On assessing weaker logical status claims in Wikidata cultural heritage records

Alessio Di Pasquale [a], Valentina Pasqual [b,*], Francesca Tomasi [b] and Fabio Vitali [a]

[a] *Department of Computer Science, University of Bologna, Italy*
E-mails: alessio.dipasquale@studio.unibo.it, fabiovitali@unibo.it
[b] *Department of Italian Studies and Classical Philology, University of Bologna, Italy*
E-mails: valentina.pasqual2@unibo.it, francescatomasi@unibo.it

Abstract. This work presents an analysis of the use of different representation methods in Wikidata to encode information with weaker logical status (WLS, e.g. uncertain information, competing hypothesis, temporally evolving information, etc.). The study examines four main approaches: non-asserted statements, ranked statements, null-valued objects, and statements qualified with properties P5102 (reason of statement), P1480 (sourcing circumstances) and P2241 (reason for deprecated rank). We analyse their prevalence, success, and clarity in Wikidata. The analysis is performed over cultural heritage artefacts stored in Wikidata divided in three subsets (i.e. visual heritage, textual heritage and audio-visual heritage) and compared with astronomical data (stars and galaxies entities). Our findings indicate that (1) the representation of weaker logical status information is limited, with only a small proportion of items reporting such information, (2) the representation of WLS varies significantly between the two datasets, and (3) precise assessment of WLS statements is made complicated by the ambiguities and overlappings between WLS and non-WS claims allowed by the chosen representaions. Finally, we list a few proposals to simplify and standardize the representation of this type of information in Wikidata, with the hope of increasing its accuracy and richness.

Keywords: wikidata, ranked statements, weaker logical status, uncertainty, cultural heritage

1. Introduction

Since 2012 Wikidata [1] is one of the most outstanding platforms to collect and share Linked Open Data trough the web.

Through the years Wikidata has developed and provided a variety of representation methods that allow to encode complex structures much beyond factual descriptive metadata. According to [2], Wikidata encompasses a multitude of facts, including some that may be contrasting since coming by different and disagreeing sources. Additionally, time-sensitive information can also be added through the use of qualifiers and ranks. For instance, structures to represent temporally evolving information (e.g., the number of followers of a Youtube Channel that is updated year after year) or multiple coexisting (and possibly competing) claims over the same subject (e.g., maintaining both the old as well as a new theory over some topic). In many such cases, multiple information items are present, yet newer or better information is not replacing older or less true assertions, but they coexists next to each other, and one or more mechanisms are used to signal their simultaneous presence, and, when appropriate, the currently adopted stance.

*Corresponding author. E-mail: valentina.pasqual2@unibo.it.

We understand these enunciates as enjoying a somehow *weaker logical status* than simply asserted statements: they are neither absolutely true nor absolutely false, but they are, e.g., true from a specific moment onward but not earlier, or true up to a given moment but not afterwards, or accepted as true by most people but not everybody, etc.

It is a cultural necessity in many (if not all) fields of knowledge to have access to available data about a complex topic in a complete and objective manner, as they evolve over time, as they are interpreted by different scholars or models, as they represent available hypothesis rather than a positive certainty. For instance, cultural heritage scholars study attributions, the temporal context of events, the temporal evolution of content, the contradictions of opinions and assertions, so that expressing weak statements, i.e., claims we are not certain about, becomes a necessary tool to increase precise awareness of the currently available data for those who consult or reuse it. Interpretation thus plays a central role in humanities disciplines, yet cultural heritage knowledge graphs and domain ontologies frequently limits the formalisation of these phenomena or only partially represent them ([3, 4], cf. section 2). Recently, a rekindled interest is shown in the formalisation of uncertain statements [5–7], claiming that interpretation constitutes a focal point in humanities data and metadata.

Wikidata supports a number of patterns to represent situations best expressed with weaker logical status claims. In this paper we analyse some of these patterns as they are employed in actual collections, both in the humanities and, as a comparison, in hard sciences. A factor that increases complexity is that many of these approaches have partially overlapping semantics, i.e., they can be used also for other purposes beyond weaker logical status claims, and this muddles the correct identification and interpretation of the situations we are interested in. We therefore want to discuss both the expected application of each approach, its relative success, as well as the impact of their ambiguous applications due to the coexistence of multiple uses for the same techniques.

In particular, we analysed four main families of approaches to weaker logical status of statements, *asserted vs. non-asserted statements*, *ranked statements*, *null-valued objects* and *qualified statements*. In this paper we try to give an answer to the following research questions:

- (RQ1) How widespread and successful is each of these approaches in the current state of Wikidata?
- (RQ2) How does the cultural domain of the Wikidata topics (and, presumably, of the individuals contributing to the data regarding the entities) affect and reflect on the relative success and richness of some approaches over the others?
- (RQ3) How clean and easy to differentiate are the applications of each approach to an actual weaker logical status versus to another of the designed uses of that approach?
- (RQ4) Is there a way to improve the clarity and cleanliness of such differentiation?

In order to perform such analysis, we accessed and downloaded two large sets of topics from Wikidata, one belonging to the cultural heritage (visual works of art such as painting and statues, text documents and audio-visual entities), and another from astronomy (celestial bodies such as stars and galaxies). Both make some use of multiple fuzzy assertions and hypothesis, and therefore are in need of assertions with weaker status (e.g., attributions uncertainties or physical locations moving over time for paintings, vs. spectral class or radial velocity for stars).

Overall our findings show that the amount of weaker logical status statements in Wikidata seems suspiciously low, as only 0,4% of visual artworks report attribution debates, a fairly low figures compared to, e.g., a more reasonable 8,5% coming from the RKD images collection[1], a difference that could be attributed to the difficulty and ambiguities in the procedures to report such complex information. We propose here also a way to simplify, streamline and homogenize such complexity, with the hope of increasing the abundance, the richness and the correctness of the representation of such phenomena in Wikidata.

The paper is structured as follows: in the state of the art (2) relevant data sources KGs and data models are presented when representing weaker logical status claims as well as schema and data assessments proposed over Wikidata. In section 3 we present the approaches provided in Wikidata to encode weaker-logical status claims. In section 4 the reserach objective are outlined, the data acquisition process is briefly described and the analysis of our Wikidata sample dataset is presented. In section 5 we present our proposal for improving the quality in annotating weaker-logical status knowledge. Finally, in section 6 we summarize our findings and outline our conclusion about the work.

[1] https://rkd.nl/en/explore/images

2. State of the art

Among public Knowledge Graphs (KGs, e.g. Wikidata [1], DBpedia [8], Yago [9], Google Knowledge Graph) we find a number of collaborative public platforms, which are built and maintained by a community of contributors, and constitute publicly available KGs that can be used for research, either expressing specialist knowledge or general knowledge - such as Wikidata.

Weaker logical status statements are a natural occurrence in many contexts covered by these KGs, but the support for their representations varies considerably. Guidelines, data modelling and data harmonization (a particularly relevant need for open platforms) can help in expressing them, i.e. for concurrent opinions or uncertain claims. In the field of cultural heritage studies, the competition of knowledge is intriguing. However, some online databases or data models only partially address this issue.

Despite domain ontologies representing the cultural heritage domain hardly manage to integrate support for interpretation (i.e., hermeneutics) into their models[5], there are some exceptions [4, 10].

CIDOC CRM[4] is a conceptual model widely adopted by many knowledge graphs [11, 12] in the cultural heritage domain. It offers a formal approach to express weaker logical status claims through the use of n-ary relations (e.g., `"crm:E13_AttributeAssignment"`).

