Operationalizing Scholarly Observations in OWL

Emilio M. Sanfilippo,[1,2] Claudio Masolo,[1] Alessandro Mosca,[1] and Gaia Tomazzoli[3]

[1]CNR ISTC Laboratory for Applied Ontology, Trento and Catania, IT

[2]CESR University of Tours, FR

[3]Sapienza University of Rome, IT

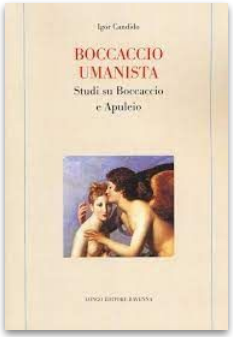
Finanziato dall'Unione europea NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

Consiglio Nazionale delle Ricerche

Introduction



<u>Our core goals</u>:

- To produce an **ontology-based** modeling framework to <u>represent</u> and logically <u>reason</u> over what scholars/critics <u>claim</u> about <u>literary texts</u>

- To contribute to the production of a <u>knowledge base</u> of <u>criticism</u>

Our approach

In a nutshell:

- **Ontology** (library of different modules) to represent *observational data* originating from literary criticism and history of literature
- **Formal mechanisms** to model and analyze scholarly arguments

↪ To primarily represent *interpretations of* literary characters found in scholars' and critics' texts

↪ The organization of interpretations as *observational data* is the first step to make the data available for further study

Examples of observational data

Boccaccio's (XIV cent) *Decameron*, *tale X.10* (*Griselda and Gualtieri*) and its interpretations.

- **Vittore Branca**: connects Boccaccio's texts with Medieval culture; connection between *Decameron* and biblic/hagiographic texts; similarities between the figures of Griselda and the Virgin Mary.

- **Michelangelo Picone**: connects Boccaccio's text with the courtly literature; in particular with Marie de France's *Lais* (XII cent); similarities between the figures of Griselda and Fresne, etc.

- **Igor Candido**: connects Boccaccio's text with some texts in classic culture, e.g. with Apuleius' (II cent.)*Metamorphoses*; similarity between the figures of Griselda and Psyche, etc.

Some research challenges

Deal with different sorts of observational data, expressed through natural language and within different methodological frames and vocabularies

↪ Intended meanings of terms may be vague or left implicit*

Deal with competing and incompatible observations based on scholarly arguments; incompatibilities cannot be always solved

↪ They must be documented together with their strengths and weaknesses**

* Pichler, A., & Reiter, N. (2022). From Concepts to Texts and Back: Operationalization as a Core Activity of Digital Humanities. Journal of Cultural Analytics, 7(4)

** Barabucci, G., Tomasi, F., & Vitali, F. (2021). Supporting complexity and conjectures in cultural heritage descriptions.

MITE observational framework (*beta version*)

Semantic Web ontology library (**OWL** + **SWRL rules**):

- **Core module**: general, domain-independent module
- **Domain-dependent modules**

↪ GitHub repository available through: https://www.loa.istc.cnr.it/mite/

Builds on previous work, in particular:

Sanfilippo, E. M., Sotgiu, A., Tomazzoli, G., Masolo, C., Porello, D., & Ferrario, R. (**2023**). Ontological Modeling of Scholarly Statements: A Case Study in Literary Criticism. In *Formal Ontology in Information Systems* (pp. 349-363). IOS Press.

Masolo, C., Botti Benevides, A., & Porello, D. (**2018**). The interplay between models and observations. Applied Ontology, 13(1), 41-71.

Observations, insights \1

Information related to the categorization of domain entities through **properties** or **relationships**, e.g.:

- Empirical scope: vase dated to 2nd century B.C. through *carbon-14 measurement*

- Literary scope: according to Vittore Branca's text *Boccaccio Medievale*, the character of Griselda in the *Decameron* is similar to the Virgin Mary in the biblic/hagiographic literature…for this-and-that reason…

> ↪ Observations may **not** be veridical with respect to the observed phenomena

Observations, insights \2

When modeling observations, it is important to track their provenance, e.g. **who** claimed them and in **which source**.

