



knowledge organisation

Here you can find information about how mythLOD data has been modelled



browse the collection

See the online catalogue and browse the collection through facets



mythLOD: from csv to rdf

The collection workflow

Data form: [Mythologiae](#)

Valentina Pasqual and Francesca Tomasi
/DH.arc Digital Humanities Advanced Research Centre
University of Bologna

THE MYTHLOD WORKFLOW



From CSV to RDF with
Mythologiae data

Factual Data

Title: L'enigma della Sfinge

Author: Tuset i Suau, Joan, 1957-^{VI}_{AF}

Keywords: sfinge enigma tuset

Typology: Pittura

Period: Arte contemporanea, XXI secolo (2004)

Description In questa peculiare rappresentazione, la Sfinge assume la forma di una donna nuda, sola e pensosa. I... [Read More](#)

See Also: [Responsible entity website](#)

Assertion

Categories: L'enigma della Sfinge

Canonical Citations: Odissea, XI, 273

Apollodoro, Biblioteca, III, 5-8

Erodoto, Storie, II, 175

Odissea, X

General References:

Cocteau, Jean, 1889-1963. | La Machine Infernale

Apollodorus. | Bibliotheca

Herodotus. | History

Homer. | Odyssey

Provenance

Interpretation Type: Iconographical Approach

Interpretation Criterion: Associazione di Fonti

Interpretation Performer: Morelli, Martina



mythLOD dataset stores
**4260 heterogeneous
artefacts** along with
their metadata

—
An Example

INPUT: THE MYTHOLOGIAE COLLECTION

Mythologiae Collection stores in a **relational database** a set of **heterogeneous artworks** spanning through several cultural institutions, artistic movements, time periods, typologies.

What artworks have in common in the collection is that each of them depicts a **mythological scene or theme**.

Annotators collected the artworks and their **descriptive metadata** (e.g. title, creation date, former location), **interpreted the artworks' content** (e.g. mythological scene) and associated a list of **literary references** that refer to the same mythological scene.

ARIANNA, VENERE E DIONISO

MAY 24, 2022

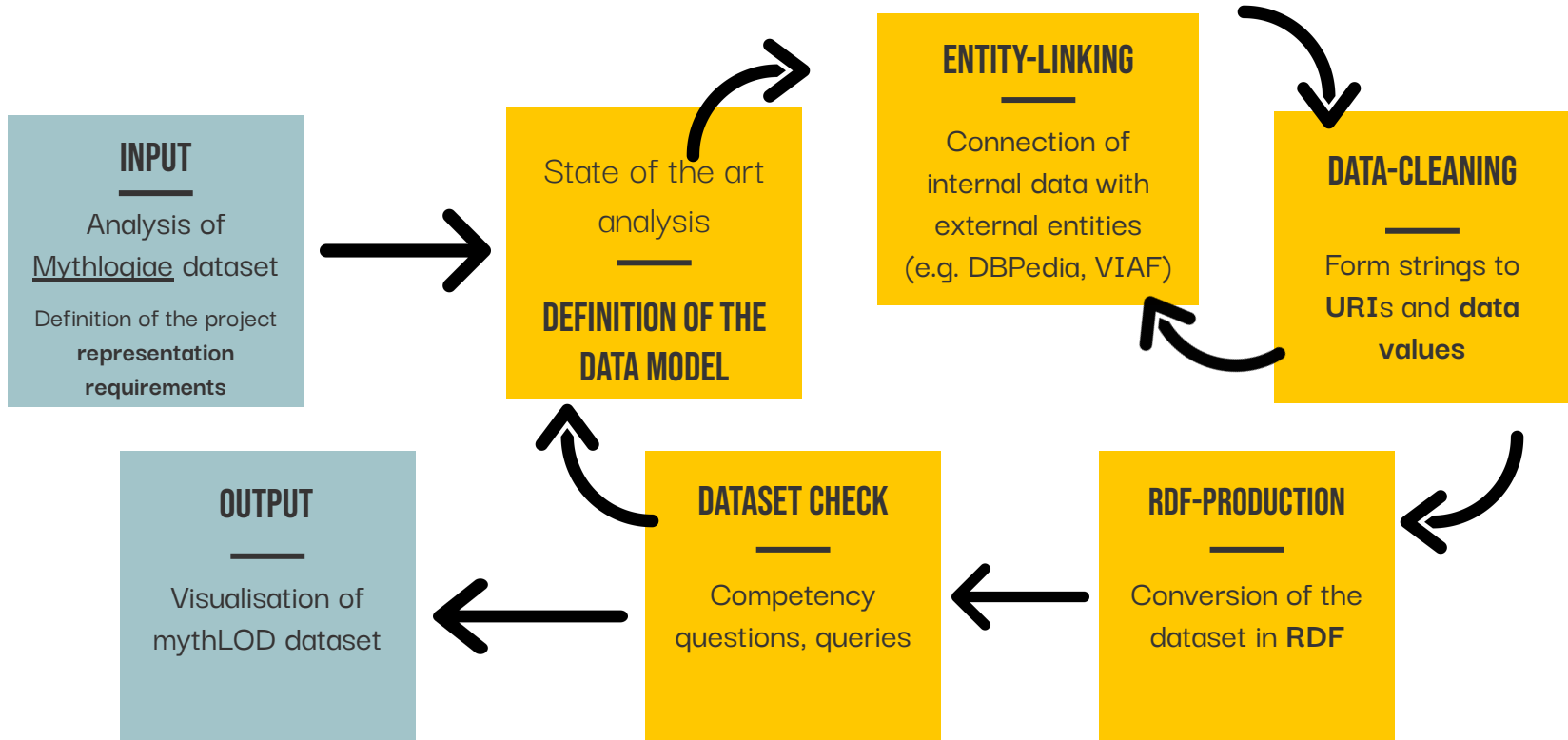
- **Titolo:** Arianna, Venere e Dioniso
- **Autore:** Tintoretto
- **Descrizione:** Venere prende la mano di Arianna, anche Dioniso vuol mano dell'amata
- **Periodo:** Arte moderna
- **Secolo:** XVI secolo
- **Data:** 1576
- **Tipologia:** Pittura
- **Collocazione:** Palazzo Ducale, Venezia
- **Note:** Olio su tela
- **Fonti letterarie classiche:** Publio Ovidio Nasone, Le Metamorfosi, VIII, vv. 177-179
- **Fonti medievali e moderne:** Giovanni Boccaccio, "Amorosa Visione", cap XII vv 14-15, Torquato Tasso, "Rime d'amore" 484, "Feci de'miei desideri" vv. 6-8
- **Riscritture letterarie:** Giovan Battista Marino, Poesie varie, Gli Incontro con Bacco*, p.186-206 *Bacco: Dioniso
- **Riscritture cinematografiche:** Nostra Signora dei Turchi, lung. drammatico, 1968, Italia, Carmelo Bene, 125 min
- **Link al museo:** [click here](#)