Europeana [10], stores approximately 50 million heterogeneous digitized items from museums, libraries, and archives across Europe. Data is collected by content providers (i.e. cultural institutions) using the EDM data model [3] and the use of proxies [13] allows to express conflicting information and track data provenance. However, concurrent statements are not visible on the online pages, and no mechanism is in place to determine which proxy will be made visible when multiple exist.

Another instance of an EDM collection is the RKD catalogue, a comprehensive collection of data about Dutch works of art throughout history. RKD gathers and represents many contested and discarded attributions of paintings and portraits. Although at the moment there is no SPARQL endpoint available for querying the collections, the data can be explored through an online catalog using traditional relational query patterns. Interestingly, about 30,000 artwork descriptions from RKD have been imported into Wikidata, representing ~3% of the total of its visual artworks.

Despite the support of representational definitions of weaker logical status claims in EDM, CIDOC-CRM and RDK data model, these weaker forms of information are often poorly reported (*reticence*) or are expressed in textual annotations rather than being modeled in the data structure (*dumping*)[14].

The widespread adoption of Wikidata within the cultural heritage community has been well-documented [15][2]. Wikidata is seen not only as a valuable tool for data publishing, alignment and enrichment, but also as a means of gaining valuable insights into cultural heritage data and the community itself [16]. Given the significance of comprehensive data in knowledge bases, there has been a focus on improving and evaluating their schema and data quality [17]. Improving the representation of complex knowledge has been tackled e.g., by [18], who compared the efficiency of several reification methods (e.g., singleton properties, n-ary relations, named graphs and standard reification) on Wikidata data. Weaker logical status claims may make good use of reification approaches.

Additionally, the representation of complex of data scenarios in knowledge bases often needs to be evaluated according multiple metrics. For instance, [19] survey quality metrics from 28 scientific publications on the topic and categorizes quality assessments into three dimensions: intrinsic (accuracy, trustworthiness, consistency), context (relevance, completeness and timeliness) and representation (ease of understanding and interoperability). Among quality measures, evaluation of completeness, defined in [20] as the "presence of all required information in a given dataset", has been approached through various methods and assessments as comparing data for similar entities [21], measuring entity relatedness [22], evaluating thoroughness of information by determining the completeness of specific attributes of objects [23], assessing low-quality statements thought the analysis of items' discussion pages, deprecated statements and constraint violations [24], and assessing and comparing data quality across large knowledge bases [20, 25].

[2] See also the list of cultural institutions involved in Wikidata can be found at https://www.wikidata.org/wiki/Wikidata:GLAM

Overall, little or no evaluation has been conducted on the representation of weaker logical status claims in Wikidata, nor has a comprehensive analysis been carried out to assess the amount knowledge related to WLS status in the field of cultural heritage. In the next section we detail our own proposal to address these shortcomings.

3. Representing weaker logical statuses in Wikidata

Wikidata represents weaker logical status statements (e.g. for uncertain or debated assertions) using at least three different representation methods: ranked statements (section 3.2), statements with specific qualifiers (section 3.3) and statements with a null-valued object (section 3.4).

3.1. Asserted vs. non-asserted statements

All claims in Wikidata are expressed through *statements*, a custom reification method[3] to express contextual information (e.g. qualifiers, rankings, references) about the claim itself. Statements connect the claim's subject and the claim's predicate to a Statement entity which refers to the claim's object and can be further used as subject of other triples. Statements, therefore, do not actually assert the corresponding claim. To do so it is necessary to also add a triple that (using a different prefix) relates the claim's subject to the claim's object through the claim's predicate, thus enabling simple query support for asserted facts. The separation between Statements and their assertion is optional, which allows to easily support claims presented as facts (we can find both the Statement and the assertion triple) as well as claims not meant to be considered facts (the Statement is there, but no assertion triple is added). This approach works well with the ranking of assertions, (see 3.2) as modelled vy the Wikibase data model[4]. As shown in listing 1, a retracted attribution of the painting "Madonna with the Blue Diadem" (`Q738038`) to Raphael is represented only by a Statement (`wd:Q738038 p:P170 s:Q738038-7729b786-4d4f-a0ca-2ded-4ea2c6307e1c`), while the accepted attribution to Gianfranco Penni is represented with both a Statement, as before (`wd:Q738038 p:P170 s:Q738038-7729b786-4d4f-a0ca-2ded-4ea2c6307e1c`) and also a plainly asserting triple directly connecting the painting to Penni (`wd:Q738038 wdt:P170 wd:Q2327761`). The different prefixes of the predicate P170 allow to differentiate a Statement from a plainly asserted triple.

3.2. Ranked statements

Competing claims over some data are represented via a ranking mechanism (e.g., *Preferred*, *Normal* and *Deprecated*). Rankings [26] communicate the consensus opinion for a statement as reached by the scientific community or Wikidata annotators. Disputes are separately hosted in the corresponding discussion page, in plain text. Many possible combinations of variously ranked competing statements can be found in the Wikidata collection, with various and debatable interpretations, but clearly Preferred statements are meant to be chosen for claims with a stronger status, and Deprecated for weaker ones: thus Statements ranked as Preferred are asserted (the assertion triple is added), while Statements ranked as Deprecated are not (no assertion triple is added). The interpretation of Normal statements varies depending on whether they coexist or not with competing Preferred and/or Deprecated claims, and similarly may vary the presence or absence of assertion triples. For example, in listing 1 the attribution to Gianfrancesco Penni enjoys both a Preferred rank as well as an assertion triple, while the attribution to Raphael is ranked as Normal and has no assertion triple. Even though the first attribution is ranked Normal rather than Deprecated, we must consider it as a superseded claim.

```
# attribution to Raphael
wd:Q738038 p:P170 s:q738038-121B92D0-E6E1-4514-960C-AE34F50054E5 .
s:q738038-121B92D0-E6E1-4514-960C-AE34F50054E5 a wikibase:Statement ;
    wikibase:rank wikibase:NormalRank ;
```

[3] https://www.wikidata.org/wiki/Help:Statements
[4] https://www.mediawiki.org/wiki/Wikibase/DataModel#Statements

```
    ps:P170 wd:Q5597 .              # creator: Raphael

# attribution to Gianfrancesco Penni
wd:Q738038 wdt:P170 wd:Q2327761 .    # creator: Gianfrancesco Penni (assertion)
wd:Q738038 p:P170 s:Q738038-7729b786-4d4f-a0ca-2ded-4ea2c6307e1c .
s:Q738038-7729b786-4d4f-a0ca-2ded-4ea2c6307e1c a wikibase:Statement;
    wikibase:rank wikibase:PreferredRank ;
    ps:P170 wd:Q2327761.            # creator: Gianfrancesco Penni
```

Listing 1: Preferred and Normal ranks

3.3. Qualifiers

Statements, independently of rank, can be decorated with additional triples annotating the nature of the statement by using predicate `P5102` (*reason of statement*[5]) and/or `P1480` (*sourcing circumstances*[6]), providing contextual information regarding the claim.

For example, in listing 2 we see that the painting "Abstract Speed + Sound" (`Q19882431`) by Giacomo Balla is described as *possibly* part of a triptych. The use of a qualifier with a Normal ranking seems to imply that the statement is considered true and therefore it is also asserted.

```
wd:Q19882431 wdt:P361 wd:Q79218 .     # part of: triptych (assertion)
wd:Q19882431 p:P361 s:Q19882431-1ac26ff2-4981-ff79-4fae-9d411ae34296 .
s:Q19882431-1ac26ff2-4981-ff79-4fae-9d411ae34296 a wikibase:Statement;
    wikibase:rank wikibase:NormalRank ;
    ps:P361 wd:Q79218 ;               # part of: triptych
    pq:P5102 wd:Q30230067 .           # circumstance: possibly
```

Listing 2: A qualified statement in Wikidata

Additionally, the properties `P2241` (*reason for deprecated rank*[7]) and `P7451` (*reason for preferred rank*[8]) are provided to annotate contextual information about superseded and preferred claims, respectively.

