Evaluation of Expressing without asserting approaches in RDF. The case of Conjectures.

Valentina Pasqual[1], Gerald Manzano[2], Eduart Uzeir[2], Francesca Tomasi[1] and Fabio Vitali[2]

[1]Digital Humanities Advanced Research Center, Department of Classica Philology and Italian studies, University of Bologna

[2]Department of Computer Science, University of Bologna

Abstract

In this paper, we evaluate the existing reification approaches for *expressing without asserting* (EWA) statements in RDF along with their related contextual information and compare them with a new method, called *Conjectures*. Conjectures express RDF statements with three states of knowledge: undisputed claims, disputed claims, and settled claims. Conjectures extend the semantics of RDF Named graphs and introduce a new syntactical form to represent both conjectural and asserted information. Our evaluation tests were performed on a large sample of Wikidata entities about artworks interspersed with additional dummy statements to simulate alternative or abandoned claims and enrich the set of non-asserted claims. Our study evaluates metrics such as the total number of triples, loading time, dataset weight, and in particular query execution time for many different and meaningful types of queries. Results show that Conjectures is competitive with existing methods and outperforms other methods in terms of efficiency when retrieving debated statements, thereby demonstrating its potential as an effective tool for expressing nuanced RDF statements.

1. Introduction

RDF [1, 2] is a powerful tool for expressing statements as absolute, asserted relationships between entities. Yet, it falls short when it comes to representing nuances about such relationships, e.g., to enrich them with contextual information, or to represent in full how they relate to each other. While some statements may never need to be questioned (or maybe there is no interest in questioning them), and therefore can be represented adequately with plain RDF triples, frequently on the other hand scientific and critical discourse is filled with concurrent opinions and interpretations. This knowledge cannot be simply expressed as plain RDF triples but requires the specification of contextual information (e.g. who is the author of such claims), and possibly of multiple and competing statements, which the community may settle or discard over time.

This is for instance the case of the hypotheses made by scholars about the many (still undetermined) locations of the Mona Lisa painting when it was stolen in 1911 [3], or of the attribution of the painting *Salvator Mundi*, possibly by Leonardo da Vinci and world-famous for having been the most expensive painting ever sold at public auction. Its attribution is now under discussion by many scholars [4]. Our guiding example in this paper, on the other hand, is the debated attributions of the painting *Napoleon crossing the Alps*[1], currently attributed to Jacques-Louis David, but previously attributed to Jérôme-Martin Langlois and the workshop of Jacques-Louis David.

RDF Reification[2] methods [5] have been used to *make statements about statements in RDF* [6] and even allowing the coexistence of multiple opinions in RDF. For instance, reification enables to express sentences like "JocondeLab[3] reports that *Napoleon crossing the alps* author is Jacques-Louis David". Usually, reification methods typically involve the use of extra triples to provide information about the original triple being reified. Listing 1 exemplifies the statement above presented using traditional reification. In the example, a new entity is introduced as an instance of the class `rdf:Statement` to introduce "JocondeLab reports" and the original triple (*Napoleon crossing the alps* author is Jacques-Louis David) is expressed by the means of three additional triples whose predicates are strictly `rdf:subject`, `rdf:predicate`, `rdf:object`.

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

:statement1 rdf:type rdf:Statement ;
            rdf:subject wd:Q19801150 ; # Napoleon Crossing the Alps
            rdf:predicate wd:P170 ; # creator
            rdf:object wd:Q67215 . # Jacques-Louis David