No standard has been defined by the annotators

CATEGORIE: **ARIANNA E DIONISO**
SI SPOSANO. KEYWORDS:
ARIANNA, DIONISO, VENEZIA.



WORKFLOW AND DATA-MANAGEMENT [1]



INPUT ANALYSIS, STUDY YOUR DATA

wp_post_status	wp_post_title	wp_post_content	wp_p	wp	wp_post_name	wp_po	wp	wp_p	wp_co	wp	wp_post_author	tx_category
auto-draft	Auto Draft			0		post		open	open	0	Manuela	
auto-draft	Auto Draft			0		post		open	open	0	Gamba Hubert	
publish	Testa di Circe			0	testa-di-circe	post		open	open	0	Gamba Hubert	circe:Circe
inherit	Testa di Circe			3479	3479-revision-v1	revision		closed	closed	0	AnnaZadra	
inherit	Circe invidiosa			6796	6796-revision-v1	revision		closed	closed	0	AnnaZadra	
publish	Circe invidiosa			0	circe-invidiosa	post		open	open	0	Gamba Hubert	circe:Circe
publish	The magic circle			0	the-magic-circle	post		open	open	0	Gamba Hubert	arianna-arianna:immagini
inherit	The magic circle			12156	12156-revision-v1	revision		closed	closed	0	AnnaZadra	
inherit	Emma Hart come Circe	<!-- wp:tadv/classic-f		12172	12172-revision-v1	revision		closed	closed	0	AnnaZadra	
publish	Emma Hart come Circe	<!-- wp:tadv/classic-f		0	emma-hurt-come-ci	post		open	open	0	Gamba Hubert	circe:Circe
inherit	Emma Hart come Circe	<!-- wp:tadv/classic-f		12172	12172-revision-v1	revision		closed	closed	0	AnnaZadra	
publish	La maga Circe			0	la-maga-circe	post		open	open	0	Gamba Hubert	circe:Circe
publish	Circe pensosa, seduta accanto ad			0	circe-pensosa-sedut	post		open	open	0	Gamba Hubert	circe:Circe
inherit	La maga Circe			575	575-revision-v1	revision		closed	closed	0	AnnaZadra	
publish	La maga Circe			0	la-maga-circe-2	post		open	open	0	Gamba Hubert	circe:Circe
inherit												
publish												
inherit												
inherit												
publish												
publish												
inherit												
inherit												
publish												
publish												
inherit												
inherit												
inherit	Disegno: Circe			14036	14036-revision-v1	revision		closed	closed	0	AlessandraPacini	

MAINTAIN

- Artworks metadata (title, date and place of creation, author..)
- Literary references metadata
- Conceptual categories (mythological themes)

AVOID/DISCARD

- Meaningless information (CMS inherited info)
- Noise in data
- Duplicates (post, revisions etc)

REPRESENTATIONAL REQUIREMENTS

In **information modeling**, a representational requirement refers to the need to accurately represent the **relevant information and data** of a **domain or problem space**. It involves **identifying the entities, attributes, relationships, and constraints involved**, and creating a formal model that represents this information in a structured and organized manner.

Representational requirements are important to **ensure** that the resulting **model is comprehensive, accurate, and effective in representing** the domain.

MYTHLOD PROBLEM SPACE

- (1) The formalisation of the relationships (identified by domain experts) between visual artworks attested in the collection and related literary sources.

REPRESENTATIONAL REQUIREMENTS IN MYTHLOD

- (1) Artwork metadata (time, place, author)
- (2) Artwork interpretation (theme)
- (3) Literary Reference (author, title, book, chapter, lines)

REVIEWING STATE OF ART, REUSE OR CREATE?

Now that the representational requirements of the project are defined, it's time to **define a data model**.

Should I reuse existing ontologies or I create a new one from scratch?