- A single observation can be claimed by different observers at different times and in different sources

↪ Observations with the same "contents" are the same independently from observers/observation acts

Observations, insights \3

Our approach does **not** cover an evaluative dimension for observations

↪ The framework does not aim to tell whether an observation is "good" or "bad" with respect to the observed entities (texts)

↪ The purpose of the framework is to document observations and provide formal mechanisms for supporting their analysis and comparison

Types of observations

At present, the core module covers three types of observations:

1. **Basic** observations (BasicObs)

2. **Source** observations (SourceObs):
   a. **Assertion** observations (AssertObs)
   b. **Rejection** observations (RejectObs)

3. **Argumentation** observations (ArgumentationObs):
   a. **Support** observations (SupportObs)
   b. **Defeat** observations (DefeatObs)

Basic Observations

↪ Represent **what scholars claim** about domain entities.

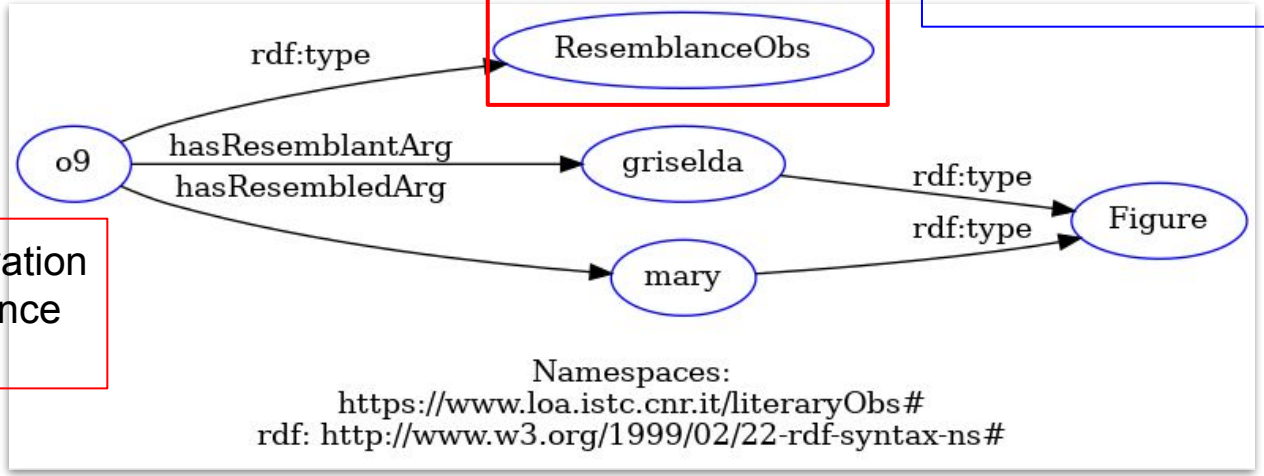
E.g., a character is the *protagonist* of a narrative, a character *resembles* another character, etc.

To represent basic observations:

- Vocabulary shared among experts: **observational vocabulary** (language)
- **Library** of **multiple vocabularies**: to represent different sorts of observations

Example: Resemblance Griselda and Virgin Mary

A subclass of BasicObs



Basic observation of Resemblance

RDF graph for the observation of resemblance between the figures Griselda and Mary

Source Observations

↪ Represent **the provenance** of an observation, i.e. the textual source where an observation is claimed

Two types of source observations:

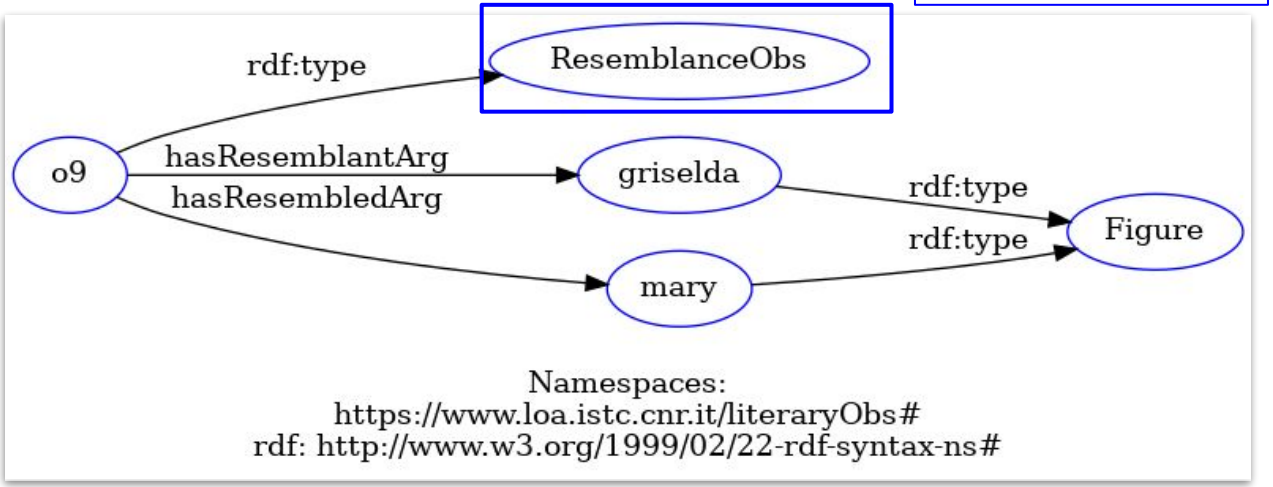
- **Assertion observation**: the source s asserts the observation o
- **Rejection observation**: the source s rejects the observation o