:statement1 wdt:P248 wd:Q29633776 . # stated in JocondeLab
```

Listing 1: Representation of the claims "XXX states that Napoleon Crossing the Alps" author is Jacques-Louis David using standard reification

Accepted statements can be therefore characterised by a plain triple to assert the statement, and the same triple under reification to provide additional information about the claim itself. This would be the case of the settled attribution of *Napoleon crossing the alps* currently attributed to Jacques-Louis David. Statements that represent alternative or historical information would then be represented with just the claim in reified form and no assertion through a plain RDF triple. This would be case of the earlier attributions of *Napoleon crossing the alps* to Jérôme-Martin Langlois and the workshop of Jacques-Louis David. We call this approach *Expressing Without Asserting* (EWA)[7] and we consider it as a powerful tool to express claims with different degrees of validity and even critical debate.

In addition to analysing the expressivity and effectiveness of the many existing reification approaches, their efficiency is also an open issue. To the best of our knowledge, no study has been proposed on the evaluation of the efficiency of EWA mechanisms with ontology-independent approaches. This paper evaluates existing reification methods for expressing without asserting in RDF with the related contextual information and compares them with a new approach, called *Conjectures*.

The paper is structured as follows: in section 2 we discuss of a number of syntactical

---

[1]http://www.wikidata.org/entity/Q19801150
[2]https://www.w3.org/wiki/RdfReification
[3]http://www.wikidata.org/entity/Q29633776

Figure 1: Concurring attributions concerning *Napoleon crossing the Alps* in Wikidata

approaches to EWA and benchmarks adopted in the literature for their evaluation. In section 3 we briefly introduce and compare conjectures as an alternative method to represent EWA statements in addition to existing proposals such as RDF reification, RDF star, and others. In section 4 we discuss the dataset and the metrics used for our comparison of the various EWA approaches, and in section 5 we analyze our findings. In section 6 we discuss the results before drawing some conclusions in section 7.

2. State of the art

Many Reification methods have been implemented [8, 9, 10, 11, 12, 13] to represent statements about RDF claims. Reification provides additional triples about such claims enabling querying and reasoning mechanisms to integrate the following key attributes[12, 14, 15]: *provenance*: enables the identification and representation of the source of a fact or statement (i.e. *Who claimed this fact?*), *time*: we can communicate the fact's time-related information (i.e. *Which is the most up-to-date claim about this fact?*), *location*: reveals locational information about an event (i.e. *Which is the location in which this claim is applicable?*), *certainty*: indicate the level of confidence that is attributed to a statement (i.e. *how confident are we about a certain event?* or *is a given statement true?*), *versioning*: it is helpful to keep track of RDF datasets' updating history (e.g. *what data version I am using right now?*)

The scientific community has suggested the following major reification strategies to be able to encapsulate all this variety of contextual information about RDF statements.

Among existing Knowledge Graphs, Wikidata provides RDF data with its reification method[11], and a long list of qualifiers about provenance, time, and place to represent complex and competing information. Additionally, each claim validity is provided by a customised approach to

express without asserting [7] such claims (via ranking mechanism[4]) which truth state cannot be given for granted. Consider for example the case of concurring attributions of *Napoleon crossing the Alps*, as shown in figure 1, is represented in Wikidata with three different claims: the currently accepted attribution to Jacques-Luois David (marked with a preferred rank and therefore asserted), the former attribution to the his workshop (marked with a normal rank and non-asserted) and the former attribution to Jérôme-Martin Langlois (marked with a deprecated rank and therefore non-asserted). The two former attributions are therefore marked with contextual information about the claim (e.g. reason for deprecated rank, former attribution) to express additional information on the claim itself.

To be widely accepted, every data representation method must be both effective and efficient. While a method's effectiveness can be formally demonstrated, we must rely on empirical observations to determine its efficiency. Every new method that is suggested typically includes performance data as well. This is accomplished by carrying out the necessary tests to see how the approach performs on performance indicators like the *number of triples*, *query execution time*, *query complexity*, *dataset storage consumption*, *support by existing tools and implementations* and other pertinent metrics. There is a vast number of academic papers that suggest this kind of experimental setup to benchmark the aforementioned reification methods[16, 12, 6, 14, 17, 18, 19]. Additionally, all the benchmarks mentioned above rely on the same four components, namely *datasets, queries, triplestores* and *reification methods*.

3. Conjectures

Conjectures is an approach intended to express RDF statements with three main states of knowledge: undisputed claims, disputed claims or evolving situations, and settled claims. Conjectures is a RDF 1.1 compliant characterization of Named graphs (the weak form), and an extension of RDF Named graphs semantics to distinguish plain Named graphs from disputed (conjectural) and settled (the strong form)[5].

Undisputed claims Non-disputed claims are expressed as asserted Named graphs (and therefore introduced by the keyword GRAPH). For example, the main subject of *Napoleon Crossing the Alps* has been recognised as Napoleon itself without any doubt as shown in listing 2.

Disputed claims Conjectures is a prototypical extension of the syntax of Trig, where the keyword GRAPH is replaced with CONJECTURE in front of a graph whose contents is expressed but not asserted and expressing those statement which is disputes or conveys evolving situations. For example, *Napoleon crossing the Alps* former attribution to Jérôme-Martin Langlois can be represented via a Conjecture as in listing 2. Naturally, graphs that are not marked as conjectures maintain the same (locally-decided) semantics as before, and therefore they may or may not contribute to the truth value of the entire dataset depending on such choice. A key aspect is that Conjectures do not use reification, *n*-ary relationships or *ad hoc* classes For example,

---

[4]https://www.Wikidata.org/wiki/Help:Ranking

[5]A complete overview of Conjectures semantics is available at https://conjectures.altervista.org//CONJ_semantics.pdf

the `rdf:Statement` class, as employed in standard reification and illustrated in Listing 1, and therefore they are orthogonal to, and fully compatible with, most of the other approaches.

```
# The painting main subject is Napoleon
GRAPH s:claim1 {
    wd:Q19801150 wdt:P921 wd:Q517
}

# the painting (wd:Q19801150) creator (wdt:P170) was said to be Jérôme-Martin Langlois (wd:Q672158)
CONJECTURE s:claim2 {
    wd:Q19801150 wdt:P170 wd:Q672158
}

# the painting creator (wdt:P170) was said to be from the school of Jacques-Louis David
CONJECTURE s:claim3 {
    wd:Q19801150 wdt:P170 _:blank
}

