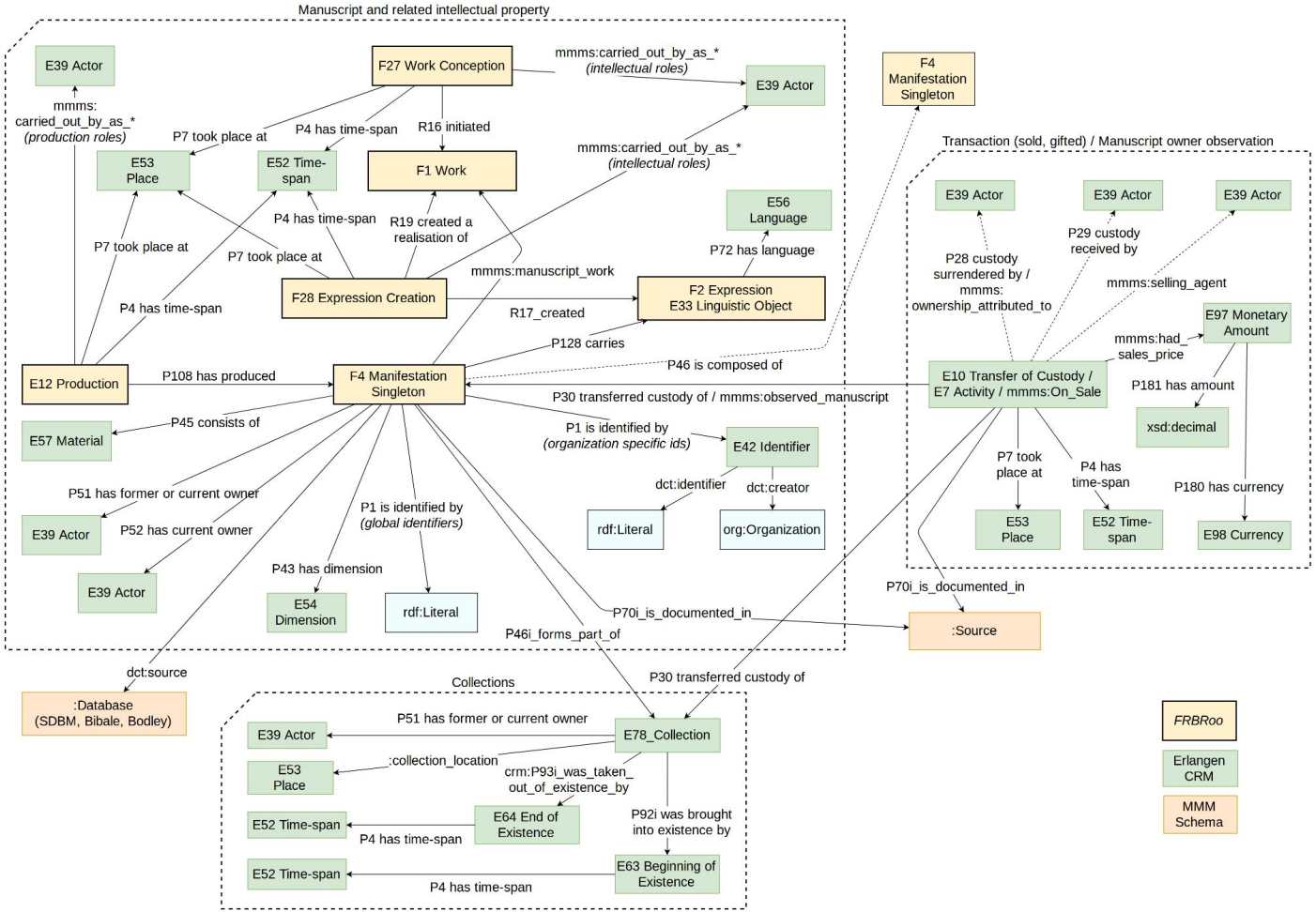
Geographical Locations

Names
Alexander Severus (4)

State
Italy (277)