For the **arguments** (related entities) of source observations, we use:

- hasObs (*has observation*): the observation that is asserted or rejected
- hasSrc (*has source*): the source that either asserts or rejects an observation

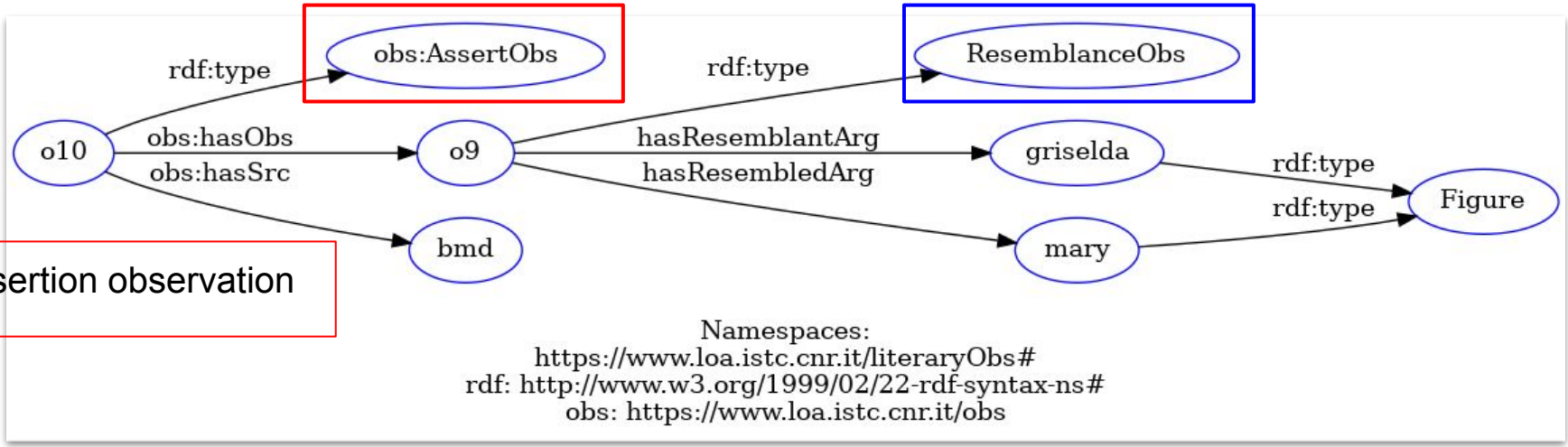
No source observation yet …

Basic observation of Resemblance



RDF graph: according to the Resemblance Observation o9, the (figure of) Griselda resembles (the figure of) Mary

Basic observation of Resemblance



Assertion observation

RDF graph: the resemblance *o9* between Griselda and Mary is asserted by the text *bmd* (*Boccaccio Medievale*, by Vittore Branca)