# After debate, the painting (wd:Q19801150) creator (wdt:P170) is Jacques-Louis David (wd:Q83155)
SETTLED s:claim4 {
    wd:Q19801150 wdt:P170 wd:Q83155
}
```

Listing 2: Representation of the disputed claim about "Napoleon crossing the Alps" creator with Conjectures

Settled claims Settled claims (introduced by the keyword SETTLED) record both the dispute, *as well as* its subsequent resolution. This is specifically and intentionally different from a trivial re-assertion of the disputed claims, in which we do not acknowledge or mention the dispute at all (the case of GRAPH in 2). To handle settled disputes we introduce *Settled Conjectures*, a third type of named graph that is at the same time conjectured *and* asserted. The collapse graphs allow us to both represent the conjectural triples (inside the usual conjectural graph) as well as the same triples but completely asserted (inside the collapse graph). In addition, the *conj : settles* relation connects the conjecture and its settlement, simplifying the task of exploring the relationships between disputes and their settlements. The rationale behind Settled Conjectures is two-fold: on the one hand, to stress the difference between claims that have not been challenged and claims that emerged as winning among competing and incompatible hypotheses and on the other to represent the dual nature of settled claims as both conjectures and assertions. Consider for example "Napoleon Crossing the Alps" current attribution to Jacques-Louis David can be represented via a Settled conjecture as in listing2.

4. Testing EWA approaches

The methodology adopted to run the tests outlined in this study can be summarized as follows: first, a series of datasets were generated, scaled, and converted in RDF to mimic a set of EWA approaches (see Section 4.1). Next, the hardware and software employed to run our experiments on the collected data has been setup (see Section 4.2). A set of metrics has been defined to evaluate the efficiency of EWA approaches (see Section 4.3). Lastly, the outcomes of the tests are presented in the next section (see Section 5).

4.1. Data Acquisition, scaling and conversion

The dataset on which these experiments have been run is composed as follows and has been named D3:

- **Art**: A thematic set of claims about 300k artwork entities in Wikidata (i.e., painting, manuscripts, books). This corresponds to about 10% of all artwork entities currently present inside Wikidata.
- **Random**: after considerable deliberation, we concluded that adding some kind of entropy to the dataset would make it more representative. This dataset contains the claims of 300k Wikidata random entities.
- **Dummy**: a selection of dummy statements regarding the artwork attributions (represented by the property `wdt:P50` and `wdt:P170` and including from 1 to 4 authors in each claim and the source of the claim) and artworks locations (represented by the property `wdt:P276`, including 1 possible location, time constraints and source) has been created[6]. Those new statements contain dummy arbitrary information ranked as deprecated and therefore non-asserted to represent alternative or historical claims to those contained in Art dataset. This design choice was made to increase the number of conjectural statements in the final dataset.

An excellent way to evaluate an algorithm's performance is to observe how it responds to variations in input size[6]. We started by downloading the whole subset of artwork entities, related individuals (basically, attributed authors) and locations. This dataset, called D4, is composed of about 3,5 million artwork entities and 188 thousand related entities (humans and locations). We have not used this dataset for our comparison due to the excessive number of timeouts in many of the queries and methods we used. Thus we scaled the dataset logarithmically in three further sizes:

- Dataset D3 : D3 is obtained by extracting one tenth of the data in D4 (D3 = D4/10).
- Dataset D2 : D2 is obtained by extracting one tenth of the data in D3 (D2 = D3/10).
- Dataset D1 : D1 is obtained by extracting one tenth of the data in D2 (D1 = D2/10).

We then surveyed the state of the art regarding reification methods to express without asserting and selected a set of methods for our analysis: Singleton properties [12], Named graphs[10] (using Wikidata rankings to decide whether a triple is asserted or not), Wikidata [11] and the recent RDF-star[13] approach. We converted Wikidata JSON files into the six selected reification methods through automatic scripts. In table 1 we provide some data about our datasets. At the end of this process, we obtained 18 new method-specific datasets. In other words, for each dataset *Dn*, $n \in [1, 3]$, we constructed the following datasets:

Consider the case of *Napoleon crossing the alps* and its concurring attributions. In addition to the statements present in Wikidata, the listings below present an additional dummy statement reporting the historical attribution of the painting to Sophie Chéradame (Q60804575), a statement claimed by the source *ContrivedAttributionsInArtHistory-VP* and never adopted by the

---

[6]The choice of adding the dummy claims is that of non-asserted statements in the Wikidata dump was circa 1%, a low figure for this experiment

name	Serialization	Reification	EWA	# RDF stmts in D3
Dn-Wikidata	Turtle	Wikidata	yes	66,768,937
Dn-rdfStar	Turtle	RDF-star	yes	29,779,850
Dn-conjStrong	TriG	Conjectures - strong form	yes	29,058,944
Dn-nGraphs	TriG	Named graphs	via ranking	28,896,268
Dn-conjWeak	TriG	Conjectures - weak form	yes	29,199,650
Dn-Singleton	Turtle	Singleton properties	yes	55,325,270

Table 1
Datasets created ($n \in [1, 3]$)

scholars (and therefore marked with deprecated rank). An example of these concurring claims is represented in the listing below with a different reification method: Wikidata statements (listing 3), RDF-star (listing 4), Conjectures in strong form (listing 5), Named graphs with the (listing 6), Conjectures in weak form (listing 7) and Singleton properties (listing 8)[7].

```
wd:Q19801150  wdt:P170  wd:Q83155 .
s:Q19801150  s:P170  wd:Q19801150-s1 ;
    ps:P170  wd:Q83155 ;
    wikibase:rank  wikibase:NormalRank .

wd:Q19801150  s:P170  s:Q19801150-s2 ;
    ps:P170  "unknown value" ;
    pq:P1774  wd:Q83155  ;
    pq:P3831  wd:Q4233718 ;
    wikibase:rank  wikibase:DeprecatedRank .

wd:Q19801150  s:P170  s:Q19801150-s3 ;
    ps:P170  wd:Q672158 ;
    wikibase:rank  wikibase:NormalRank .

wd:Q19801150  s:P170  s:Q19801150dummy ;
    ps:P170  wd:Q60804575 ;
    pq:P248  conj:ContrievedAttributionsInArtHistory-VP ;
    wikibase:rank  wikibase:DeprecatedRank .
```

Listing 3: Wikidata statements

```
wd:Q19801150  wdt:P170  wd:Q83155 .
<< wd:Q19801150  wdt:P170  wd:Q83155 >>
    wikibase:rank  wikibase:PreferredRank .