Wikidata provides a list of 96 recommended values for `P1502` and 83 recommended values for `P1480` in their respective *Property Talk* pages, while no list of recommended terms is provided for `P2241` nor for `P7452`. Even at first glance it is possible to notice a very wide range of types and specificities (e.g., qualifiers such as *possibly*, *presumably*, and *probably* versus, say, *prosopographical phantom*, *project management estimation* or *archive footage*), and many are not connected to weaker logical status assessments. In addition, semantic overlaps can be noticed on many of these terms, e.g. between *allegation* and *allegedly*, or between *hypothesis*, *hypothetical entity*, *hypothetically* and *scientific hypothesis*. These overlaps support arbitrariness of choice for contributors, increasing the ambiguity of the resulting annotation.

3.4. Null values

Wikidata statements can be associated with a blank node. This is meant to imply that the statement is associated with an unknown value, rather than a missing statement. For example, the "Missal for the use of the ecclesiastics of Clermont' (`Q113302686`), an illuminated manuscript from the 14th century, has been recorded with both an unknown creator and unknown author, as shown in listing 3.

```
wd:Q113302686 wdt:P50 _:4c60f23d697d2d89d9fe49824c8f3a01 .   # author: unknown (blank node - asserted)
wd:Q113302686 p:P50 s:Q113302686-032e3cc5-4fd6-1f20-8830-0909945ba683 .
s:Q113302686-032e3cc5-4fd6-1f20-8830-0909945ba683 a wikibase:Statement;
    wikibase:rank wikibase:NormalRank ;
```

[5] the underlying circumstances of this statement, see https://www.wikidata.org/wiki/Property_talk:P5102
[6] a qualification of the truth or accuracy of a source, see https://www.wikidata.org/wiki/Property_talk:P1480
[7] to indicate why a particular statement should have deprecated rank, see https://www.wikidata.org/wiki/Property_talk:P2241
[8] to indicate why a particular statement should have preferred rank, see https://www.wikidata.org/wiki/Property_talk:P7451

```
    ps:P50 _:f8c6b698b13ef3dd3738e025df3a2d5d .              # author: unknown (blank node)

wd:Q113302686 wdt:P170 _:759d5c5c7a58a8a286512c257514463a .  # creator: unknown (blank node - asserted)
wd:Q113302686 p:P170 s:Q113302686-8d47e883-4566-bc8b-cd8f-6cffebc5414c .
s:Q113302686-8d47e883-4566-bc8b-cd8f-6cffebc5414c a wikibase:Statement;
    wikibase:rank wikibase:NormalRank ;
    ps:P170 _:28d04a432a3589d30a5c6da79d3fac50 .             # creator: unknown (blank node)
```

Listing 3: Null-valued statement in Wikidata

Null values are used to express a specific nuance, e.g. to distinguish between explicit representation of ignorance (a null value) versus simply disregarding a property (the absence of a triple of the relevant predicate). Yet, we see numerous problems in how null values are used for many other purposes in addition to that. For instance, null values are used in some predicates to represent values that cannot exist, e.g. when signaling the start (`P155`: *follows* + null value) of the end (`P156`: *followed by* + null value) in sequences. The overabundance of applications of this method increases the chances of misuses by contributors or of ambiguities in queries.

3.5. Discussion

Even before checking on the actual usage patterns of these methods, we can immediately notice the richness of annotations made possible by them, the subtle nuances they afford, but at the same time the variety of (potential) sources of ambiguities, overlapping connotations and representation vagueness. In particular, we can summarise three specific problems that are worth further discussions:

1. Although the separate uses of Normal, Preferred and Deprecated rankings are clear and practical, there are uncertainties when they coexist on the same predicate, especially for the different representation of Normal statements when Preferred are also present, or when all three rankings are present.
2. The sheer number of qualifiers, the differing level of their respective specificities, the manifest semantic overlapping of many of them makes it quite hard to guarantee homogeneity and precision in their use.
3. The subtlety in the semantic differences between providing no value and providing a null value for a property of a wikidata item, as well as their other types of applications makes the use of null values particularly complicated and ambiguous.

Yet, all these reflections are somehow empty and pointless unless we examine how contributors are actually using these methods for expressing real weak logical status claims in their wikidata contributions. This topic is covered in the next section.

4. Usage patterns of WLG in Wikidata datasets

In order to generate some analysis about the actual usage of WLS claims, and to provide an initial answer to our research questions, we collected two datasets of wikidata items, one about Cultural Heritage items (visual arts, text documents and audio-visual entities) and another about Astronomical objects (galaxies and stars). The datasets were selected so as to be approximately comparable in size and in number of individual statements, and under evidences that both types of entities rely on weaker logical status claims when entities undergo re-evaluations due to new evidences or the recording of different opinions.

4.1. Data Acquisition

```
SELECT DISTINCT ?artwork ?type
WHERE {
    ?artwork wdt:P31 ?type.
    ?type (wdt:P279*) wd:Q838948.
    hint:Prior hint:rangeSafe true
    }
```

Listing 4: SPARQL query retrieving Wikidata entities to subclasses of work of art (`Q838948`)

We first created a dataset of Cultural Heritage items. All Wikidata entities belonging to the class *work of art* (`Q838948`) or any of its sub-classes were collected using a SPARQL query (listing 4). This cultural heritage dataset has been semi-automatically divided into three sub-datasets, due to the wide diversity of cultural properties and their associated claims:

- *Audio-Visual heritage* (CHav): This collection holds information about audio-visual materials that have cultural, historical, or artistic value. They include movies, videos, recordings of music or spoken words, and other audio-visual materials that provide a record of a particular event in a specific time or place. The dataset contains 1.251.626 entities and 17.141.394 statements organized in 25.033 json files.
- *Visual heritage* (CHv): This collection holds information about visual artifacts that have cultural, historical, or artistic value. They include paintings, drawings, sculptures, photographs, decorative arts, etc. The dataset contains 1.078.855 entities and 12.850.825 statements organized in 21.579 json files.
- *Textual heritage* (CHt): This collection holds information about written and printed materials that have historical or cultural significance. They include books, manuscripts, letters, and other written documents. The dataset contains 625.110 entities and 4.584.444 statements organized in 12.503 json files[9].

We also downloaded Wikidata entities of architecture-related classes; they were later discarded due to their fairly lower number as well as for the presence of many statistical ambiguities that could make their evaluation useless (e.g., many entities belonging to these classes should not be considered relevant to cultural heritage collections).

In order to verify our assumptions with a diverse datasets of similar size, we acquired an additional collection of astronomical entities, organized in two datasets:

- *Stars* (As): This collection holds a random selection of 1.199.950 Wikidata entities (of the ~3.3 million existing) belonging to the class `Q523` - *Star*, The dataset contains 27.470.140 statements in 23.999 json files[10].
- *Galaxies* (Ag): This collection holds a random selection of 1.200.000 Wikidata entities (of the ~2 million existing) belonging to the class `Q318` - *Galaxy*, The dataset contains 14.439.421 statements in 24.000 json files.

