



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

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

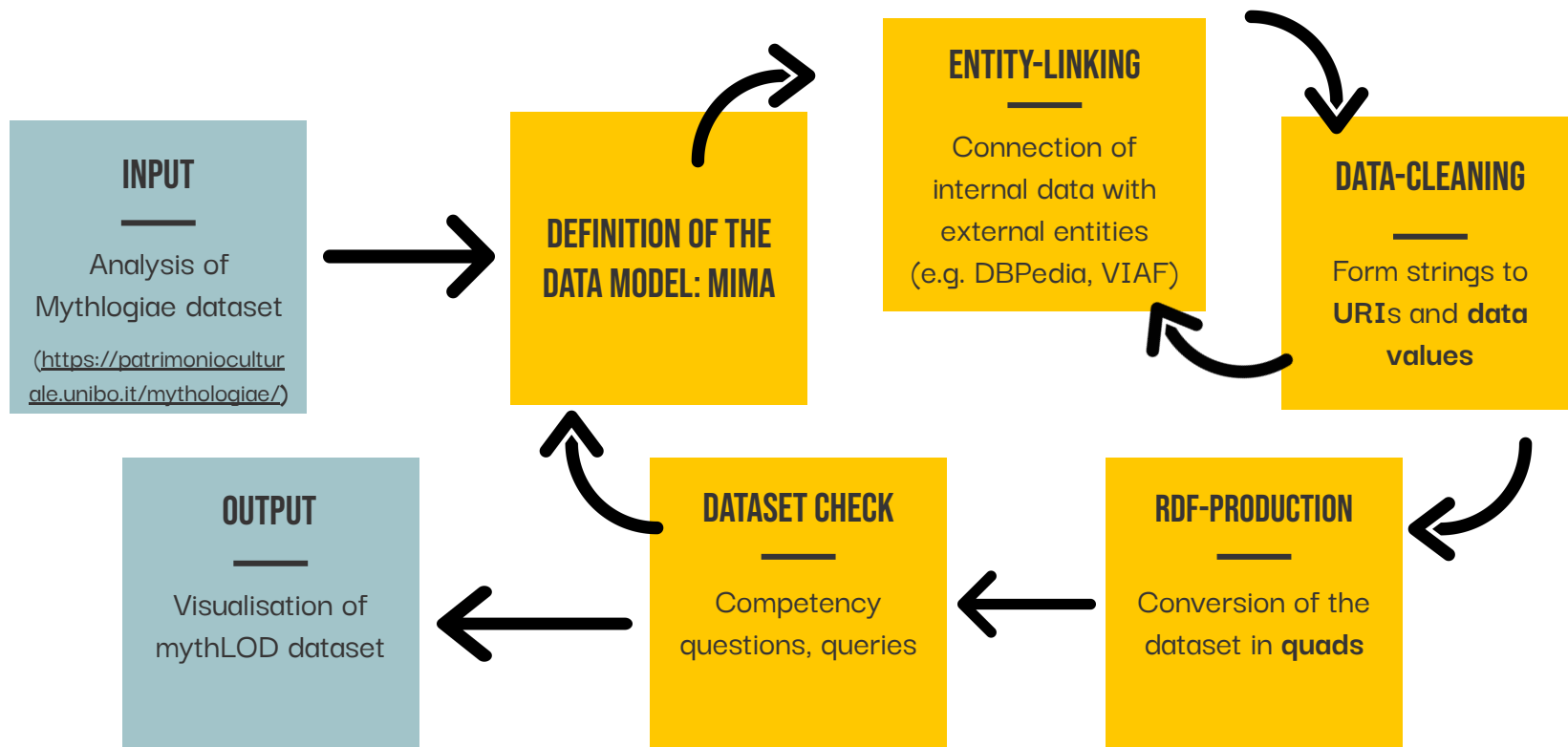
Interpretation Performer: Morelli, Martina