<< wd:Q19801150  wdt:P170  "unknown value" >>
    pq:P1774  wd:Q83155 ;
    pq:P3831  wd:Q4233718 ;
    wikibase:rank  wikibase:NormalRank .

<< wd:Q19801150  wdt:P170  wd:Q672158 >>
    wikibase:rank  wikibase:NormalRank .

<< wd:Q19801150  s:P170  wd:Q60804575 >>
    pq:P248  conj:ContrievedAttributionsInArtHistory-VP ;
    wikibase:rank  wikibase:DeprecatedRank .
```

Listing 4: RDF-star

```
SETTLED  s:Q19801150-s1  {
    wd:Q19801150  wdt:P170  wd:Q83155 .
}
s:Q19801150-s1  wikibase:rank  wikibase:PreferredRank .

CONJ  s:Q19801150-s2  {
    wd:Q19801150  wdt:P170  "unknown value" .
}
s:Q19801150-s2  pq:P1774  wd:Q83155 ;
    pq:P3831  wd:Q4233718 .
    wikibase:rank  wikibase:NormalRank .

CONJ  s:Q19801150-s3  {
    wd:Q19801150  wdt:P170  wd:Q672158 .
}
s:Q19801150-s3  wikibase:rank  wikibase:NormalRank .

CONJ  s:Q19801150dummy  {
    wd:Q19801150  s:P170  wd:Q60804575
}
s:Q19801150dummy  wikibase:rank  wikibase:DeprecatedRank ;
    pq:P248  conj:ContrievedAttributionsInArtHistory-VP ;
```

Listing 5: strong form of Conjectures

```
GRAPH  s:Q19801150-s1  {
    wd:Q19801150  wdt:P170  wd:Q83155 .
}
s:Q19801150-s1  wikibase:rank  wikibase:PreferredRank .

GRAPH  s:Q19801150-s2  {
    wd:Q19801150  wdt:P170  "unknown value" .
}
s:Q19801150-s2  pq:P1774  wd:Q83155  ;
    pq:P3831  wd:Q4233718 .
    wikibase:rank  wikibase:NormalRank .

GRAPH  s:Q19801150-s3  {
    wd:Q19801150  wdt:P170  wd:Q672158 .
}
s:Q19801150-s3  wikibase:rank  wikibase:NormalRank .

GRAPH  s:Q19801150dummy{
    wd:Q19801150  s:P170  wd:Q60804575
}
s:Q19801150dummy  wikibase:rank  wikibase:DeprecatedRank ;
    pq:P248  conj:ContrievedAttributionsInArtHistory-VP ;
```

Listing 6: Named graphs

---

[7]The approach adopted for the data acquisition and conversion is documented at https://github.com/conjectures-rdf/expressing-without-asserting-efficiency-datasets

```
GRAPH s:Q19801150-s1  {
  wd:Q19801150 conj01:P170 wd:Q83155.
  conj01:P170 conj:isAConjecturalFormOf wdt:P170.
}
GRAPH s:collapseOfQ19801150-s1 {
  wd:Q19801150 wdt:P170 wd:Q83155.
  s:collapseOf-s1 conj:collapses s:Q19801150-s1
}
s:Q19801150-s1 wikibase:rank wikibase:PreferredRank.
GRAPH s:Q19801150-s2  {
  wd:Q19801150 conj02:P170 "unknown value".
  conj02:P170 conj:isAConjecturalFormOf wdt:P170.
}
s:Q19801150-s2 pq:P1774 wd:Q83155;
  pq:P3831 wd:Q4233718.
  wikibase:rank wikibase:NormalRank.

GRAPH s:Q19801150-s3  {
  wd:Q19801150 conj03:P170 wd:Q672158.
  conj03:P170 conj:isAConjecturalFormOf wdt:P170.
}
s:Q19801150-s3 wikibase:rank wikibase:NormalRank.

GRAPH s:Q19801150dummy {
  wd:Q19801150 conj04:P170 wd:Q60804575.
  conj03:P170 conj:isAConjecturalFormOf wdt:P170.
}
s:Q19801150dummy wikibase:rank wikibase:DeprecatedRank;
  pq:P248 conj:ContrievedAttributionsInArtHistory-VP;
```

Listing 7: weak form of Conjectures

```
wd:Q19801150 wdt:P170 wd:Q83155.
wd:Q19801150 sng:Q19801150-s1 wd:Q83155.
Q19801150-s1 sng:isASingletonPropertyOf wdt:P170.
s:Q19801150-s1 wikibase:rank wikibase:PreferredRank.

wd:Q19801150 s:Q19801150-s2 "unknown value".
s:Q19801150-s2 sng:isASingletonPropertyOf wdt:P170.
s:Q19801150-s2 pq:P1774 wd:Q83155;
  pq:P3831 wd:Q4233718.
  wikibase:rank wikibase:NormalRank.