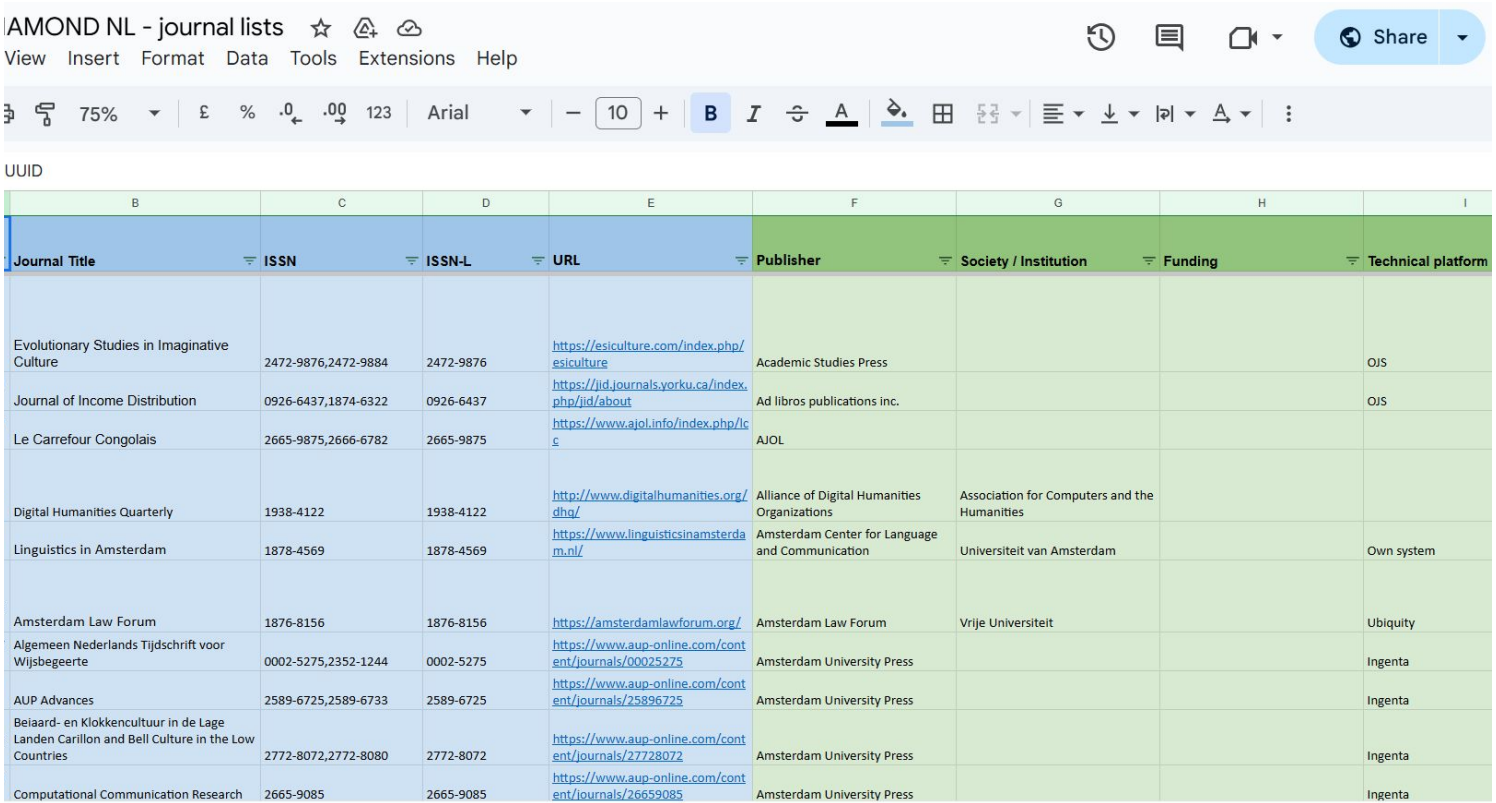
Region
Lazio (239)

City
Rome (239)