Before deciding, you should review the state of the art and understand if you can model your data on your representational requirements with an ontology that already exists. If no, the solution can be to create a new ontology.

STATE OF THE ART SURVEY IN MYTHLOD

- (1) Artwork metadata and interpretation
 - (a) CIDOC-CRM
 - (b) DCterms
 - (c) EDM (Europeana Data Model)
 - (d) FRBRoo
- (2) Literary Reference
 - (a) FRBR
 - (b) FRBRoo
 - (c) SPAR ontologies (Fabio, Cito)
 - (d) HUCit

MUSEUM DOMAIN

LIBRARY
DOMAIN

REUSING, PROS AND CONS

PROS

- (1) Reuse of fully functional data model (implemented, tested, documented) - time saving
- (2) In particular, domain ontologies are developed by outstanding institutions, this guarantees trust
- (3) Guarantees interoperability

CONS

- (1) Verbose solutions
- (2) Modelling is sometimes too general for the project purposes
- (3) Not always guaranteed that the resource will be online in the future

REUSED ONTOLOGIES IN MYTHLOD

- (1) Artwork metadata and interpretation
 - (a) FRBRoo [2]
 - (b) DCterms
- (2) Literary Reference
 - (a) FRBRoo
 - (b) DCterms
 - (c) HUCit [3]

CIDOC CRM AND FRBR

The CIDOC Conceptual Reference Model (CRM) is a theoretical and practical tool for **information integration in the field of cultural heritage**. [...]

The CIDOC CRM achieves this by providing definitions and a **formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation** and of general interest for the querying and exploration of such data.

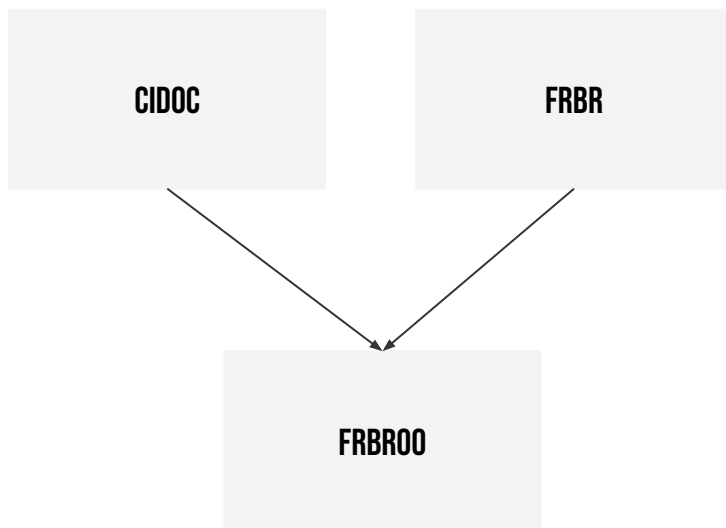
Built and maintained by **ICOM** (International Council of Museums) and adopted from several museums world-wide

The FRBR conceptual model [...] proposes the creation of bibliographic concepts ("work", "expression", "manifestation" and "item") and a new way to **formalise relations between these bibliographic entities**.

The so called WEMI-Model (Work, Expression, Manifestation and Item model) tries to identify the core aspects of **publications** and is the foundation of the FRBR family.

Built and maintained by **IFLA** (International Federation of Library Associations and Institutions) and adopted from several libraries world-wide

FRBROO



It is mythLOD backbone ontology

FRBROO is the **object-oriented** version of the **FRBR family** of conceptual models **harmonised with CIDOC CRM**. It is a **formal ontology** that captures and represents the underlying semantics of bibliographic information and therefore facilitates the integration, mediation, and **interchange of bibliographic and museum information**.

With FRBROO it is possible to express both CIDOC and FRBR domains plus some new aspects not concerned by the two previous models.

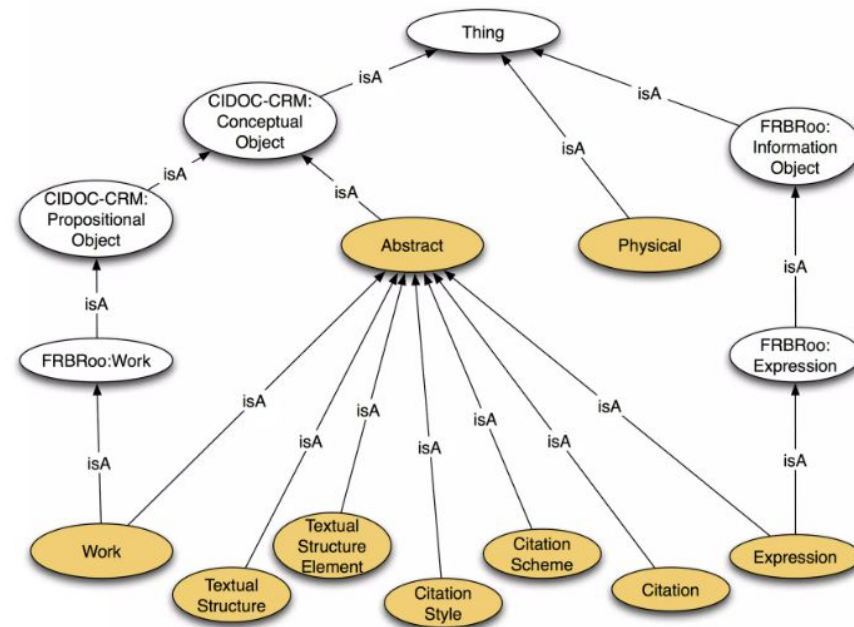
- (1) Textual elements as an unicum (e.g. manuscripts)
- (2) Performing arts

HUCIT

HuCit's Overview

HuCit is a formal ontology which extends CIDOC and FRBRoo models to represent canonical citations.