We decided to limit the number of astronomical entities to 1.200.000 so as to approximately balance them to each other (although the CHt is about half in size with 625.110 entities), as well as the average number of statements for each entity (CHav: 13,7, CHv: 11,9, CHt: 7,3, As: 22,9, Ag: 12,). In table 1 we show a summary of basic information about these collections.

The statements for all selected entities were downloaded in JSON format[11]. Data is stored in numerous files in JSON format and each file stores a complete representation of exactly 50 Wikidata entities with their labels, descriptions and statements.

4.2. Analysis

In the following we will describe as WLS statements all wikidata statements showing the use of one of the methods described in section 3, regardless of whether they have actually been used to make weaker logical status claims. A tabular presentation of our analysis is shown in table 1.

Even though critical analyses are a pivotal element in humanities discourses, plainly stated statements with no competing claims are largely the most represented information in the dataset: the vast majority of statements in CH dataset (>99%, in particular 99.75% in CHav, 99.93% in CHv and 99.69% in CHt) are plainly asserted statements with no WLS additions. In contrast, the Astronomical datasets show a reasonably different situation, 83% overall of plainly asserted statements, and specifically As at 72,58% and Ag at 95%.

[9] All data used in the analysis are available at https://doi.org/10.5281/zenodo.7624784 [27]. All Python scripts used for this paper are accessible for inspection and reuse at https://github.com/alessiodipasquale/Wikidata_WLS

[10] the As dataset was meant to be composed of 24.000 files with 50 entities each, but after running our tests we noticed that a file was corrupt and we chose to simply discard that contribution.

[11] via https://www.wikidata.org/wiki/Wikidata:Data_access

	Cultural Heritage			Astronomy	
	Audio-visual (*CHav*)	Visual (*CHv*)	Textual (*CHt*)	Stars (*As*)	Galaxies (*Ag*)
Entities	1.251.626	1.078.855	625.110	1.199.950	1.200.000
Statements	17.141.394	12.850.825	4.584.444	27.470.140	14.439.421
Weaker Logical Status (*WLS*)	**95.908 (0.56%)**	**122.666 (0.95%)**	**16.679 (0.36%)**	**7.532.173 (27.41%)**	**721.504 (5.00%)**
Non-asserted statements	43.211	9.056	14.055	7.532.107	721.503
Ranked as Deprecated	7.622	3.057	1.568	2.768.829	189.691
Deprecated with a reason	4.949	769	715	2	0
Null values	50.611	1.969	1.356	4	0
Qualified statements	2.406	114.674	1.556	532	1
WLS qualified statements	2.086	111.641	1.318	62	1
WLS qualifiers w/o *circa*	719	3.988	330	35	0

Table 1

Entities, statements and types of WLS statements

Non-asserted statements: of the methods previously listed (cf. section 3), non-asserted statements (i.e. variously ranked statements with no corresponding asserted triples) are largely the most frequent method for representing competing information in both As and Ag (almost the totality of WLS statements, with 99.99% in both Av and Ag). The situation is fairly different in the CH collections, non-asserted statements being the most frequently used method in CHt (84.01%), but much less so in CHav (only 45.05%) and almost unused in CHv (7.38%).

Deprecated statements: Deprecating a statement (section 3.3) is a major method to avoid assessing a claim and, in addition, assert that the claim has a weaker logical status than others in the same entity. Deprecated claims are visibly a small portion of the overall non-asserted statements, occurring only in 20% of the non-asserted statements of the Cultural Heritage entities and in 30% of the non-asserted statements of Astronomical entities. At the same time, about half of the deprecated statements were annotated with the corresponding `q:P2241`(in particular, 45.59% CHt, 25.15% CHv, 64.93% CHav - compare this with basically 0% in both A datasets), proving that scholars in the humanities have a solid interest in annotating provenance of WLS claims on CH data. Yet, only less that 1% of preferred statements have been annotated with the corresponding qualifier `P7452` *reason for preferred rank*.

Null-valued statements: Null-valued statements are almost non-existent in Astronomical data (exactly 4 occurrences in As and an absolute 0 in Ag out of more than 7 millions Weaker Logical Status claims), and very sparsely used in the Digital Humanities as well: 1,6% in CHv and 8,10% in CHt. Surprisingly higher is the result for the CHav dataset, with 52.77% of the overall WLS claims using this method. This outlier value will be commented later in this section.

Qualifiers: Statements qualified with `P5102` (*reason of statement*) and `P1480` (*sourcing circumstances*) predicates are the least employed representation method out of the surveyed ones, being used in 7.87% of the WLS statements in CHt, and in 2.17% of the CHav statements, present in 0.0002% of the As statements and only in one Ag statement. Yet this approach is used in 91.01% of the WLS statements of the CHv dataset. This value will be commented later on in this section.

We further surveyed the terms actually used as values for the qualifiers. We witnessed the use of respectively 200 different values for qualifier `P5102`, 419 for `P1480` and 588 for `P2241`. These values largely exceed the proposed values specified in the corresponding Wikidata property talk pages (respectively, 194 values for `P5102` and 175 for `P1480` - there are no suggested values `P2241`). Furthermore, the three sets of actual terms show considerable overlap of values between them (in our datasets, but also over all of Wikidata), as shown in figure 1. This seem to imply that the semantics associated to these values, and indeed to the properties themselves, may have been unclear to contributors, who then in some cases selected the qualifier in non-predictable ways. Therefore, we took the decision to group all three sets into a single category (shown as *WLS qualified statements* in table 1).

Overall, the three sets contain a variety of terms such as generic contextual information items, such as provenance details, as well as domain-specific terms not relevant to our purposes (e.g. *show election*, *declared deserted*, or *text exceeds character limit*), as well as qualifiers we can truly consider suggesting weaker logical statuses (e.g., *possibly*, *disputed*, *expected*, etc.).

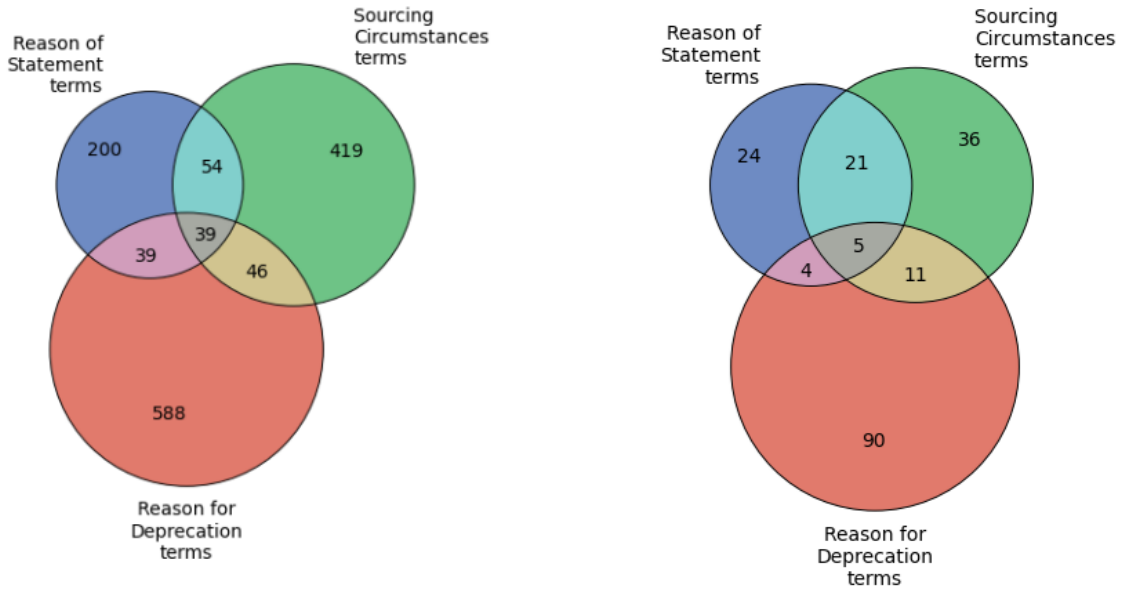

Fig. 1. Terms used in qualifiers `P1502`, `P1480` and `P2241` throughout Wikidata (*left*) and in the CH datasets (*right*)

Therefore, ignoring the list of suggested values provided by the Property Talk pages and focusing on the actual values found in our datasets, we surveyed the list of terms and selected a subset of 101 terms that seem to concretely refer to WLS claims. This subset of WLS terms seem to be widespread in CH datasets (2.086 occurrences in CHav, 111.641 occurrences in CHv and 1.318 occurrences in CHt), while almost not employed in Astronomical datasets (62 occurrences in As and only 1 in Ag).