https://mapping-manuscript-migrations.github.io/data_model/mmm.html

Before adding information



	B	C	D	E	F	G	H	I
	Journal Title	ISSN	ISSN-L	URL	Publisher	Society / Institution	Funding	Technical platform
	Evolutionary Studies in Imaginative Culture	2472-9876,2472-9884	2472-9876	https://esiculture.com/index.php/esiculture	Academic Studies Press			OJS
	Journal of Income Distribution	0926-6437,1874-6322	0926-6437	https://jid.journals.yorku.ca/index.php/jid/about	Ad libros publications inc.			OJS
	Le Carrefour Congolais	2665-9875,2666-6782	2665-9875	https://www.ajol.info/index.php/lcc	AJOL			
	Digital Humanities Quarterly	1938-4122	1938-4122	http://www.digitalhumanities.org/dhq/	Alliance of Digital Humanities Organizations	Association for Computers and the Humanities		
	Linguistics in Amsterdam	1878-4569	1878-4569	https://www.linguisticsinamsterdam.nl/	Amsterdam Center for Language and Communication	Universiteit van Amsterdam		Own system
	Amsterdam Law Forum	1876-8156	1876-8156	https://amsterdamlawforum.org/	Amsterdam Law Forum	Vrije Universiteit		Ubiquity
	Algemeen Nederlands Tijdschrift voor Wijsbegeerte	0002-5275,2352-1244	0002-5275	https://www.aup-online.com/content/journals/00025275	Amsterdam University Press			Ingenta
	AUP Advances	2589-6725,2589-6733	2589-6725	https://www.aup-online.com/content/journals/25896725	Amsterdam University Press			Ingenta
	Beiaard- en Klokkencultuur in de Lage Landen Carillon and Bell Culture in the Low Countries	2772-8072,2772-8080	2772-8072	https://www.aup-online.com/content/journals/27728072	Amsterdam University Press			Ingenta
	Computational Communication Research	2665-9085	2665-9085	https://www.aup-online.com/content/journals/26659085	Amsterdam University Press			Ingenta

entity

What are you modeling?

