CHAPTER 3

Art description and analysis in the computational era

In recent years, increasing attention has been paid to a structured representation of art-related topics. In the context of the description of artwork subject matter and content, several notable works were conducted, providing domain-specific cataloging standard guidelines, controlled vocabularies, and ontologies. Diverse attempts in applying Computer Vision to artworks to recognize a subject show promising results, too. Furthermore, data analyses of the art market and cultural-related context hint at the kind of understanding that data may provide. However, little data about iconography and iconology are available in the Knowledge Graphs.[1] Therefore, we conducted a study to assess their presence and correctness among the available KGs.

3.1 Cataloguing standards

To be able to share and relate artwork data on the Web, each institution must annotate its holdings according to shared standards, which can be done at different levels of detail. Cataloging standards offer a set of guidelines for entities that describe a resource. For what concerns the cataloging of works of art, an international reference point is the Getty's *Categories for the Description of Works of Art* (CDWA),[2] partially incorporated into the guide *Cataloguing Cultural Objects* (CCO) (Harpring et al., 2006). The CDWA are a set of guidelines created by the Getty Vocabulary Program to standardize the descriptions of works of art held in various institutions. It comprises about 540 categories and subcategories of information organized around basic categories that can be used

1 Knowledge graphs can be defined as 'massive semantic nets that integrate various and heterogeneous information sources to represent knowledge about specific domains of discourse' (Fensel et al., 2020). On a technical level, KGs are sets of triples (subject, predicate, object) encoded in a serialization of the Resource Description Framework (RDF) (McBride, 2004)

2 https://www.getty.edu/research/publications/electronic_publications/cdwa/definitions.pdf

Published in: Sofia Baroncini, Diving into artworks interpretations through the lenses of semantic data. An Application on Panofsky's Iconological Studies, Heidelberg: arthistoricum.net 2026, https://doi.org/10.11588/arthistoricum.1462

to identify a work of art. The *Subject Matter* category[3] allows us to describe the subject and meaning of an artwork. In detail, the *Specific Subject Type* subcategory allows classifying different reading levels, as defined by Panofsky (description, identification, interpretation), to distinguish subject matter from content. Both CDWA and CCO underline that adopting a simplified description of the approach by Panofsky 'can be helpful in indexing subjects for purposes of retrieval.'[4] Following the alignment first proposed by Shatford (1986), they define the second and third level, that is, the identification of themes, narratives, iconographies, and meanings, as the *aboutness* (i.e. what the work is about). In contrast, the first level and eventually the second corresponding to the *ofness* (viz. what can be seen by a non-expert interpreter [Žumer et al., 2012, pp. 207–208; Klenczon & Rygiel, 2014]). If the subject corresponds to the work itself (e.g., the term *architecture* used for describing a cathedral) and does not refer to a subject depicted by the object (e.g., a drawing representing a cathedral), the term *isness* shall be used. The concepts of *ofness*, *aboutness*, and *isness* are a core aspect of knowledge organization initiatives (ISKO, IFLA) and are discussed in more detail in Zeng et al. (2009) and Hjorland (2016).

3.2 Controlled vocabularies

Controlled lists of terms are important components of the knowledge organization field, as they provide uniquely defined descriptors of concepts to allow interoperability and an efficient information retrieval of the metadata of the cataloged entries. A controlled vocabulary is 'an organized arrangement of words and phrases used to index content and/or to retrieve content through browsing or searching' (Harping, 2013, p. 13). In this sense, it includes any kind of controlled list. Authority files are controlled vocabularies presenting preferred versions of the terms, and relations with their synonyms. If the terms are organized in a clear hierarchical structure, the controlled list is defined as a taxonomy. Furthermore, those controlled lists of terms that present explicit relations among the concepts identified by such terms are named thesauri.[5] In the art history sector, the most widely adopted vocabularies for the description of art history object, and specifically for the description of artworks' subject matter, are the Getty Vocabularies and Iconclass.

The Getty Research Institute released five vocabularies,[6] specifically for the description of artworks (Cultural Objects Name Authority), iconography (Iconography Authority), general terms for describing artworks (Art and Architecture Thesaurus), artists (Union List Of Artists Names), and places (Thesaurus of Geographical Names).