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

—
An Example

WORKFLOW AND DATA-MANAGEMENT



INPUT ANALYSIS

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

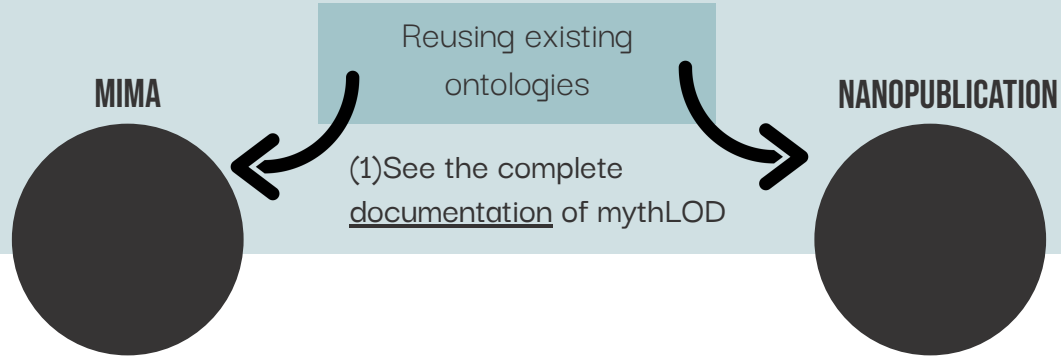
MAINTAIN

- Artworks metadata (title, date and place of creation, author..)
- Literary works metadata
- Conceptual categories
- Meta-information (e.g. author and date of each record)

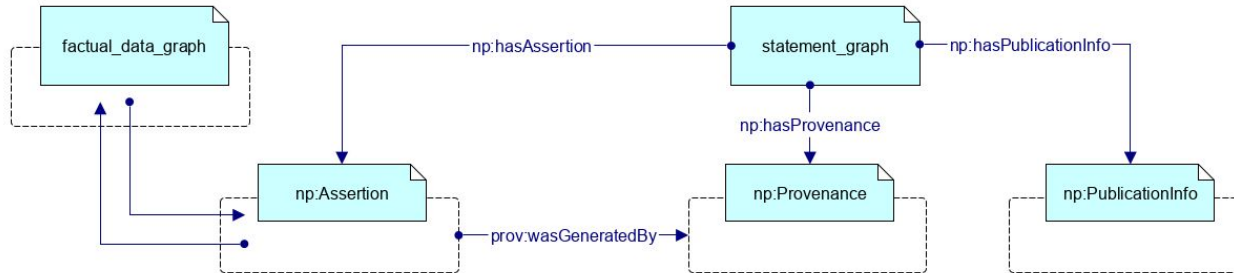
AVOID/DISCARD

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

DEFINITION OF THE DATA MODEL: MIMA



ORGANISING KNOWLEDGE INTO GRAPHS



MIMA DATA MODEL - FOUR LAYERED MODEL TO REPRESENT HERMENEUTICS

LEVEL 0: FACTUAL DATA

Artwork 246, **type** painting, **title** "L'enigma della sfinge", **author** Joan Tuset i Suau, **date of creation** 2004 etc..

All information which is not deemed to be questionable

LEVEL 1: ASSERTION

Artwork 246, represents **the Sphinx**. The same theme (the Spinx) is attested in **literary sources** such as **Aeneid, Odyssey etc...**

All information which is deemed to be questionable/subjective

LEVEL 1: PROVENANCE

The assertion author is the researcher Martina Morelli, which stated it 20/03/2020. His approach to the research was Iconographic analysis and he performed an hermeneutic analysis + literary sources linking