AMOND N - journal lis

View Insert Format Data Tools Extensions Help

Share

UUID

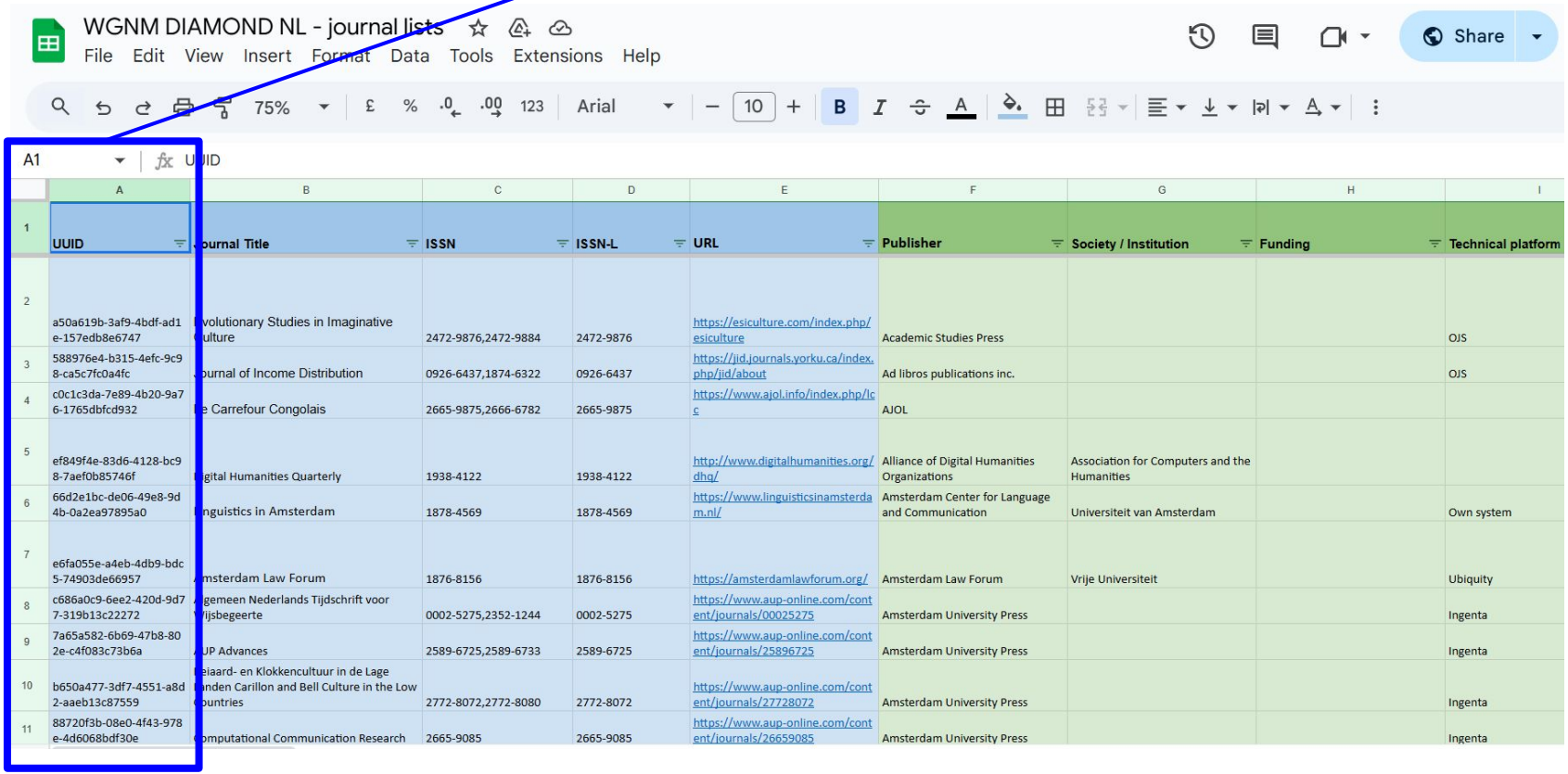
B	C	D	E	F	G	H	I
Journal Title	**ISSN**	**ISSN-L**	**URL**	**Publisher**	**Society / Institution**	**Funding**	**Technical platform**
Evolutionary Studies in Imaginative Culture	2472-9876,2472-9884	2472-9876	https://esiculture.com/index.php/esiculture	Academic Studies Press			OJS
Journal of Income Distribution	0926-6437,1874-6322	0926-6437	https://jid.journals.yorku.ca/index.php/jid/about	Ad libros publications inc.			OJS
Le Carrefour Congolais	2665-9875,2666-6782	2665-9875	https://www.ajol.info/index.php/lcc	AJOL			
Digital Humanities Quarterly	1938-4122	1938-4122	http://www.digitalhumanities.org/dhq/	Alliance of Digital Humanities Organizations	Association for Computers and the Humanities		
Linguistics in Amsterdam	1878-4569	1878-4569	https://www.linguisticsinamsterdam.nl/	Amsterdam Center for Language and Communication	Universiteit van Amsterdam		Own system
Amsterdam Law Forum	1876-8156	1876-8156	https://amsterdamlawforum.org/	Amsterdam Law Forum	Vrije Universiteit		Ubiquity
Algemeen Nederlands Tijdschrift voor Wijsbegeerte	0002-5275,2352-1244	0002-5275	https://www.aup-online.com/content/journals/00025275	Amsterdam University Press			Ingenta
AUP Advances	2589-6725,2589-6733	2589-6725	https://www.aup-online.com/content/journals/25896725	Amsterdam University Press			Ingenta
Beiaard- en Klokkencultuur in de Lage Landen Carillon and Bell Culture in the Low Countries	2772-8072,2772-8080	2772-8072	https://www.aup-online.com/content/journals/27728072	Amsterdam University Press			Ingenta
Computational Communication Research	2665-9085	2665-9085	https://www.aup-online.com/content/journals/26659085	Amsterdam University Press			Ingenta