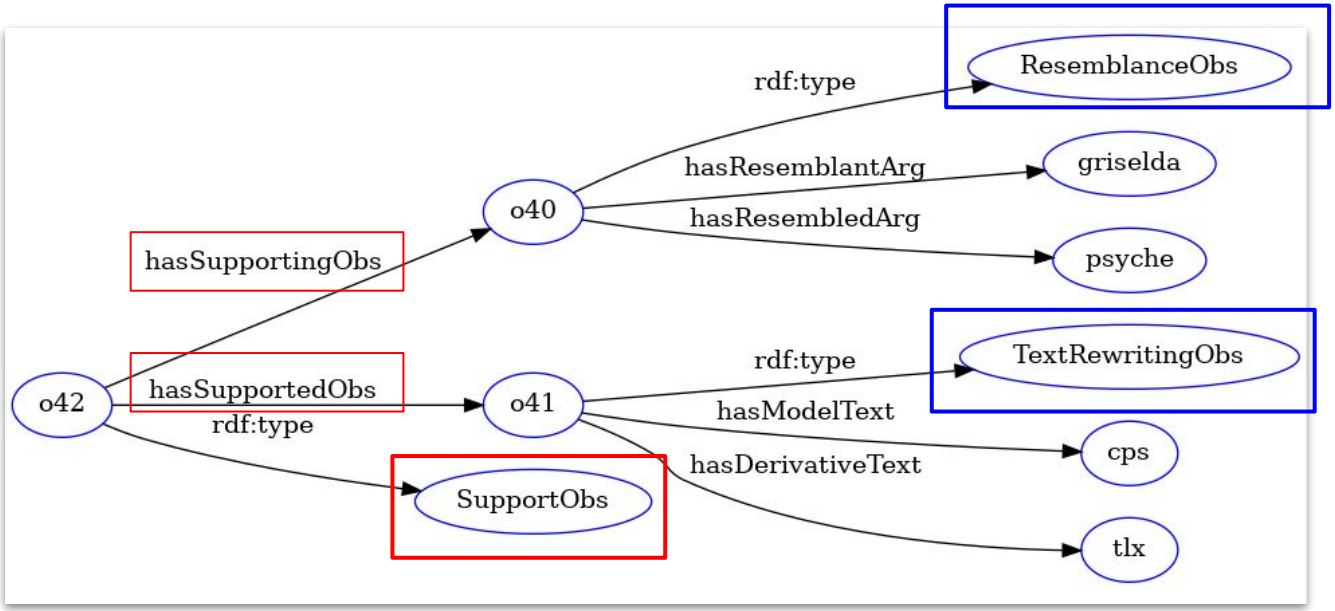
Argumentation Observations

↪ To represent "positive" or "negative" interactions between observations:

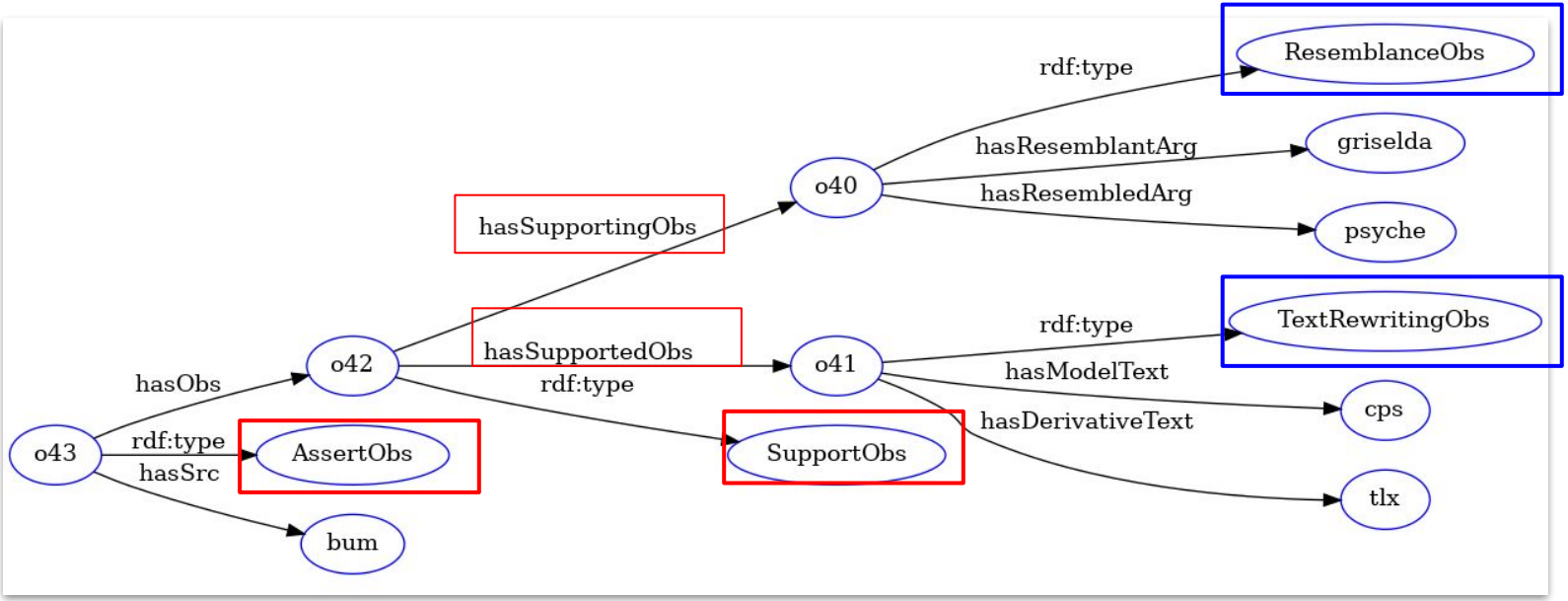
- **Support observation**: o *increases* the plausibility of o', i.e., o supports o'
- **Defeat observation**: o *decreases* the plausibility of o', i.e., o defeats o'