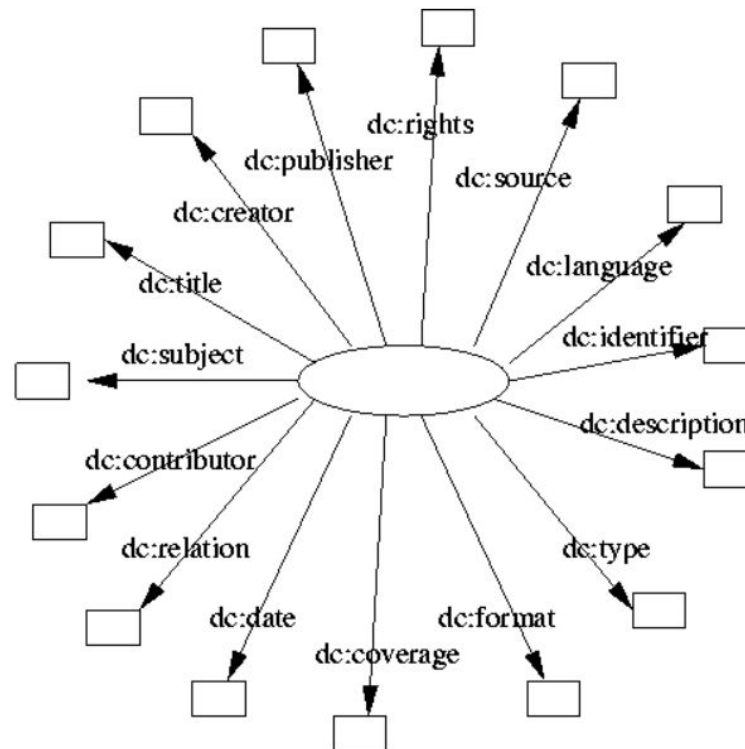
It is HuCit has been reused to specialise FRBRoo expressivity on canonical citations



DCTERMS

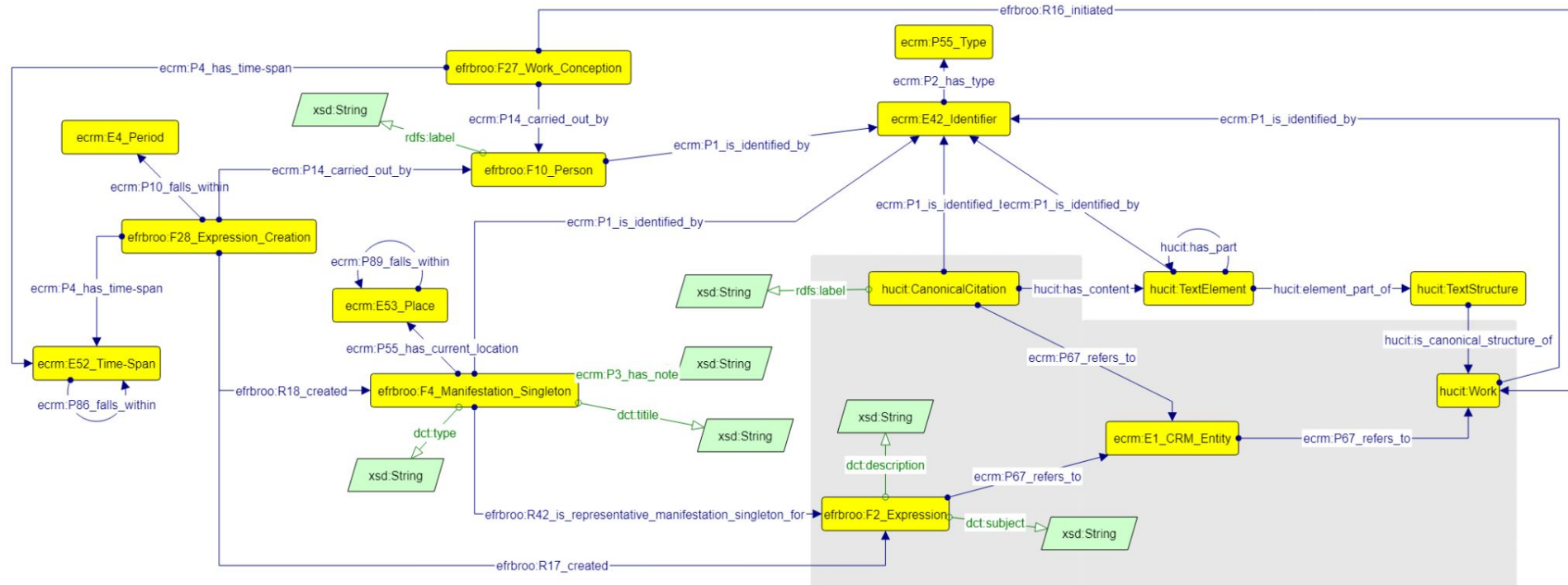
Dublin Core Terms (DCTerms) is a lightweight ontology which records basic metadata about **cataloging information**.