Metadata, Attributes, Labels

What do you want to know about the entity?

ID

How do you uniquely identify each entity?



WGNM DIAMOND NL - journal lists

File Edit View Insert Format Data Tools Extensions Help

Share

A1 UUID

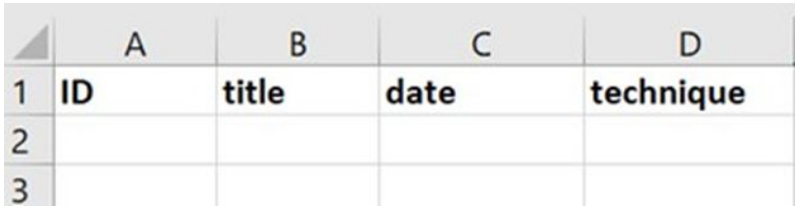
	A	B	C	D	E	F	G	H	I
1	UUID	Journal Title	ISSN	ISSN-L	URL	Publisher	Society / Institution	Funding	Technical platform
2	a50a619b-3af9-4bdf-ad1e-157edb8e6747	Evolutionary Studies in Imaginative Culture	2472-9876,2472-9884	2472-9876	https://esiculture.com/index.php/esiculture	Academic Studies Press			OJS
3	588976e4-b315-4efc-9c98-ca5c7fc0a4fc	Journal of Income Distribution	0926-6437,1874-6322	0926-6437	https://jid.journals.yorku.ca/index.php/jid/about	Ad libros publications inc.			OJS
4	c0c1c3da-7e89-4b20-9a76-1765dbfcd932	Le Carrefour Congolais	2665-9875,2666-6782	2665-9875	https://www.ajol.info/index.php/lcc	AJOL			
5	ef849f4e-83d6-4128-bc98-7aef0b85746f	Digital Humanities Quarterly	1938-4122	1938-4122	http://www.digitalhumanities.org/dhq/	Alliance of Digital Humanities Organizations	Association for Computers and the Humanities		
6	66d2e1bc-de06-49e8-9d4b-0a2ea97895a0	Linguistics in Amsterdam	1878-4569	1878-4569	https://www.linguisticsinamsterdam.nl/	Amsterdam Center for Language and Communication	Universiteit van Amsterdam		Own system
7	e6fa055e-a4eb-4db9-bdc5-74903de66957	Amsterdam Law Forum	1876-8156	1876-8156	https://amsterdamlawforum.org/	Amsterdam Law Forum	Vrije Universiteit		Ubiquity
8	c686a0c9-6ee2-420d-9d77-319b13c22272	Algemeen Nederlands Tijdschrift voor Wijsbegeerte	0002-5275,2352-1244	0002-5275	https://www.aup-online.com/content/journals/00025275	Amsterdam University Press			Ingenta
9	7a65a582-6b69-47b8-802e-c4f083c73b6a	AUP Advances	2589-6725,2589-6733	2589-6725	https://www.aup-online.com/content/journals/25896725	Amsterdam University Press			Ingenta
10	b650a477-3df7-4551-a8d2-aaeb13c87559	Beiaard- en Klokkencultuur in de Lage Landen Carillon and Bell Culture in the Low Countries	2772-8072,2772-8080	2772-8072	https://www.aup-online.com/content/journals/27728072	Amsterdam University Press			Ingenta
11	88720f3b-08e0-4f43-978e-4d6068bdf30e	Computational Communication Research	2665-9085	2665-9085	https://www.aup-online.com/content/journals/26659085	Amsterdam University Press			Ingenta

Identifiers (IDs)

Identifier: what is an ID?

An ID variable is a variable that identifies each entity in a dataset (person, household, etc) with a distinct value.

- Unique
- Universal
- *Persistent*



	A	B	C	D
1	**ID**	**title**	**date**	**technique**
2				
3				

After adding information

Cleaning and Normalization

Consistency

- **Use of capital letters**
  Jan Steen / jan steen

- **Dates**
  03/05/1865
  05/03/1865
  1865/05/03

- **Language**
  Musée du Louvre
  Louvre Museum

- **Accents**
  Diego de Velázquez
  Diego de Velazquez

Duplicates

- **Same record with different ID**

  789393 - Rijksmuseum

  3984 - Rijksmuseum

- **Same record with different spellings and ID**

  789393 - Rijksmuseum

  3984 - Rijksmuseum Amsterdam

Controlled Vocabularies

- Helps with consistency
- Contributes to multilingual datasets
- There are created by the community or influential entities in the field and support by the community
- How can I use controlled vocabularies?

	A	B	C	D
1	**ID**	**artists name**	**artists last name**	**ULAN**
2				
3	**ID**	**artists name**	**artists last name**	**Wikidata**
4				

- Library Catalog
- Getty Research Portal
- Finding Aids for Research Collections
- Photo Archive
- Research Guides & Bibliographies
- Digital Collections
- Article & Research Databases
- Getty Provenance Index
- BHA & RILA
- Getty Vocabularies
  - Art & Architecture Thesaurus (AAT) ®
  - Cultural Objects Name Authority (CONA) ®
  - Getty Thesaurus of Geographic Names (TGN) ®
  - **Union List of Artist Names (ULAN) ®**
    - About ULAN
    - ULAN: Frequently Asked Questions
    - How to use ULAN
  - Contribute

Union List of Artist Names® Online

Search the ULAN Help

Find Name or ID: Search

Role: Clear
Lookup

Nationality:
Lookup

Pop-up Search Browse the ULAN hierarchies

Copyright information

Search Tips

For the Find Name field, you may use **AND** and **OR** [e.g., 1) eldon garnet, 2) garnet OR carnet, 3) garnet OR carnet OR karnette, 4) eldon AND carnet, 5) eldon AND (garnet OR carnet), 6) (eldon OR elton) AND (garnet OR carnet)]. Boolean operators must be in all caps (AND and OR). Wildcard is the asterisk (*); right truncation only. To find an exact match rather than a key word, use quotes [e.g., "carlos"]. There is an implied AND between the Find Name, Role, and Nationality fields.