Interestingly, the value `Q5727902`: *circa* is by far the most employed value in CHv, appearing 107.653 times in `P1480`. This brings the overall count of this value completely out of scale with respect to other values (e.g., the second most frequent WLS term in CHv is "*probably*", occurring only 1.676 times). By removing specifically the value "circa" from the others in the last line of table 1, we see a much homogeneous distribution of values across the three CH datasets. On the contrary, many others terms in the list are present only once in the whole dataset and contribute very little to the overall impact of the *Qualified statements* method.

Another outlier seems to be the fact that null-valued statements are present in the CHav dataset with a much higher proportion than elsewhere. Null-valued statements are heavily employed in some specific properties that appear frequently in the CHav dataset and do not appear elsewhere, such as `P364`: *original language of film or TV show*, `P155`:*follows*, and `P155`:*followed by* In the CHt and CHv datasets these properties do not appear with the frequency and we observe a more heterogeneous distribution of methods (cf. figure 2).

In theory, the methods for WLS are **not** meant as alternative to each other and to be used in an exclusive way. It would be perfectly acceptable and reasonable to use them on the same statement in the same entity, e.g., to describe a deprecated null-valued statement that then results as non asserted). Yet, method overlaps in the surveyed datasets are very poorly represented: both datasets show almost no overlap in the methods employed: no overlap could be found between null-valued statements and the other approaches, and there is a little overlap between deprecated & WLS qualified statements: 0.19 in CHav, 0.15% in CHt, none (0%) in CHv, 1.61% in As and none (0%) in Ag).

To summarise, it becomes manifest that the WLS representation methods employed are quite diverse, even between the datasets of the same domain. Specifically, in CHav the most commonly used WLS representation method is Null Values (52.77%), in CHv it is the WLS Qualified statement (91.61%), and in CHt it is Non-asserted (84.26%). In the astronomy datasets non-asserted statements overwhelmingly represent WLS claims, but Deprecated statements have a much larger impact on them than in the Cultural Heritage domain.

The property analysis provides valuable insights, too, as shown in Figure 2. We divided actual usage of WLS methods by the property in which said method appears. The x-axis contains, for each dataset, the ten most frequent properties in which WLS statements appear, and the y-axis shows in logarithmic scale the number of occurrences of such statements, organized by color: non-asserted statements (with rank normal), non-asserted statements (with rank Deprecated), statements with qualifiers (only WLS-related qualifiers), and null-valued statements.

The analysis of the datasets was performed through a systematic evaluation of the properties associated with the WLS representation methods. Each dataset was analyzed with the aim of identifying (1) the most prominent properties for each dataset, (2) the most prominent properties for each dataset with each method.

We can immediately notice the predominance of null-valued statements in CHav (P364: *original language of film or TV show*, P155: *follows*, P156: *followed by*,P162: *producer* and P345: *IMDb ID*), which goes to prove the peculiarity of the use of null-valued statements in the CHav dataset previously described. The dataset CHt has a considerable number of null-valued statements, too, but only on properties P1476: *title* and P50: *author*, for untitled and/or anonymous documents.

Qualified statements are largely present in CHv and CHt on properties P571: *Inception*, P577: *Publication date*, and P625: *coordinate location*, where, as mentioned, the Q5727902: *circa* qualifier dominates the occurrences.

Normally ranked, yet non-asserted statements appear in large numbers in CHav for P8687: *Social media followers*, P348: *software version identifier*, P175: *performer* and P1476: *title*. They represent peculiar uses of the non-asserted Normal ranks for statements that represent multiple, independent values for the same property, none of which is "more important" than the others. Similar reflections can be made for P18: *image* on dataset CHv, and on properties P1433: *published in* and P921: *main subject* in dataset CHt. Property P1215: *apparent magnitude* dominates this category for astronomical data. Most of the remaining properties employ a Deprecated ranking for evolving or uncertain information.

To summarize, we show here a list of some of the complexities and ambiguities we identified in both the CH and the A datasets.

– **Null-valued statements**.

  * *Data entry errors*: Data include errors probably introduced during the annotation. For instance, the novel "Invisible Monsters" (Q2600527) is both attributed to Chuck Palahniuk (the actual author) and an unknown and probably erroneous entity.
  * *Dumping from pre-existing databases*: some null values may be the result of an error in the conversion or of an empty field of a record after importing an existing database into wikidata. For example the painting "Marshy Landscape" (Q6773948) has a null-valued statement for the catalog code (P528) property.
  * *Model fitting*: When the model does not fully support the situation to be described some arrangements were taken, such as the use of a null value for the property *original language of film or TV shows* P364 when the entity is a silent movie. For example see "Silent Tests" (Q390207), whose (P364) predicate is null-valued and additionally qualified with *applied to part* (P518): *dialogue* (Q131395).
  * *The value does not exist*. To mark the beginning and the end of a sequence, properties such as *follows* (P155) and *followed by* (P156) are entered with a null value for the first and last entities of the sequence. For example, the statement *follows* (P155) for the song (Q5616371) is null-valued.
  * *The value exists but is not known*. For example, the painting "The Welcome Home"(Q110041706) is marked to have an unknown *creator* (P170). This is probably the only true WLS use of null-valued statements.

– **Ranked statements**

  * *Evolving situation*: the claim is not true at the moment, but was correct at some point in the past, and keeping this information is deemed interesting to maintain. For example the number of *social media followers* (P8687) of artists and politicians, the change of *location* (P276) of a movable cultural object such as a painting or a statue, or the change of its *copyright status* (P6216), may change over time and this change is recorded via differently ranked statements. For instance, the print "Races: Anteriel" (Q79471408) recently shifted from the public domain to copyrighted.

* *Evolving knowledge*: because of a new observation or theory, a previous value is considered superseded. This situation is mainly connected to new observations, theories, measurements, guesses and interpretations. For example, the introduction of a new accepted attribution of a work of art means that the previous one is now deemed as false or at least deprecated, or, in astronomy, the object "15 Orionis" (`Q6675`) was previously considered an *instance of* (`P31`) an *infrared source* (`Q67206691`), but it is now fully considered as a *star* (`Q523`);
* *Less favoured versions*: competing claims are not described as neither false nor true, but one of them is preferred over the others so that they are marked as preferred and asserted while the others are non-asserted. For example, the *titles* (`P1476`) of textual works are often provided in different languages, and the title in the original language is marked as the preferred version, while the translated titles in other languages are not asserted.