It is DCTerms has been reused in mythLOD to represent the items' basic information instead of reusing FRBRoo (too verbose for some simple tasks)



MYTHLOD MODEL

Modelling of (1) artwork metadata, (2) interpretation of the artwork theme and (3) bibliographic references metadata



DATA-CLEANING: FROM STRINGS TO URIS

SPLIT MULTIPLE VALUES

Unchaining multiple values in a single cell.

For example:

The string "Pan, Gli Dei" refers to two separate categories: "Pan" and "Gli Dei".

We split this values and recorded as two different URIs.

```
https://purl.org/vpq/mythlod/data/categ/pan and  
https://purl.org/vpq/mythlod/data/categ/gli-dei
```

CASE BY CASE DISAMBIGUATION

Disambiguating different strings referring to the same value

For example,

1. Giacomo, Leopardi, Canti
2. Giacomo Leopardi, Canti
3. Leopardi Giacomo, Canti

Refer to the same entity, they need to be uniformed (e.g.

```
https://purl.org/vpq/mythlod/data/work/leopardi-giacomo-canti)
```

STRINGS CONVERSION INTO URIS

For URIs: replacing spaces, special characters, accents etc.. (e.g. "Enea nella penisola italica" → `https://purl.org/vpq/mythlod/data/categ/enea-nella-penisola-italica)`

For Literals (other data-types): dates, numbers, URLs sometimes need to be adjusted to be machine-readable (e.g. "1998" → `"01-01-1998"^^xsd:date)`

ENTITY LINKING

Manual, automatic, semi-automatic.

I used **OpenRefine** (<https://openrefine.org>)

This helps you with disambiguation and also to get external information from external dataset.

ALIGNMENTS IN MYTHLOD

Artworks' authors → VIAF ID and DBpedia

References' authors → VIAF ID and DBpedia

Places → Wikidata and Geonames

ENTITY LINKING, SOME EXAMPLES

Reconciliation of different strings for the same **author**, for example

- Labels in the CSV: "Ovid" or "Ovidio"
- mythLOD URI for Ovid: <https://purl.org/vpq/mythlod/data/person/ovid>
- VIAF URI for Ovid: <https://viaf.org/viaf/88342447>
- VIAF label for Ovid: "Ovid, 43 B.C.-17 or 18 A.D."

Record in mythLOD:

```
<https://purl.org/vpq/mythlod/data/person/ovid> a efrbroo:F10_Person ;
```

```
    rdfs:label "Ovid, 43 B.C.-17 or 18 A.D."^^xsd:string ;
```

```
    owl:sameAs <https://viaf.org/viaf/88342447> .
```

ENTITY LINKING, SOME EXAMPLES

Getting external information (e.g. lat and long) for artwork's **holding institutions locations**, for example:

- Labels in the CSV: "Kunsthistorisches Museum"
- mythLOD URI for this institution: <https://purl.org/vpq/mythlod/data/place/kunsthistorisches-museum>
- Wikidata URI for this institution: <https://www.wikidata.org/Q95569>
- Wikidata latitude and longitude values: 48.203611111111,16.361666666667

Record in mythLOD:

```
<https://purl.org/vpq/mythlod/data/place/kunsthistorisches-museum> a ecrm:E53_Place ;  
  rdfs:label "Kunsthistorisches Museum"^^xsd:string ;  
  ecrm:P2_has_type <https://purl.org/vpq/mythlod/data/type/collocazione> ;  
  ecrm:P89_falls_within <https://purl.org/vpq/mythlod/data/place/austria>,  
<https://purl.org/vpq/mythlod/data/place/wien> ;  
  owl:sameAs <http://sws.geonames.org/6354990>,  
  <http://worldcat.org/identities/lccn-n50056014>,  
  <https://id.loc.gov/authorities/names/n50056014>,  
  <https://www.wikidata.org/Q95569> ;  
  wdt:P625 "48.203611111111,16.361666666667"^^xsd:string .
```

RDF PRODUCTION

WHERE ARE WE IN THE WORKFLOW?

- (1) Input data has been surveyed and analysed
- (2) The data model has been defined
- (3) Input data has been cleaned, disambiguated and enriched



RDF-PRODUCTION

Creation of the actual Knowledge Base which stores the cleaned data following the data model.

RDF-PRODUCTION IN MYTHLOD

mythLOD stores 4260 artworks (=4260 rows in the input CSV), an automatic solution was needed