About the ULAN

Learn about the purpose, scope and structure of the ULAN. The ULAN is an evolving vocabulary, growing and changing thanks to contributions from Getty projects and other institutions. Find out more about the ULAN's contributors.

Have a Question?

Contact the Vocabulary Program

By Diego Velázquez - Selected work 17 from Self Portrait: Renaissance to Contemporary (Anthony Bond, Joanna Woodall, ISBN 978-1855143579)., Public Domain, https://commons.wikimedia.org/w/index.php?curid=11883169

Research

Union List of Artist Names® Online
Full Record Display

Click the icon to view the hierarchy.

Semantic View (JSON, JSONLD, RDF, N3/Turtle, N-Triples)

ID: 500016881
Page Link: http://vocab.getty.edu/page/ulan/500016881

Record Type: Person

Velázquez, Diego (Spanish painter, 1599-1660)

Note: Son of Juan Rodríguez da Silva, who came from a Portuguese family that had settled in Andalusia in the early 16th century. Diego had a classical education, but preferred painting; he studied. under Francisco de Herrera the Elder. Velázquez was one of the most important European artists of the 17th century, spending his career from 1623 in the service of Philip IV of Spain.

Names:

- **Velázquez, Diego** (**preferred**,V,index,LC)
- **Diego Velázquez** (V,display)
- **Velázquez, Diego Rodríguez de Silva y** (V)
- **Diego de Silva Velàzquez** (V)
- **Velázquez** (U)
- **Diego Velasquez** (V)
- **Velázquez, Diego de Silva y** (V)
- **Diego de Silva Velazquez** (V)
- **Diego Rodriguez de Silva y Velasquez** (V)
- **ולאסקז, דיאגו** (U)
- **Diego Velazquez** (V)
- **Diego Valasquez** (V)
- **Valesquez, Diego** (V)
- **Valaquez, Diego** (V)
- **Valasques, Diego** (V)

Critical approaches to datasets

- **What** has been recorded in the dataset?
- **Where** has been the dataset compiled?
- **Who** compiled it, and **how**?
- **Why** was the dataset compiled?

AUTOR: **LOPEZ GARCIA, DIEGO.**
TITULO: **RETRATO DE SU ESPOSA CON MANTILLA.**
INVTª Nº: 534.
ESCUELA: SEVILLANA.
FECHA: AÑO 1922.
MEDIDAS: 159 x 133 cm.
MATERIA: OLEO/LIENZO.
PROCEDENCIA: COLECCION DE D. DIEGO LOPEZ GARCIA.
CONCEPTO DE INGRESO: DONACION DE D. RAFAEL LOPEZ GARCIA.
FECHA DE INGRESO: 1 DE ENERO DE 1957.
– FIRMADO Y FECHADO EN LA ZONA INF. DRCHA.: "A mi esposa. Diego Lopez. 1922".

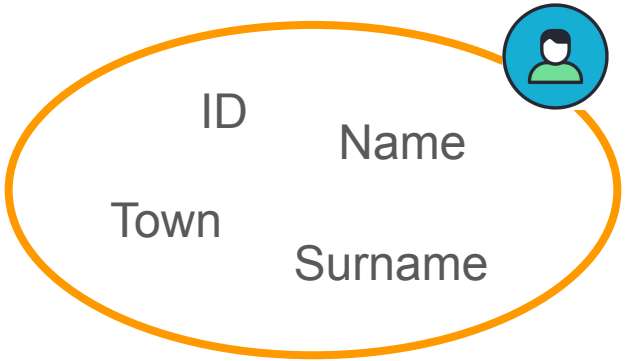
– 274 –

AUTOR: **LOPEZ GARCIA, DIEGO.**
TITULO: **RETRATO DE D. JOSE LOPEZ (HERMANO DEL AUTOR).**
INVTª Nº: 530.
ESCUELA: SEVILLANA.
FECHA: AÑO 1898.
MEDIDAS: 60 x 46 cm.
MATERIA: OLEO/LIENZO.
PROCEDENCIA: COLECCION DE D. DIEGO LOPEZ GARCIA.
CONCEPTO DE INGRESO: DONACION DE D. DIEGO LOPEZ GARCIA.
FECHA DE INGRESO: 1 DE ENERO DE 1897.
– FIRMADO EN EL ANG. INF. DRCHO.: "Diego Lopez a su hermano Pepe (+ en 10 de Enero de 1903). /Hecho en 1898.".