– **Qualified statements**

* *Imprecision*: for instance, the hypothetical entity "IRAS 17163-3907"(`Q540167`) has an observed *luminosity* (`P2060`) property set to "500,000 solar luminosity" with a `P1480` qualifier *circa*; similarly, the painting "Girl Reading a Letter at an Open Window" (`Q700251`) by Johannes Vermeer is dated (`P571`, *inception*) 14th century with a `P1480` qualifier *circa*.
* *Uncertainties*: for instance, the painting "Madame Antoine Arnault"(`Q109252498`) has *creator* (`P170`) set to Jean-Baptiste Regnault (`Q453485`) with a `P5102` qualifier *disputed*;
* *Contextualizing qualifiers*: for instance, the star "Altair" (`Q12975`) has a *flattening* (`P1102`) property set to 0.2 with a `P5102` qualifier *greater than*;
* *Cautioning qualifiers*: for instance, the "Frontispiece to Christopher Saxton's Atlas of the Counties of England and Wales State I" (`Q105949375`) has the *creator* (`P170`) property set to Remigius Hogenberg (`Q18576859`), with the contributor cautioning through a `P1502` qualifier that this is only an *attribution*.

4.3. Discussion

The datasets presented in the previous section and the analysis we performed on their content allows us to reach some conclusions on the research questions specified in the introduction.

RQ1 - *How widespread and successful is each of these approaches in the current state of Wikidata?* - The current state of WLS claims in Wikidata is poor. Even though Wikidata focus on established knowledge (community consensus), rather conjectural or controversial information[12], in many cases it is objective and scientifically precise to represent the complexity of uncertainty and evolving knowledge, rather than omitting information because they are not completely established. In these cases, Wikidata seems doing poorly, as <1% of the claims we analyzed in CH datasets show weaker logical status characteristics, a much lower figure than the 5% in the Ag dataset or the 27,41% figure of As data. Does this show an intrinsic difference in the two cultural domain or is there something else underneath? To provide an answer to this further question, we turned to the RKD database. RKD[13] holds detailed data about Dutch and Flemish paintings, drawings and prints throughout the ages, from XVI Centuries artworks to modern ones. Overall, more than 260.000 items belonging to the images collections are described, and through the use of its EDM-inspired datamodel a particular attention is given to multiple competing attributions, e.g. authorship attributions (counting more than 317.000 recorded attributions). Although Dutch and Flemish artworks may not be representative of the full scale of world-wide collections of artworks represented in the CHv dataset, they can provide an interesting comparison with it. We were able to discover that deprecated authorship attributions are present in about 8,5% of the attributions in the RKD image collection (circa 290.00 former attributions vs. 27.000 discarded attributions on 265.000 items stored in the RKD images collection), a conspicuously higher figure than the meager

[12] https://www.wikidata.org/wiki/Help:Ranking#What_ranks_are_not
[13] see https://rkd.nl/en/

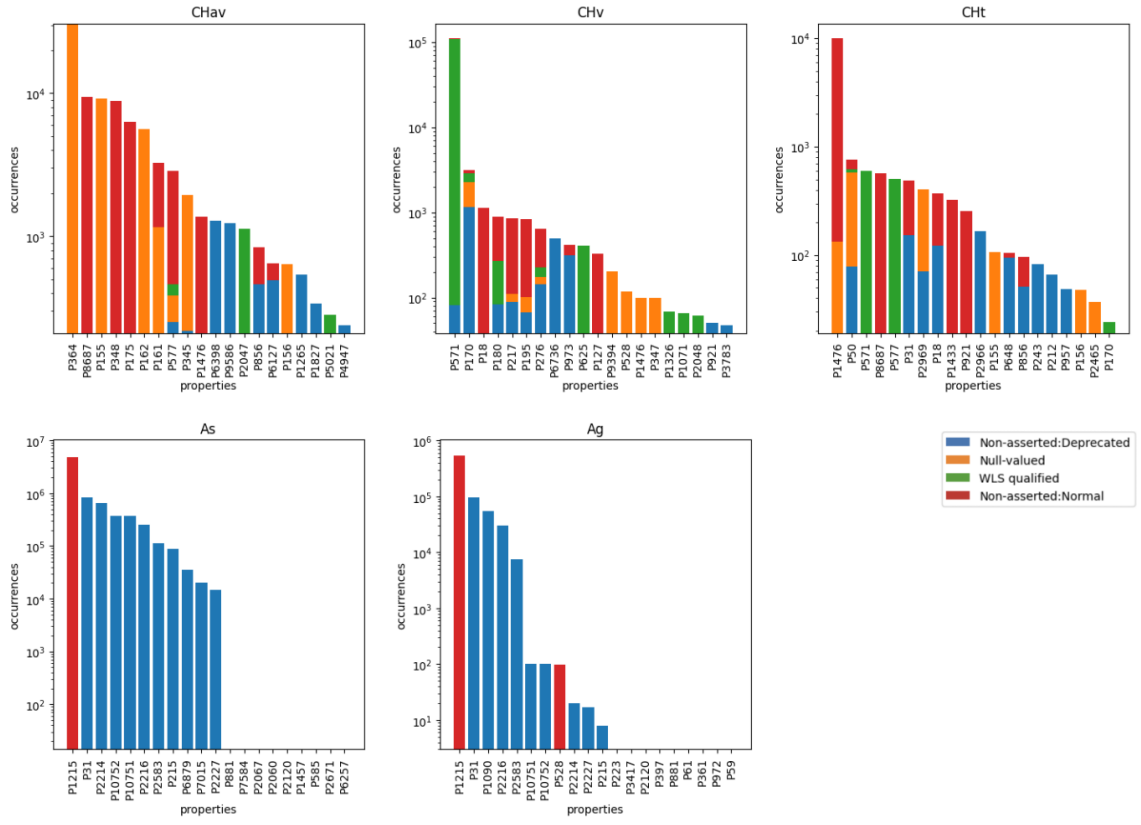

Fig. 2. Top 10 most recurrent properties implied in WLS claims in each dataset

0.95% WLS statements of the CHv dataset. This fact may indicate a radical underrepresentation of complex attributions within Wikidata entities. We may conclude WLS statement are not particularly widespread nor successful in Wikidata collections within the Cultural Heritage domain, and they are possibly misrepresenting the complexity and variety of situations that exist in this domain.

RQ2 - *How does the cultural domain of the Wikidata topics (and, presumably, of the individuals contributing to the data regarding the wikidata topics) affect and reflect on the relative success and richness of some approaches over the others?* - Our analysis of data highlighted a number of peculiarities between the Cultural Heritage datasets and the Astronomical ones. The two families of datasets present many different representational artefacts: while the CH datasets seem to employ, with variable proportions, all the listed methods, the astronomical datasets employ almost exclusively ranked statements. Additionally while WLS statements in A datasets affect a fairly small number of properties, they cover a much wider range of properties in CH, as shown in figure 2. These aspects seem to highlight key differences in what the two communities consider weaker logical status: without committing too much to interpretations outside our competency, we may hypothesize that deprecations in astronomical data mostly reflect the result of newer and better data (more recent observations, maybe?) while the humanities community uses WLS statements for a much larger set of uncertainties of due to ignorance, scholarly interpretations and disagreements. Thus it may occur that the specification of P5102 (reason of statement) and P2241 (Reason for deprecated rank) qualifiers may seem overkill in astronomical data, and a real necessity for some annotations in the humanities.