mythLOD Knowledge Base has been produced via [RDFLib Python library](#).

```
@prefix co: <http://purl.org/co/> .
@prefix crm: <http://www.cidoc-crm.org/cidoc-crm/> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix ecrm: <http://erlangen-crm.org/current/> .
@prefix efrbroo: <http://erlangen-crm.org/efrbroo/> .
@prefix hico: <http://purl.org/emmedi/hico/> .
@prefix hucit: <http://purl.org/net/hucit#> .
@prefix myth: <https://purl.org/vpq/mythlod/data/> .
@prefix np: <http://www.nanopub.org/nschema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix prov: <http://www.w3.org/ns/prov#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix schema: <http://schema.org/> .
@prefix wdt: <http://www.wikidata.org/prop/direct/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

myth:assertion3343 {
  <https://purl.org/vpq/mythlod/data/item/3343-expression> ecrm:P67_refers_to <https://purl.org/vpq/mythlod/data/item/3343-expression> .
}

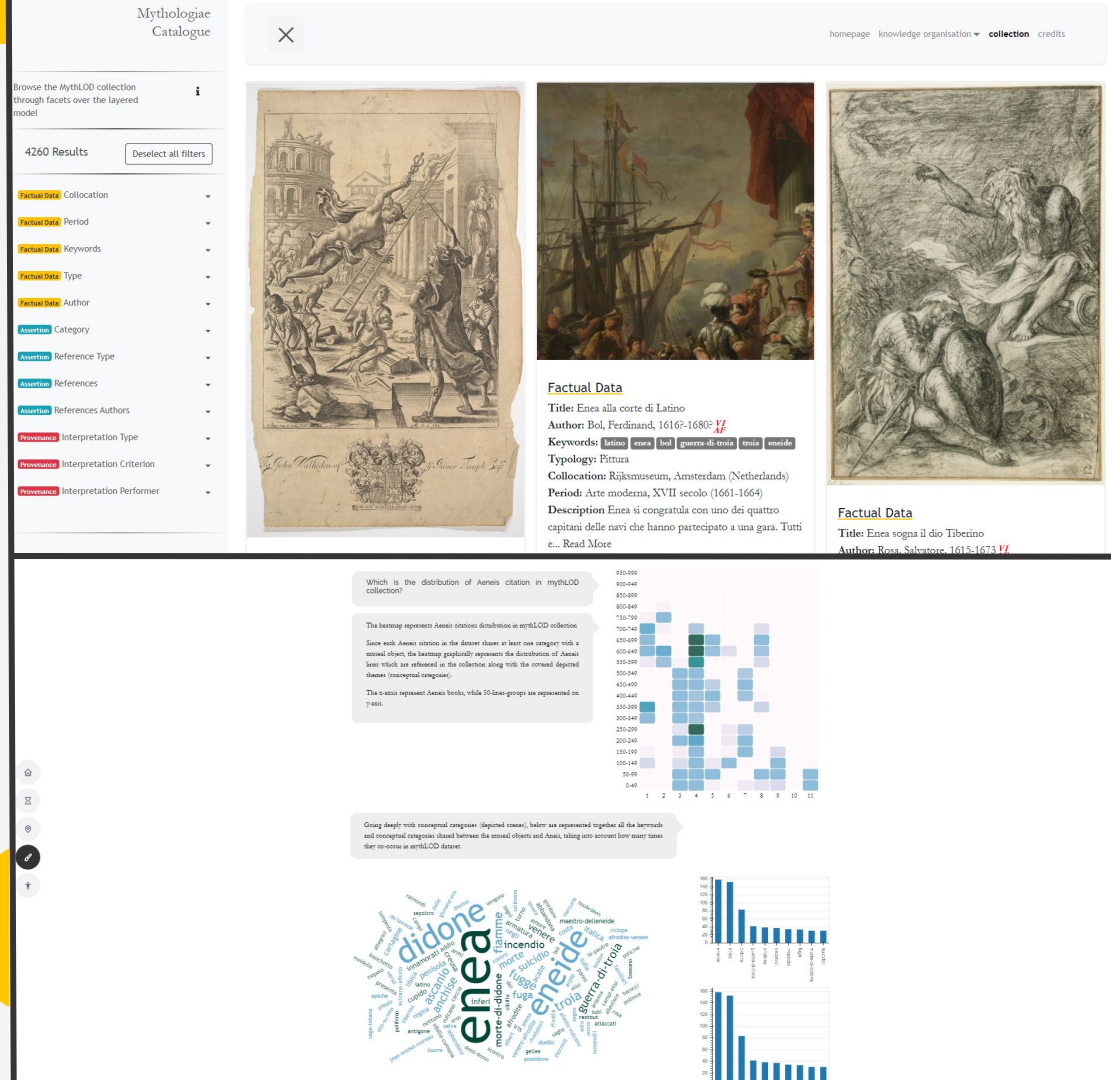
myth:pubInfo280 {
  myth:np-280 prov:wasAttributedTo <https://purl.org/vpq/mythlod/data/person/dharc> .
  prov:wasGeneratedAtTime "2020-08-24T09:00:00"^^xsd:dateTime .
}

myth:provenance622 {
  myth:assertion622 prov:wasGeneratedAtTime "2019-05-16T13:37:00"^^xsd:dateTime ;
  prov:wasGeneratedBy <https://purl.org/vpq/mythlod/data/int-act/622> .

  <https://purl.org/vpq/mythlod/data/int-act/622> a prov:InterpretationAct ;
  hico:hasInterpretationCriterion myth:hermeneutic-analysis ;
  hico:hasInterpretationType myth:iconographic-approach ;
  prov:wasAttributedTo <https://purl.org/vpq/mythlod/data/person/gamba-hubert> .
}
```

VISUALISING DATA

The dataset can be reused to become the source of web visualisations



REFERENCES

[1] Pasqual, Valentina, e Francesca Tomasi. 2022. "Linked Open Data Per La Valorizzazione Di Collezioni Culturali: Il Dataset MythLOD". AIB Studi 62 (1):149-68. <https://doi.org/10.2426/aibstudi-13301>.

[2] Bekiari, Chrysoula, et al. 2017. "Definition of FRBROo: A conceptual model for bibliographic information in object-oriented formalism." <https://repository.ifla.org/handle/123456789/659>.