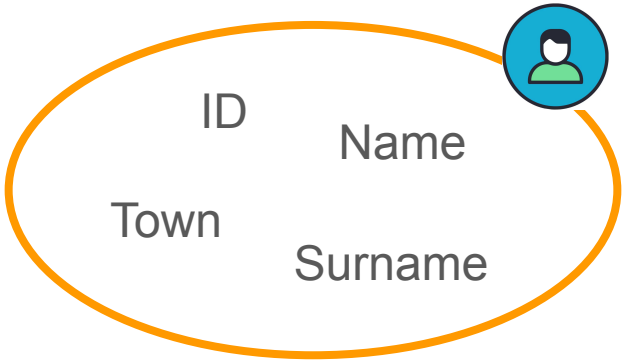
– 275 –

Páez Morales, L. (2018). De mujeres, museos y redes sociales. Porque Museo vie ne de Musa. In *Investigación y g énero. Reflexiones desde la investigación para avanzar en igualdad: VII Congreso Universitario Internacional Investigación y Género (2 018), p 5 77-588*. SIEMUS

Tables vs. Triplets

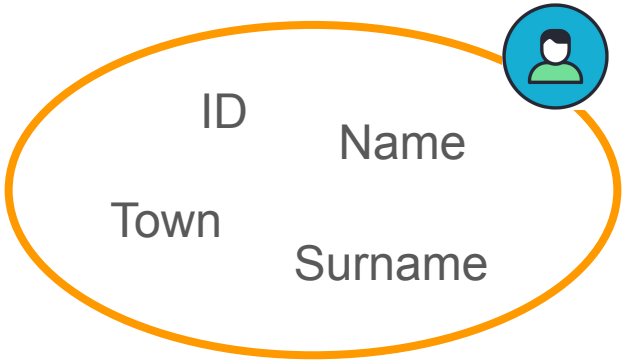


Tables vs. Triplets



ID	Name	Surname	Town
C0001	Stefano	Rapisarda	Ede

Tables vs. Triplets



ID	Name	Surname	Town
C0001	Stefano	Rapisarda	Ede

<subject>, <predicate>, <object>

C0001, hasName, Stefano
C0001, hasSurname, Rapisarda
C0001, livesIn, Ede

Linked Open Data (LOD)

https://en.wikipedia.org/wiki/Linked_data

Linked Open Data (LOD)

https://en.wikipedia.org/wiki/Linked_data

C0001, LivesIn, Ede

Linked Open Data (LOD)

https://en.wikipedia.org/wiki/Linked_data

C0001, LivesIn, Ede

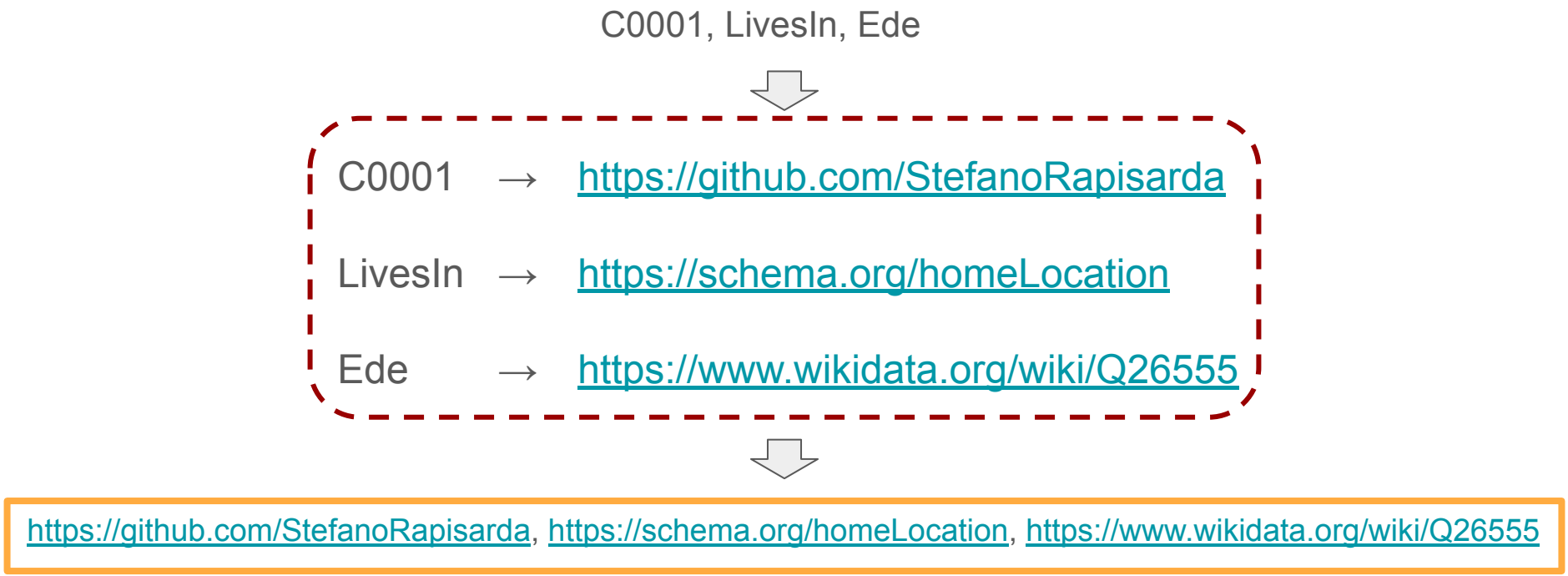
C0001 → https://github.com/StefanoRapisarda