RQ3 - *How clean and easy to differentiate are the applications of each approach to an actual weaker logical status versus to another of the designed uses of that approach?* - Unfortunately, there is much noise and ambiguity in how Wikidata contributors have used WLS methods in the datasets we studied. This makes it very difficult to

differentiate and search WLS data. The variety of cases listed at the end of section 4.2 summarizes a probably lacking yet vast panorama of WLS and non-WLS knowledge situations modelled through WLS representation methods. It is therefore fairly difficult not only to search for specific data patterns over the full dataset, but even to interpret correctly individual entities in a proper manner.

Furthermore, the use of the same methods for WLS and non-WLS related characterizations makes complex patterns very hard to express and identify. For instance, if an artwork A was *supposedly* moved from location X to location Y, but we are not certain, then both location X and location Y must be represented as deprecated, the first because of an evolving situation (A is not at location X anymore) and the second because of uncertainty, since the new location Y is only guessed. Without a complete and thorough contextual annotation (e.g. why each claim is discarded) disambiguation and full understanding is impossible (it is also worth noting that only 50% of deprecated statements are annotated with a *reason of deprecation* qualifier `P2241`).

RQ4 - *Is there a way to improve the clarity and cleanliness of such differentiation?* - getting down to detailing workable solutions to improve the situation for WLS statements in a project as large and as complex as Wikidata is always running the risk of becoming an exercise in futility. In the next section we will try to propose a list of possible actions for WLS statements, starting from very conservative proposals with limited impact, up to bolder and more impacting changes.

5. Towards a leaner and harmonic support for WLG in Wikidata

In this section we enumerate a short list of possible remediation activities to be performed over the Wikidata data model and the collection itself so as to simplify and disambiguate WLS assertions from the rest. We approach such a complex endeavor with humility and caution, as we are well aware that it may be hard to assess from our vantage point both the impact and the difficulty of the implementation of each suggested step.

For this reason, we express our suggestions as an ordered list whose first items are meant as simple cleaning up activities of little impact, and then progress to bolder and more impacting actions that sometimes require not just a modification in the data model, but possibly also the systematic update of small, but still numerically relevant, selections of the current datasets.

1. Reorganize, simplify, and re-categorize the suggested values for qualifying properties `P5102`, `P1480`, `P2241`, and `P7452`. Provide a list of suggested values for `P2241` and `P7452`. The lists should be clearly differentiated and with no semantic overlaps neither between lists nor within each list.
2. Require a P7452 qualifier in all Preferred statements and a P2241 qualifier in all Deprecated statements. Provide simple to use interface widgets for their specification. Make sure that no such statements can be saved without a qualifying proposition.
3. Require the specification of a P5102 in all null-valued statements. Extend the list of suggested values for P5102 with values specifically addressing null-valued statements, allowing to easily distinguish the various situations we encountered, in particular between actual WLS uses (e.g., anonymous: author exists but is not known) and a non-WLS uses (e.g., last of a sequence: following entity does not exist).
4. Require the specification of `P5102` and `P1480` qualifiers for all WLS-related rankings: only asserted statements with Normal rank are allowed to remain without qualifiers.
5. Create a new and separate *Certainty Degree* qualifier specifically for WLS statements, separating the reason for the chosen qualification from the certainty or confidence degree of the qualification. Such certainty degree should be scalar and use a limited number of values, avoiding any complexity in distinguishing between, e.g.: possibly, presumably, hypothetical, dubious, etc.). A 5- or 7-item scale would suffice, e.g.: *non accepted, highly unlikely, unlikely, possible, probable, almost surely*, and *accepted*. Different labels, and even the use of numerical values instead of labels, would be perfectly acceptable.
6. Reorganize the values of P5102 and P1480 to remove values merely representing a uncertainty (replaced by the new *Certainty Degree* qualifier). To this end an initial list of values is being created. The current list has been generated by following a Grounded Theory approach [28]: first, labels, definitions and usage data

of suggested and actually used qualifiers have been collected and categorized to represent different macro-themes or concepts. These concepts allowed theories to emerge and be developed from the coded data with an iterative process that continued until the theory was "grounded" in the data. The resulting list in its current state, collecting the surveyed terms from the Wikidata *property talk* pages and the terms actually used in the CH datasets, contains 150 values referring to WLS claims and organized in 18 theories, and can be accessed in the Github folder of the project[14].

7. Restrict ranking for competing statements to just three (possibly four) different patterns, and prevent any other variant:
   - *Preferred* + *Deprecated*: to be used whenever there is a number of competing statements and some of them are clearly chosen to be the best ones. Accepted statements are set to Preferred (and asserted) while the rest is set to Deprecated (and not asserted); there are no Normal ranks. Both Preferred and Deprecated statements are fully qualified with P5102, and with P7452 and P2241 respectively, and the new Certainty qualifier. Preferred statements would be assigned a *accepted* or *almost surely* degree, while Deprecated ones would assigned a *not accepted* or *highly unlikely* certainty degree. Intermediate degrees would not be used.
   - *Normal rank* + *asserted*: this would be the default situation, to be used when no dispute or disagreement exists and the statement(s) are all equally accepted. All statements are also asserted. Since this is the default no qualifier is necessary, but it is still possible to specify a P5102 or a P1480 value. No certainty degree is necessary.
   - *Normal rank* + *non-asserted*: to be used when there is a number of competing statements but none of them stands above the rest as being the most likely. This would be the case, for instance, of a work of art not definitely attributed to anyone, but for which several competing hypothesis exist, but none seem more convincing than the others. No statement is asserted, and P5102 and/or P1480 values are required. All statements would be assigned a value from the central ones, from *highly unlikely* to *probable*, to the exclusion of the extremes.

   A fourth pattern could be in theory allowable, that of a claim for which the only reported value is clearly wrong, but no acceptable alternative exists. In this case, we could use a Deprecated statement for the reported wrong value, and a null-valued statement with Normal ranking to represent the unknown correct value.

6. Conclusions and future works

Our work is the first systematic study about the representation of weaker logical status claims (WLS) over cultural heritage data in Wikidata. Through WLS claims it is possible to express uncertain information, competing hypothesis, temporally evolving information, etc. for which a plain and direct assertion is inappropriate. We analysed four patterns used in Wikidata for WLS claims, asserted vs. non-asserted statements, ranked statements, null-values and qualifiers.

In our analysis we found out a number of interesting facts. First of all, very few statements are expressed using weaker logical status than could have been expected by comparing other similar sources. Second, the wikidata data model, far from being too poor for expressing WLS claims, has shown to provide, in fact, a overabundance of methods, but there seem to a large overlapping in uses between themselves and also towards non-WLS applications. Finally, there are important differences in how different datasets coming from different domain employ these methods for weaker logical status claims. It seems that domain-specific non-WLS situations can be considered as a justification for much of this variety, and this contributed to the idea that WLS-specific features should be introduced in the Wikidata model to address specifically weaker logical status claims. We proposed a set of increasingly impacting modifications to the data model aiming towards a leaner and more accurate representation of these phenomena with the expectation that they can manage to improve data quality and information retrieval specifically over uncertain, evolving and competing statements.

[14] https://github.com/alessiodipasquale/Wikidata_WLS

We are still working toward a full taxonomy of values for qualifying ranked predicates, as this seems to be to our eyes the most rapid and solid way to fully represent both the weaker logical status of a claim and its underlying nature and justification. We plan to publish such taxonomy with a proposal for mapping existing data points into such taxonomy in order to loose no information in the conversion.