[3] Matteo Romanello and Michele Pasin. 2013. "Citations and annotations in classics: old problems and new perspectives. In Proceedings of the 1st International Workshop on Collaborative Annotations in Shared Environment: metadata, vocabularies and techniques in the Digital Humanities (DH-CASE '13). Association for Computing Machinery, New York, NY, USA, Article 2, 1-8. <https://doi.org/10.1145/2517978.2517981>.

[] Daquino, Marilena, Valentina Pasqual, and Francesca Tomasi. 2020. "Knowledge Representation of Digital Hermeneutics of Archival and Literary Sources". JLIIS.It 11 (3):59-76. <https://doi.org/10.4403/jlis.it-12642>.

THANK YOU FOR YOUR ATTENTION

Any questions?

Valentina Pasqual valentina.pasqual2@unibo.it

Francesca Tomasi francesca.tomasi@unibo.it

/DH.arc Digital Humanities Advanced Research Centre
University of Bologna

EXERCISE



From CSV to RDF with
Mythologiae data

INPUT DATA ANALYSIS (CSV)

item_id	item_title	item_categories	century	author	classic_references	mythologiae_link
246	Lo sbarco di Enea in Italia	Enea nella Penisola Italica	XVIII secolo	Mason, James Lorrain, Claude	Virgilio, Eneide	https://mythologiae.unibo.it/index.php/2019/05/15/la-terra-di-enea-in-italia/
854	Vaso raffigurante Prometeo	Prometeo incatenato	XIX secolo	Simyan, Victor Etienne	Eschilo, Prometeo incatenato	https://mythologiae.unibo.it/index.php/2019/06/20/vaso-prometeo/
942	Flora in giardino	Flora, Gli Dèi	XVII secolo	Van Avont, Peeter e Brueghel, Jan		https://mythologiae.unibo.it/index.php/2019/06/12/flora-in-giardino/
3997	Pan e Siringa	Pan, Gli Dei	XVIII secolo	Boucher, François	Ovidio, Metamorfosi	https://mythologiae.unibo.it/index.php/2017/07/06/pan-e-siringa-7/
2149	Amore divino e profano	Eros, Gli Dèi	XIX secolo	Mandrado y Agudo, Josè de		https://mythologiae.unibo.it/index.php/2019/05/09/amore-divino-e-profano/

Artwork
id

String
(literal)

Multiple
records in
one cell

Multiple
values per
cell: start
time, end
time, period
typology

Multiple
records in
one cell

Multiple values
per cell: author +
textual work +
work type

REPRESENTATIONAL REQUIREMENTS AND MODELLING CHOICES

Problem space

Representation of the relation between
visual and textual items

in particular
.....

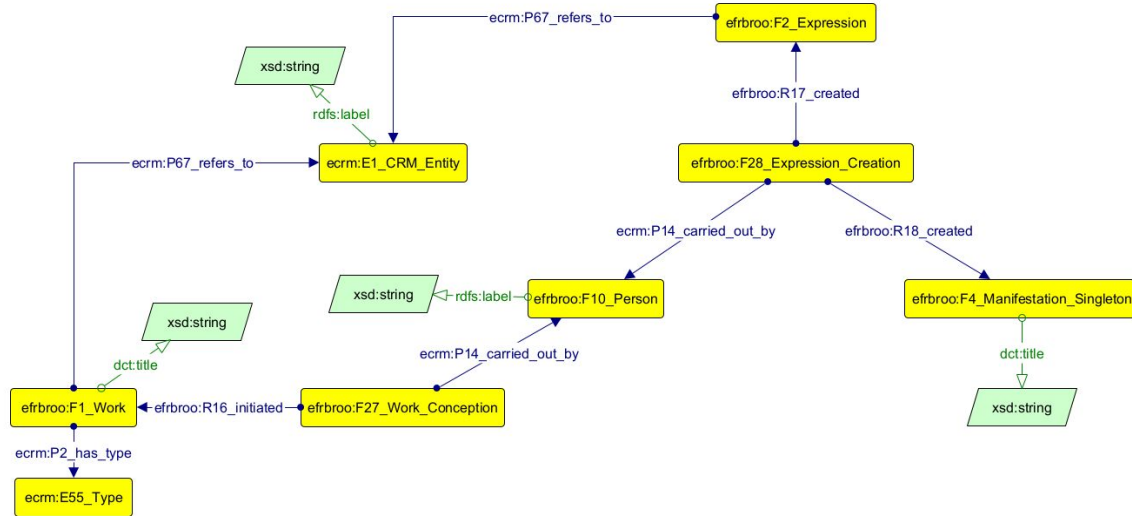
- (1) Artwork metadata (item_id, item_title, author, century)
- (2) Literary references (classic_references = title, type and author)
- (3) Artwork interpretation (item_category = mythological theme)

How can we model it?
Reuse and/or create a model?

Is it possible to establish new
relationships between data?