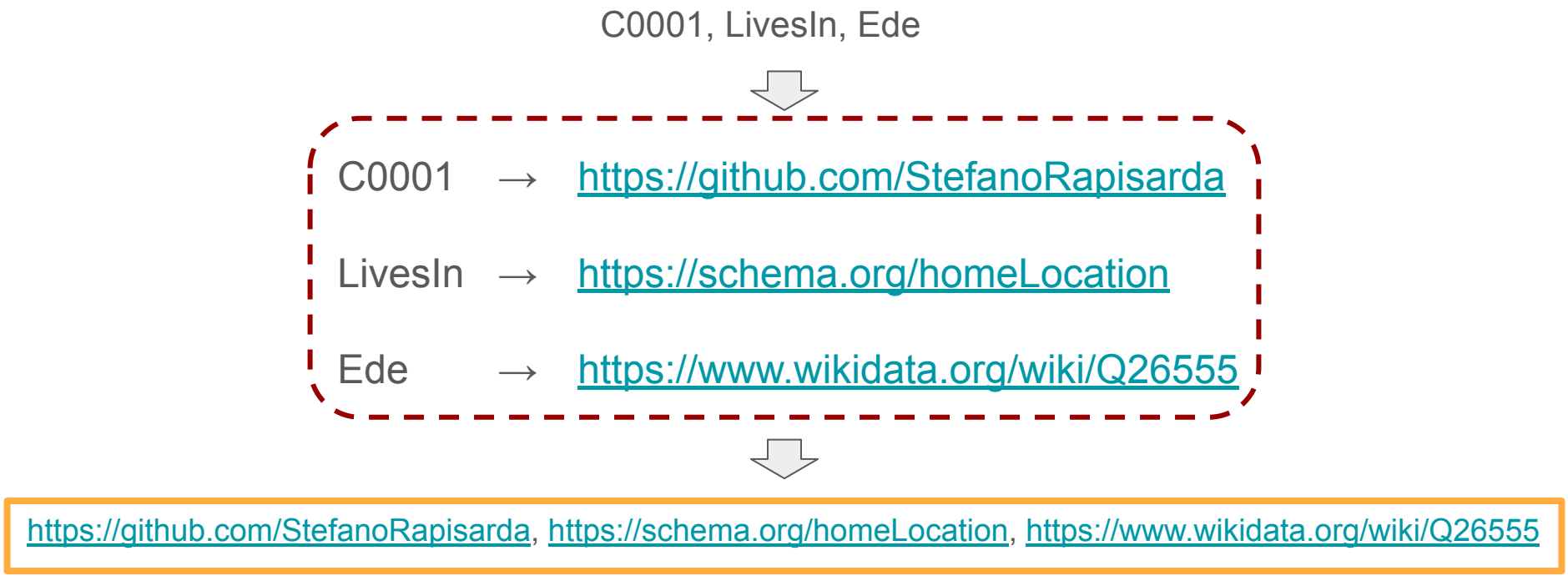
LivesIn → https://schema.org/homeLocation

Ede → https://www.wikidata.org/wiki/Q26555

Linked Open Data (LOD)

https://en.wikipedia.org/wiki/Linked_data



Linked Open Data (LOD)

https://en.wikipedia.org/wiki/Linked_data