References

[1] F. Erxleben, M. Günther, M. Krötzsch, J. Mendez and D. Vrandečić, Introducing Wikidata to the Linked Data Web, in: *The Semantic Web - ISWC 2014*, P. Mika, T. Tudorache, A. Bernstein, C. Welty, C. Knoblock, D. Vrandečić, P. Groth, N. Noy, K. Janowicz and C. Goble, eds, Springer International Publishing, Cham, 2014, pp. 50–65. ISBN 978-3-319-11964-9. doi:doi.org/10.1007/978-3-319-11964-9_4.

[2] C. Möller, J. Lehmann and R. Usbeck, Survey on English Entity Linking on Wikidata: Datasets and Approaches, *Semantic Web* **13** (2022). doi:10.3233/SW-212865.

[3] M. Doerr, S. Gradmann, S. Hennicke, A. Isaac, C. Meghini and H. Van de Sompel, The europeana data model (EDM), in: *World Library and Information Congress: 76th IFLA general conference and assembly*, Vol. 10, 2010, p. 15.

[4] M. Doerr, C.-E. Ore and S. Stead, The CIDOC conceptual reference model-A new standard for knowledge sharing, in: *26th international conference on conceptual modeling (ER 2007)*, 2007.

[5] M. Piotrowski and M. Neuwirth, Prospects for computational hermeneutics, in: *Proceedings of the 9th AIUCD Annual Conference*, 2020. http://amsacta.unibo.it/6316/.

[6] M. Fafinski and M. Piotrowski, Modelling Medieval Vagueness, in: *INFORMATIK 2020*, R.H. Reussner, A. Koziolek and R. Heinrich, eds, Gesellschaft für Informatik, Bonn, 2021, pp. 1317–1326. doi:10.18420/inf2020_123.

[7] M. Daquino, V. Pasqual and F. Tomasi, Knowledge Representation of digital Hermeneutics of archival and literary Sources, *JLIS.it* (2020), 59–76. doi:https://doi.org/10.4403/jlis.it-12642.

[8] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak and Z. Ives, Dbpedia: A nucleus for a web of open data, in: *The Semantic Web: 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007+ ASWC 2007, Busan, Korea, November 11-15, 2007. Proceedings*, Springer, 2007, pp. 722–735. doi:doi.org/10.1007/978-3-540-76298-0_52.

[9] T. Rebele, F. Suchanek, J. Hoffart, J. Biega, E. Kuzey and G. Weikum, YAGO: A multilingual knowledge base from wikipedia, wordnet, and geonames, in: *The Semantic Web–ISWC 2016: 15th International Semantic Web Conference, Kobe, Japan, October 17–21, 2016, Proceedings, Part II 15*, Springer, 2016, pp. 177–185. doi:doi.org/10.1007/978-3-319-46547-0_19.

[10] V. Petras, T. Hill, J. Stiller and M. Gäde, Europeana–a Search Engine for Digitised Cultural Heritage Material, *Datenbank-Spektrum* **17** (2017), 41–46. doi:10.1007/s13222-016-0238-1.

[11] E. Delmas-Glass and R. Sanderson, Fostering a community of PHAROS scholars through the adoption of open standards, *Art Libraries Journal* **45**(1) (2020), 19–23–. doi:10.1017/alj.2019.32.

[12] E.E. Fink, American Art Collaborative (AAC) Linked Open Data (LOD) Initiative, Overview and Recommendations for Good Practices (2018). https://repository.si.edu/bitstream/handle/10088/106410/OverviewandRecommendationsAccessible.pdf.

[13] A. Isaac et al., Europeana data model primer (2013). https://pro.europeana.eu/page/edm-documentation.

[14] G. Barabucci, F. Tomasi and F. Vitali, Supporting complexity and conjectures in cultural heritage descriptions, in: *Proceedings of the International Conference Collect and Connect: Archives and Collections in a Digital Age*, CEUR Workshop, 2021, pp. 104–115. http://ceur-ws.org/Vol-2810/paper9.pdf.

[15] A. Stinson, S. Fauconnier and L. Wyatt, Stepping Beyond Libraries: The Changing Orientation in Global GLAM-Wiki, *JLIS.it* **9**(3) (2018), 16–34–. doi:10.4403/jlis.it-12480. https://www.jlis.it/index.php/jlis/article/view/95.

[16] M. Zhitomirsky-Geffet and S. Minster, Cultural information bubbles: A new approach for automatic ethical evaluation of digital artwork collections based on Wikidata, *Digital Scholarship in the Humanities* (2022). doi:10.1093/llc/fqac076.

[17] M. Mora-Cantallops, S. Sánchez-Alonso and E. García-Barriocanal, A systematic literature review on Wikidata, *Data Technologies and Applications* **53**(3) (2019), 250–268. doi:doi.org/10.1108/DTA-12-2018-0110.

[18] D. Hernández, A. Hogan and M. Krötzsch, Reifying RDF: What works well with wikidata?, *SSWS@ ISWC* **1457** (2015), 32–47.

[19] A. Piscopo and E. Simperl, What We Talk about When We Talk about Wikidata Quality: A Literature Survey, in: *Proceedings of the 15th International Symposium on Open Collaboration*, OpenSym '19, Association for Computing Machinery, New York, NY, USA, 2019. ISBN 9781450363198. doi:10.1145/3306446.3340822.

[20] M. Färber, F. Bartscherer, C. Menne, A. Rettinger, A. Zaveri, D. Kontokostas, S. Hellmann and J. Umbrich, Linked Data Quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO, *Semantic Web* **9**(1) (2018), 77–129–. doi:10.3233/SW-170275.

[21] V. Balaraman, S. Razniewski and W. Nutt, Recoin: Relative Completeness in Wikidata, in: *Companion Proceedings of the The Web Conference 2018*, WWW '18, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 2018, pp. 1787–1792–. ISBN 9781450356404. doi:10.1145/3184558.3191641.

[22] M. Ponza, P. Ferragina and S. Chakrabarti, A Two-Stage Framework for Computing Entity Relatedness in Wikipedia, in: *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, CIKM '17, Association for Computing Machinery, New York, NY, USA, 2017, pp. 1867–1876–. ISBN 9781450349185. doi:10.1145/3132847.3132890.

[23] L. Galárraga, S. Razniewski, A. Amarilli and F.M. Suchanek, Predicting Completeness in Knowledge Bases, in: *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, WSDM '17, Association for Computing Machinery, New York, NY, USA, 2017, pp. 375–383–. ISBN 9781450346757. doi:10.1145/3018661.3018739.

[24] K. Shenoy, F. Ilievski, D. Garijo, D. Schwabe and P. Szekely, A study of the quality of Wikidata, *Journal of Web Semantics* **72** (2022), 100679. doi:https://doi.org/10.1016/j.websem.2021.100679. https://www.sciencedirect.com/science/article/pii/S1570826821000536.
[25] D. Abián, F. Guerra, J. Martínez-Romanos and R. Trillo-Lado, Wikidata and DBpedia: A Comparative Study, in: *Semantic Keyword-Based Search on Structured Data Sources*, J. Szymański and Y. Velegrakis, eds, Springer International Publishing, Cham, 2018, pp. 142–154. ISBN 978-3-319-74497-1.
[26] Help:Ranking - Wikidata. https://www.wikidata.org/wiki/Help:Ranking.
[27] A.D. Pasquale, F. Vitali and V. Pasqual, Wikidata selection of Cultural Heritage, Stars and Galaxies entities and claims, Zenodo, 2023. doi:10.5281/zenodo.7624784.
[28] B.G. Glaser and A.L. Strauss, *The discovery of grounded theory: Strategies for qualitative research*, Routledge, 2017. ISBN 1351522167.