Meta information about the assertion

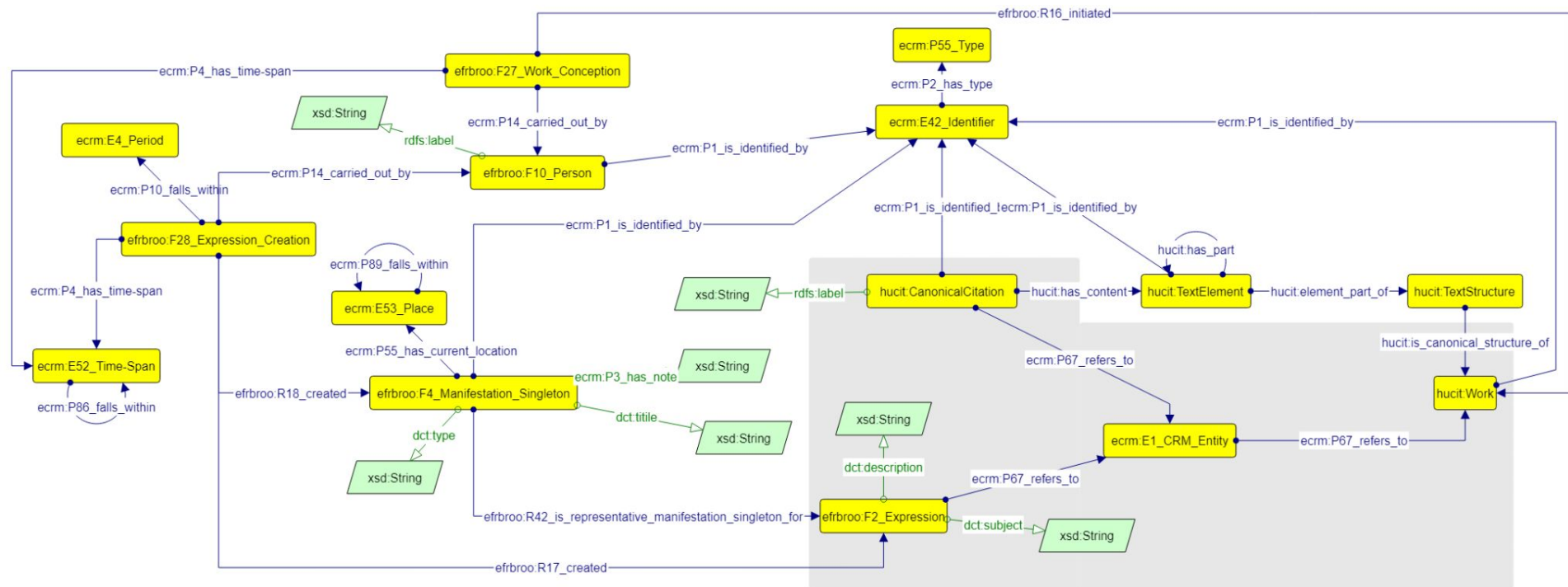
LEVEL 1: PUBLICATION INFO

Information about level 1 and 2: Valentina Pasqual interpreted the data and published them in march 2021