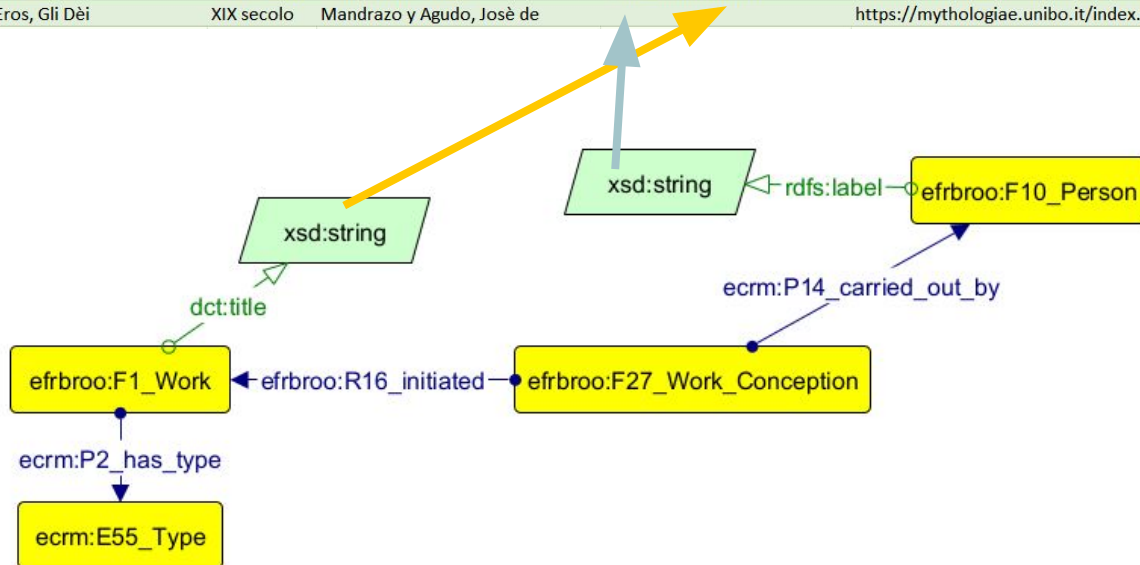
STUDYING FRBROO E DEFINING OUR DATA MODEL

FRBROO (VERSION 2.4, APPROVED BY IFLA)



MAPPING THE DATA, FROM TABLE TO TRIPLES

item_id	item_title	item_categories	century	author	classic_references	mythologiae_link
246	Lo sbarco di Enea in Italia	Enea nella Penisola Italica	XVIII secolo	Mason, James Lorrain, Claude	Virgilio, Eneide	https://mythologiae.unibo.it/index.php/2019/05/15/la-terra-di-enea-in-italia/
854	Vaso raffigurante Prometeo	Prometeo incatenato	XIX secolo	Simyan, Victor Etienne	Eschilo, Prometeo incatenato	https://mythologiae.unibo.it/index.php/2019/06/20/vaso-prometeo/
942	Flora in giardino	Flora, Gli Dèi	XVII secolo	Van Avont, Peeter e Brueghel, Jan		https://mythologiae.unibo.it/index.php/2019/06/12/flora-in-giardino/
3997	Pan e Siringa	Pan, Gli Dei	XVIII secolo	Boucher, François	Ovidio, Metamorfosi	https://mythologiae.unibo.it/index.php/2017/07/06/pan-e-siringa-7/
2149	Amore divino e profano	Eros, Gli Dèi	XIX secolo	Mandrado y Agudo, Josè de		https://mythologiae.unibo.it/index.php/2019/05/09/amore-divino-e-profano/



TRIPLES AND RDF DATA

Must be a URI	Must be a URI	Can be a URI or Literals
SUBJECT	PREDICATE	OBJECT
<code>https://w3id.org/mythlod/manif246</code>	<code>a</code>	<code>http://erlangen-crm.org/efrbroo/F4_Manifestation_Singleton</code>
<code>https://w3id.org/mythlod/cre246</code>	<code>http://erlangen-crm.org/efrbroo/R18_created</code>	<code>https://w3id.org/mythlod/manif246</code>
<code>https://w3id.org/mythlod/manif246</code>	<code>http://purl.org/dc/terms/title</code>	<code>"Lo sbarco di Enea in Italia"xsd:string</code>

URIS, DON'T

Spaces
Accents
Punctuation

URIS, DO

CamelCase
Underscore or hyphen-minus

LITERALS

Strings (natural language)
Integers (numbers)
Dates (year)
Dates (date-time)
....

ENTITY LINKING

item_id	item_title	item_categories	century	author	classic_references	mythologiae_link
246	Lo sbarco di Enea in Italia	Enea nella Penisola Italica	XVIII secolo	Mason, James Lorrain, Claude	Virgilio, Eneide	https://mythologiae.unibo.it/index.php/2019/05/15/la-terra-di-enea-in-italia/
854	Vaso raffigurante Prometeo	Prometeo incatenato	XIX secolo	Simyan, Victor Etienne	Eschilo, Prometeo incatenato	https://mythologiae.unibo.it/index.php/2019/06/20/vaso-prometeo/
942	Flora in giardino	Flora, Gli Dèi	XVII secolo	Van Avont, Peeter e Brueghel, Jan		https://mythologiae.unibo.it/index.php/2019/06/12/flora-in-giardino/
3997	Pan e Siringa	Pan, Gli Dei	XVIII secolo	Boucher, François	Ovidio, Metamorfosi	https://mythologiae.unibo.it/index.php/2017/07/06/pan-e-siringa-7/
2149	Amore divino e profano	Eros, Gli Dèi	XIX secolo	Mandrado y Agudo, José de		https://mythologiae.unibo.it/index.php/2019/05/09/amore-divino-e-profano/

Can we align this data to external resources to get new information?

How do you model this new information?

TIP 1

We did not get additional information about classical reference authors (e.g. date of birth and death) nor for items categories...

TIP 2

VIAF, Wikidata, DBpedia, Getty Thesauri etc..