wd:Q19801150 s:Q19801150-s3 wd:Q672158.
s:Q19801150-s3 conj:isAConjecturalFormOf wdt:P170.
s:Q19801150-s3 wikibase:rank wikibase:NormalRank.

wd:Q19801150 s:Q19801150dummy wd:Q60804575.
s:Q19801150dummy conj:isAConjecturalFormOf wdt:P170.
s:Q19801150dummy wikibase:rank wikibase:DeprecatedRank;
  pq:P248 conj:ContrievedAttributionsInArtHistory-VP;
```

Listing 8: Singleton properties

4.2. Hardware and software configuration

Tests have been run on a computer with processor Intel Core i5-8259U CPU @ 2.30GHz 2.30 GHz, RAM 32,0 GB, Windows 10 pro 64 bits, 1T hard disk. The TriG and SparQL parsers of our GraphDB engine were modified to parse Conjectures in strong form. Our GraphDB configuration uses 28G Ram allocated to the application,[8][9] and 10G cache size. A repository has been created for each dataset with inferences off, no rule set assigned, predicates list index enabled and (when possible) contexts enabled. All other parameters are left in their default values. Repositories are already running before their performance tests are executed.

4.3. Metrics

We decided to base the comparison of our reification methods on four major metrics. These metrics are well-established when it comes to RDF quantitative analysis. The performance-related features of the reification methods under consideration should all be covered by those criteria, which should also give us a clear picture of the benefits and drawbacks of each method.

- *Total number of triples in endpoint*: This value is particularly interesting since it makes it possible to assess the verbosity of each method.
- *Loading time*: Time consumed by each dataset to be uploaded in the SPARQL endpoint.
- *Dataset weight in triplestore*: The storage size of the dataset after it has been uploaded and stored in the triplestore.

---

[8]https://graphdb.ontotext.com/documentation/10.1/configuring-graphdb-memory.html
[9]https://graphdb.ontotext.com/documentation/10.2/getting-started.html#:~:text=the%20aforementioned%20icon. -,Configuring%20the%20JVM,Contents%2Fapp%2FGraphDB%20Desktop

- *Query execution time*: Response time on a selected set of queries. Each query is executed automatically ten times. The average value is then computed.

4.3.1. Queries

Two sets of SPARQL queries (GQn, FQn) have been designed. While GQn queries do not include any filter, FQn queries restrict the results only to paintings (Q3305213). Each query set is composed of 6 queries assessing all possible statuses of claims validity. In particular, the queries retrieve the following topics: valid claims (Q1), debated claims (Q2), debated claims with their provenance/time (Q3), currently disputed claims (Q4), accepted claims after being debated (Q5), undisputed claims (Q6). Considering that authors' and locations' attributions provide a simple, yet effective use case to test RDF representation of EWA over our dataset, GQn and FQn have been then customised on retrieving authorship attributions (GQn-P170 and FQn-P170) and artworks' locations (GQn-P276 and FQn-P276) respectively by the use of Wikidata properties P170 and P276. All queries return the same data and are all correct in their respective datasets. Each query set has been automatically run 10 times and the average times have been calculated. Table 4.3.1 summarizes the nature of the queries. All actual queries are available at https://github.com/conjectures-rdf/expressing-without-asserting-efficiency-tests as well as the full set of results.

Query	Predicate	Data selected by query
GQ1	P170	All attributions of artworks that are currently considered valid
GQ1	P276	All locations of artworks that are currently considered valid
GQ2	P170	All attributions of artworks that have been debated
GQ2	P276	All past and debated locations of artworks
GQ3	P170	All attributions of artworks that have been debated, *with provenance*
GQ3	P276	All past and debated locations of artworks, *with date of move*
GQ4	P170	All currently debated attributions of artworks
GQ4	P276	All locations of artworks whose current location is uncertain
GQ5	P170	All settled attributions of artworks
GQ5	P276	All current locations of artworks that were moved
GQ6	P170	All attributions of artworks that were never debated
GQ6	P276	All locations of artworks that never moved
FQ1	P170	All attributions of paintings (Q3305213) that currently are considered valid
FQ1	P276	All locations of paintings (Q3305213) that are currently considered valid