↪ These observations are particularly relevant when working with corpora of literary criticism, history of literature (but not only) …

… because it is commonly the case that interpreters provide arguments to support their claims or defeat others' claims



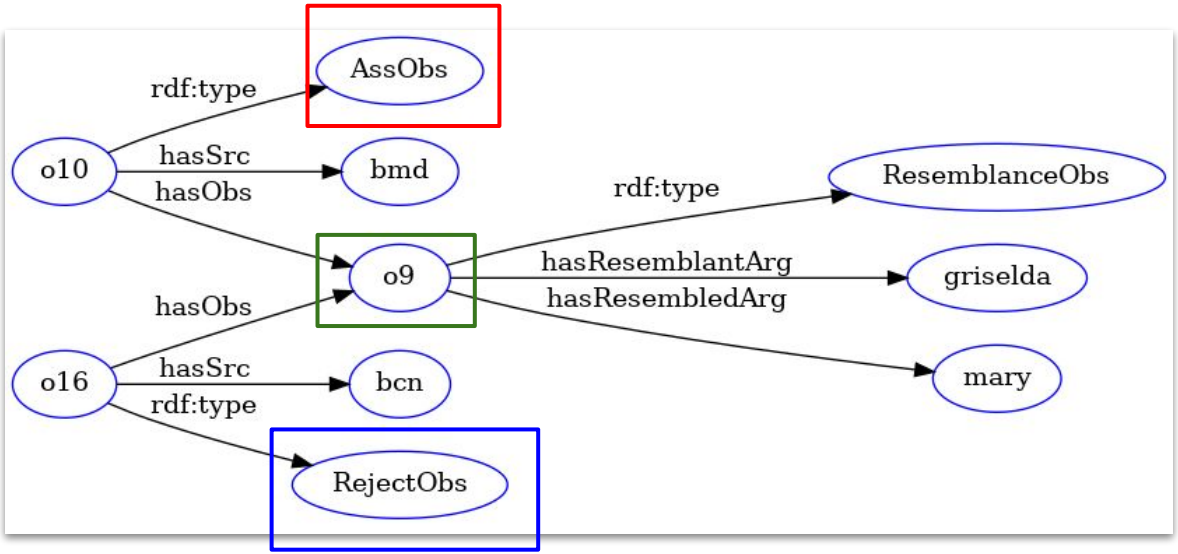
RDF graph: according to the support observation *o42*, the observation *o41* relative to the model/derivative relation between Decameron's tale *X.10* and Apuleio's *The Metamorphoses* is supported by the observation *o40* relative to the resemblance between Griselda and Psyche



RDF graph: *o43* is the source observation asserting *o42*

Criteria for analysis

A combination of **OWL** axioms and **SWRL** rules for the automatic classification of observations, sources, etc. by reasoning methods



↪ By reasoning, o9 is classified as a *disputable observation*

Conclusions

Next steps: improve and test the framework against case studies (in MITE, etc)

- Figure (class): A generic modeling element for (literary) characters; it needs to be characterized according to the analysis put forward in MITE
- Introduction of further criteria for the analysis of the data
- Modeling of the temporal dimension of scholarly observations, e.g., by reusing W3C Time Ontology for reasoning over time
- …

Thank you!

Ontology available at: https://www.loa.istc.cnr.it/mite/

Emilio M. Sanfilippo,
emilio.sanfilippo@cnr.it

Claudio Masolo, Alessandro Mosca, and Gaia Tomazzoli

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
Consiglio Nazionale delle Ricerche