Meta information about the whole publication

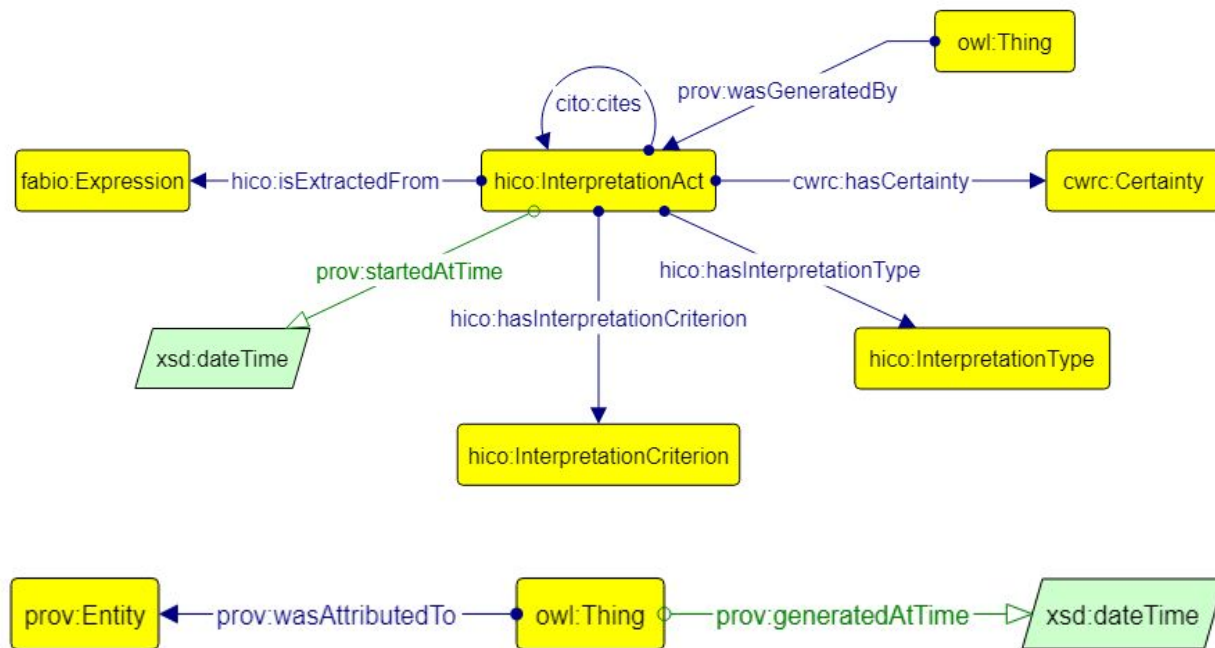
ONTOLOGIES: FACTUAL DATA (0) AND ASSERTIONS (1) MODEL

Each level reuses its own set of ontologies depending on what it is needed to be expressed



ONTOLOGIES: PROVENANCE (2) AND PUBLICATION INFO (3)

Each level reuses its own set of ontologies depending on what it is needed to be expressed



Those two levels
provide
**Accountability and
Trust** to the dataset

DATA-CLEANING: FROM STRINGS TO URIS

01 INPUT DATA ANALYSIS

Transform the dataset in a series of computable objects on which you can perform cleaning activity and disambiguation.

02 CASE BY CASE DISAMBIGUATION

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. `myth:leopardi-giacomo`)

03 FROM CSV STRINGS TO RDF URIS AND DATAVALUES

For **URIs**: replacing spaces, special characters, accents etc.. (e.g. "Enea nella penisola italica" → `myth:enea-nella-penisola-italica`);

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

04 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. Ex: aligning mythLOD places with **DBpedia** I got their lat-long coordinates

Result: fully
computable **RDF**
dataset

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

[illegible]

EXERCISE

INPUT

```
1 item_id;item_title;item_categories
2 246;Lo sbarco di Enea in Italia;Enea nella Penisola Italica
3 854;Prometeo;Prometeo, Miti di Fondazione
4 942;Flora in giardino;Flora, Gli Dei
5 3997;Pan e Siringa;Pan, Gli Dei
6 2149;Amore divino e profano; Eros, Gli Dei
```

MYTHLOD DATA

Data conversion from CSV to
RDF

OUTPUT

```
1 @prefix dct: <http://purl.org/dc/terms/> .
2 @prefix ecrm: <http://erlangen-crm.org/current/> .
3 @prefix efrbroo: <http://erlangen-crm.org/efrbroo/> .
4 @prefix myth: <https://w3id.org/mythlod/> .
5 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
6 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
7
8 myth:assertion246 {
9     myth:expr246 ecrm:P67_refers_to myth:eneaNellaPenisolaItalica .
10 }
11
12 myth:assertion3997 {
13     myth:expr3997 ecrm:P67_refers_to myth:GliDei,
14         myth:pan .
15 }
16
17 myth:assertion942 {
18     myth:expr942 ecrm:P67_refers_to myth:GliDei,
19         myth:flora .
20 }
21
22 myth:factual_data_bag {
23     myth:Eros a ecrm:E1_CRM_Entity ;
24     rdfs:label "Eros"^^xsd:string .
```

THE EXERCISE IS AVAILABLE ON [GITHUB](#)

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