FQ2	P170	All attributions of paintings (Q3305213) that have been debated
FQ2	P276	All past and debated locations of paintings (Q3305213)
FQ3	P170	All attributions of paintings (Q3305213) that have been debated, *with provenance*
FQ3	P276	All past and debated locations of paintings (Q3305213), *with date of move*
FQ4	P170	All currently debated attributions of paintings (Q3305213)
FQ4	P276	All locations of paintings (Q3305213) whose current location is uncertain
FQ5	P170	All settled attributions of paintings (Q3305213)
FQ5	P276	All current locations of paintings (Q3305213) that were moved
FQ6	P170	All attributions of paintings (Q3305213) that were never debated
FQ6	P276	All locations of paintings (Q3305213) that never moved

Table 2
Overview of query types (GQn and FQn) for artworks' attributions and locations
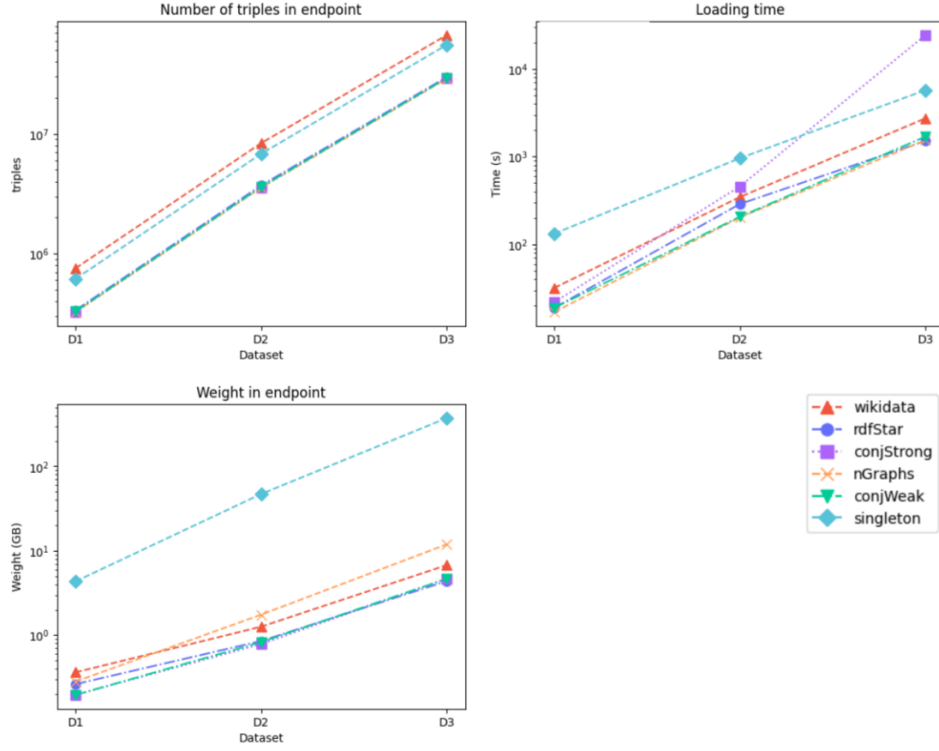
Figure 2: Dataset measurements for the surveyed reification methods. In particular, number of triples in endpoint (top-left), dataset loading time (top-right) and dataset weight in endpoint (bottom-left)

5. Test results

5.1. Number of triples in endpoint

All reification methods add additional triples to the already existing ones to represent the necessary metadata (e.g. Singleton properties) or extend RDF 1.1 syntax (e.g. RDF-star). As shown in table 2, Named graphs are the method which uses reification with the lower number of triples, but with no explicit distinction between asserted and non-asserted graphs. While other surveyed methods (in particular, RDF-star, Wikidata statements and Singleton properties) use reification methods and assert each claim with an additional triple, Conjectures uses Named graphs structure to express both statements and reification without adding additional triples resulting in the method to express without asserting with the lowest addition of triples.

5.2. Loading Time

In the context of dataset D1 and D2, Conjectures in the strong form remain competitive with the most efficient methods, notably RDF-star, and outperform Wikidata statements and Singleton properties as shown in 2. However, the loading times increase in D3. This performance discrepancy is attributed to the triplestore's parser method for recognizing conjectural data. Specifically,

the process of checking each resource's presence in a collection during loading contributes to the observed delays. In essence, the loading time of the dataset increases proportionally with the quantity of non-asserted triples (conjectures).

5.3. Dataset weights in triplestore

The Singleton method exhibits a storage size tenfold greater than alternative approaches, with Conjectures in their weak form and Wikidata occupying intermediate positions. RDF-star, Conjectures in their strong form, and Named graphs demonstrate similar sizes as shown in figure 2.

5.4. Query Execution Time

The time response average for each dataset seems to increase linearly for each surveyed dataset (D1, D2, D3). For this reason, figure 3 provides the snapshot of the execution time of queries GQn and FQn on attributions and locations only on dataset D3.

As illustrated in figure 3, the response times obtained from the execution of general queries (GQn) on dataset D3 show that the strong form of Conjectures is less efficient than other methods when retrieving asserted data, particularly in retrieving valid claims, currently disputed claims and undisputed claims (queries GQ1, GQ4 and GQ6) for both creators and locations. However, Conjectures perform better with disputed claims. In particular, Weak and Strong Conjectures outperform other surveyed methods in the retrieval of debated statements with and without provenance information (queries GQ2 and GQ3) and accepted claims after being debated (GQ5) for both locations and creators.

Similar to what was observed in GQn, Conjectures are less efficient in retrieving valid claims (FQ1). On the contrary, Conjectures strong form is the most efficient method regarding the retrieval of debated claims with and without contextual information (queries FQ2 and FQ3) and accepted claims after debate (FQ5). In the remaining queries, currently disputed claims (FQ4) and statements that have never been subject to debate (FQ6), Conjectures still maintain competitive times with the rest of the methods. Strong Conjectures, in particular, address the significant increase in response times for weak-form queries FQn[3:8]. Essentially, a notable improvement in the performance of the strong form has been detected, proving to be the most efficient method in half of the selected queries and, in the remaining ones, a valid competitor.

6. Discussion

6.1. EWA expressivity

In query performance assessment, we registered some slightly different values for the number of results of some queries, such differences are explained below highlighting some intrinsic differences in the models. Typically, a SPARQL query can still retrieve claims that won a certain debate by accessing concurring ones stored in the Knowledge Graph (KG). In cases where debates are not recorded, and only accepted statements are reported, reification approaches fall short. Reification methods do not distinguish between claims that have never been questioned and
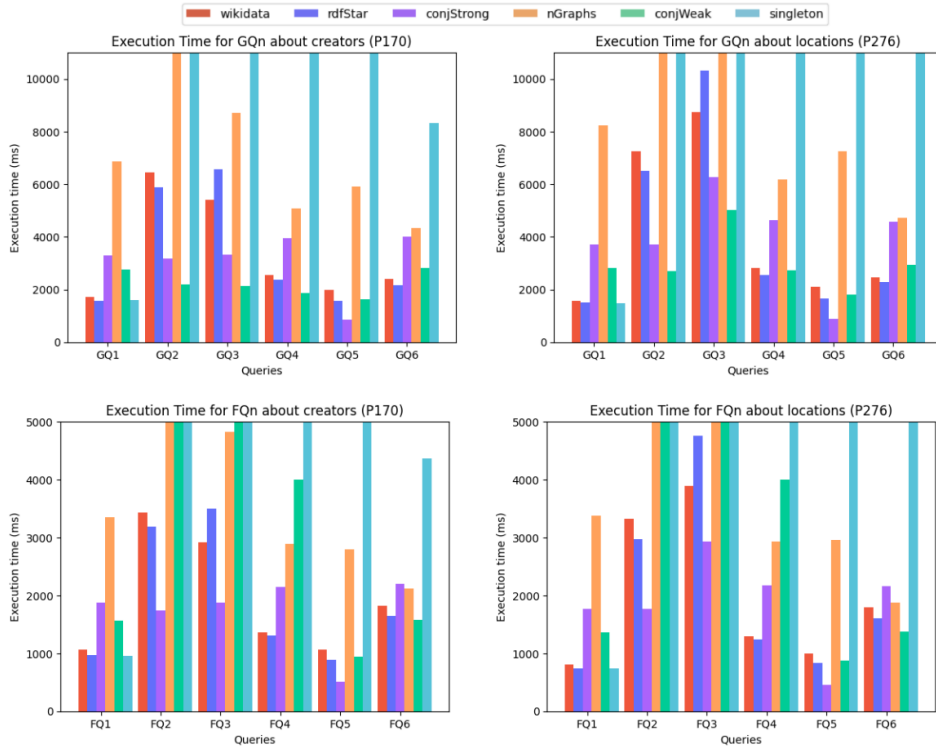
Figure 3: Time responses for queries sets GQn and FQn on creators and locations run against D3 dataset

settled claims post-debate since both are recorded as asserted triples. At this point, an ontology-dependent solution must be addressed to represent this differentiation such as Wikibase ranking. For instance, the painting *Portrait of Dona Isabel de Requesens* (Q29651096) has been attributed to Giulio Romano and marked as the settled claim, but no concurrent claim is reported. The concept of SETTLED in Conjectures can also capture this nuanced distinction in an ontology-independent fashion which is transversal to be used on whichever KG.

Another instance is when two claims express the same triple but with different qualifiers. Consider a scenario where a historical painting X was initially attributed to author Y (marked as an attribution) and afterwards this attribution has been considered settled by the community. The historical attribution cannot be retrieved in a SPARQL query since its content is also asserted. But, while Wikidata and Singleton provide unique IDs to distinguish such claims which can be involved in the query to retrieve such claim, RDF-star associates all contextual triples with the same quoted triple. This becomes more complex in multi-triple claims, where expressing them as individual Wikidata statements is not feasible. For instance, when dealing with paintings attributed to the collaboration of multiple individuals, expressing this complexity becomes particularly interesting. Conjectures use Named graphs to group statements. which allows to express and retrieve such complex statements with simple SPARQL queries. Other methods would require to adopt other types of grouping methods (e.g. a n-ary relationship and/or a

blank node) with additional complexity and execution times.

6.2. EWA efficiency

Overall, we can immediately see several trends concerning the surveyed methods' efficiency: Singleton properties are systematically slower than the others, while Named graphs and Conjectures in weak form performs at an intermediate level about the fastest methods, Wikidata, Conjectures in strong form and RDF-star. Strong form also outperforms RDF-star in many queries where the specifics of debated attributions and past locations become meaningful. Strong form is the quickest method for expressing debates (disputed claims, GQ2 and FQ2, disputed claims with provenance GQ3 and FQ3, settled claims GQ5 and FQ5), with a small loss in term of performance for what concern asserted claims (valid claims GQn and FQn, undisputed claims GQ6 and FQ6) and currently disputed claims (GQ5 and FQ5).

Conjectures in strong form are also competitive for what concerns number of triples and overall weight in the triplestore. It is competitive on loading time for, but loading times show an interesting loss of performance for large datasets, that need to be investigated further.

7. Conclusions

This work evaluates the efficiency of EWA mechanisms by comparing several reification methods (Wikidata, RDF-star, Named graphs, Singleton properties) and the novel Conjectures approach (weak and strong form) on four major metrics (number of triples, loading time, dataset weight and query execution time). Among the most efficient methods as RDF-star and Wikidata statements, the strong form of Conjectures exhibits notable performance gains, particularly in retrieving claims about debates (e.g., disputed claims with and without provenance information and settled claims). In the future, we aim to optimize some aspects concerning the efficiency of the strong form of Conjectures. In particular, we will prioritize the optimization of the loading process, aiming to reduce the loading time and enhance its overall performance.

References

[1] J. J. C. Graham Klyne, Resource Description Framework (RDF): Concepts and Abstract Syntax (2004). URL: https://www.w3.org/TR/2004/REC-rdf-concepts-20040210/.

[2] M. Giunti, G. Sergioli, G. Vivanet, S. Pinna, Representing n-ary relations in the semantic web, Logic Journal of IGPL (2019). doi:10.1093/jigpal/jzz047.

[3] A. Freundschuh, Crime stories in the historical urban landscape: narrating the theft of the mona lisa, Urban History 33 (2006) 274–292. doi:10.1017/S0963926806003816.

[4] B. Lewis, Salvator mundi: Going out on a limb, KUR - Kunst und Recht 25 (2023). URL: https://doi.org/10.15542/KUR/2023/2/3. doi:10.15542/KUR/2023/2/3.

[5] P. Hayes, J. Carroll, C. Welty, M. Uschold, B. Vatant, F. Manola, I. Herman, J. Lawrence, Defining N-ary Relations on the Semantic Web (2006). URL: https://www.w3.org/TR/2006/NOTE-swbp-n-aryRelations-20060412/.

[6] F. Orlandi, D. Graux, D. O'Sullivan, Benchmarking rdf metadata representations: Reification, singleton property and rdf*, 2021 IEEE 15th International Conference on Semantic Computing (ICSC) (2021) 233–240. URL: https://api.semanticscholar.org/CorpusID:232151947.

[7] M. Daquino, V. Pasqual, F. Tomasi, F. Vitali, Expressing without asserting in the arts, in: Proceedings of the Italian Research Conference on Digital Libraries. Padova, Italy, 2022.

[8] P. Hayes, Rdf semantics, w3c recommendation, http://www. w3. org/TR/rdf-mt/ (2004).

[9] N. Noy, A. Rector, P. Hayes, C. Welty, Defining n-ary relations on the semantic web, W3C working group note 12 (2006).

[10] J. J. Carroll, C. Bizer, P. Hayes, P. Stickler, Named graphs, Journal of Web Semantics 3 (2005) 247–267.

[11] F. Erxleben, M. Günther, M. Krötzsch, J. Mendez, D. Vrandečić, Introducing Wikidata to the Linked Data Web, in: P. Mika, T. Tudorache, A. Bernstein, C. Welty, C. Knoblock, D. Vrandečić, P. Groth, N. Noy, K. Janowicz, C. Goble (Eds.), The Semantic Web – ISWC 2014, Springer International Publishing, Cham, 2014, pp. 50–65.

[12] V. Nguyen, O. Bodenreider, A. Sheth, Don't like rdf reification? making statements about statements using singleton property, in: Proceedings of the 23rd International Conference on World Wide Web, WWW '14, Association for Computing Machinery, New York, NY, USA, 2014, p. 759–770. URL: https://doi.org/10.1145/2566486.2567973. doi:10.1145/2566486.2567973.

[13] O. Hartig, Foundations of rdf* and sparql*:(an alternative approach to statement-level metadata in rdf), in: AMW 2017 11th Alberto Mendelzon International Workshop on Foundations of Data Management and the Web, Montevideo, Uruguay, June 7-9, 2017., volume 1912, Juan Reutter, Divesh Srivastava, 2017.

[14] A.-C. Ngonga Ngomo, I. Fundulaki, A. Krithara, J. Frey, K. Müller, S. Hellmann, E. Rahm, M.-E. Vidal, A.-C. Ngonga Ngomo, I. Fundulaki, A. Krithara, Evaluation of metadata representations in rdf stores, Semant. Web 10 (2019) 205–229. URL: https://doi.org/10.3233/SW-180307. doi:10.3233/SW-180307.

[15] V. Nguyen, O. Bodenreider, K. Thirunarayan, G. Fu, E. Bolton, N. Queralt-Rosinach, L. I. Furlong, M. Dumontier, A. Sheth, On reasoning with rdf statements about statements using singleton property triples (2015).

[16] D. Hernández, A. Hogan, M. Krötzsch, Reifying RDF: what works well with wikidata?, in: T. Liebig, A. Fokoue (Eds.), Proceedings of the 11th International Workshop on Scalable Semantic Web Knowledge Base Systems, volume 1457 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2015, pp. 32–47.

[17] A. Hogan, J. Umbrich, A. Harth, R. Cyganiak, A. Polleres, S. Decker, An empirical survey of linked data conformance, J. Web Semant. 14 (2012) 14–44. doi:http://dx.doi.org/10.2139/ssrn.3198962.

[18] G. Demartini, I. Enchev, M. Wylot, J. Gapany, P. Cudré-Mauroux, BowlognaBench—Benchmarking RDF Analytics, in: K. Aberer, E. Damiani, T. Dillon (Eds.), Data-Driven Process Discovery and Analysis, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 82–102.

[19] Y. Theoharis, V. Christophides, G. Karvounarakis, Benchmarking Database Representations of RDF/S Stores, in: Y. Gil, E. Motta, V. R. Benjamins, M. A. Musen (Eds.), The Semantic Web – ISWC 2005, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005, pp. 685–701.