3 https://www.getty.edu/publications/categories-description-works-art/categories/object-architecture-group/16/

4 For the alignment of the concept of the subject matter to the main cataloging standards we refer to section 16 of *Metadata Standard Crosswalk*, available at https://www.getty.edu/research/publications/electronic_publications/intrometadata/crosswalks.html

5 The ISO standard defines a thesaurus as a 'controlled and structured vocabulary in which concepts are represented by terms, organized so that relationships between concepts are made explicit' (ISO 25964-1; clause 2.62, available at: https://www.iso.org/obp/ui/en/#iso:std:iso:25964:-1:ed-1:v1:en)

6 https://www.getty.edu/research/tools/vocabularies/

Each item described in the vocabulary has a unique identification code, related terms, and sources. When possible, alignment with other standards is provided. As terms are organized in a hierarchy, each term specifies the hierarchy to which it belongs and the type of hierarchical relationship present. Furthermore, since vocabularies adopt a thesaural structure, they provide the preferred term of a concept, enriched with variants and multilingual translations, as well as internal association relationships within the terms.

The **Iconography Authority** (IA) includes proper names, relationships, and dates for iconographical narratives, characters, events, and literary works in a standard-compliant format. As it has a thesaural structure, it includes equivalence, associative, and hierarchical relationships. The thesaurus is aligned with other ones (e.g., Library of Congress and Iconclass), and it is collaboratively updated and enriched. Although the Getty vocabularies are progressively being published as Linked Open Data, the Iconography Authority has not been published in this format yet.

The **Art and Architecture Thesaurus** (AAT) includes generic concepts, not identifiable by a proper name, needed to catalog and retrieve information about art and architecture. Its development began in the 1970s to respond to the need for art journal indexing services and art libraries, which started to automate the indexing process.[7] Its wide scope reflects the aim to meet the needs of various user groups (e.g., historians, architects, librarians, museum personnel, etc.) to allow cross-collection information retrieval. Descriptions of materials, abstract concepts, styles, roles, and techniques are included. AAT is one of the vocabularies exposed as Linked Open Data viewable in JSON-LD, RDF, N3 / Turtle, and N-Triples.

Iconclass[8] is a classification system for iconographical description with standardized controlled terminology widely used for research purposes and information retrieval in museums. Designed in the 1940s by Henri van der Waal (van Straten, 2012), it is the first controlled vocabulary developed even before modern computer science's birth. Based on a system of numbers and letters identifying each entry, it can uniquely identify each resource according to W3C Semantic Web indications. Each identifier is expressed as a URI,[9] and it is available in LOD.

Iconclass is one of the sources of the Getty Vocabularies. It presents a hierarchical structure based on ten macro-categories, identified by numbers from 0 to 9, enriched with generic associative relationships between entries. The first category was introduced later to identify contemporary art subjects (0 Abstract, Non-representational Art), including abstract shapes, colors, and concepts. The other sections describe the physical world (2 Nature, 3 Human Being, Man in General) and aspects concerning society (1 Religion and magic, 4 Society, Civilization, Culture, 5 Abstract Ideas and

7 https://www.getty.edu/research/tools/vocabularies/aat/about.html

8 https://iconclass.org/

9 A URI (Uniform Resource Identifier), is the unique identifier for resources. It diverges from a URL because it refers to the resource in spite of its location storage, as a URL does. For their stability in time, they are used in the semantic web context and can be semantically linked to the URIs of other resources using semantic web properties. For additional information on the concept of URIs, we refer to https://www.w3.org/TR/webarch/

Concepts, 6 History). The last three categories (7 Bible, 8 Literature, 9 Classical Mythology, and Ancient History) deal with the most famous literary sources of subjects in Western art. Each category is further deepened by more specific terms, identifying about 450 entries that gradually include more specific items. Specifications may refer to general concepts or individuals (e.g., *the Announcement of Christ's birth*, code 73A5), which can be more thoroughly described (e.g., *the Annunciation with Mary kneeling*, code 73A523).

Although both Getty Vocabularies and Iconclass are excellent solutions to support most iconographic recognition tasks, there are naturally situations in iconographical and iconological studies that cannot be represented through these taxonomies. The study by Baroncini et al. (2021) highlights the different situations in which a complex representation can not be fully represented. For instance, relations among subject entries and possibly shared symbolical meanings they express can not be stated. Moreover, the transmission of iconographical themes over time, treated by a consistent number of iconological studies, and their variation according to the period's style can not be described. Furthermore, Getty Vocabularies allows associating a date with a subject, but the stylistic variants cannot be related to the respective historical period. Likewise, Iconclass and Getty Vocabularies allow specifying the source of the iconography. However, it is impossible to relate a source to a variant of a subject (e.g., Durer's Rape of Europa depicted following the description provided by the Renaissance writer Poliziano). Therefore, to fully express the complexity embedded in the branch of iconological studies, other more flexible means are needed.

3.3 Ontologies

In the field of computer science, the term *ontology* identifies an explicit formalization of a domain, that is, a machine-readable representation consisting of the definition of concepts and their relations according to a shared view. As a result, the core structure of an ontology consists of a taxonomy that describes the relevant terms in a hierarchical structure (Staab & Studer, 2010, p. 2). The formalization is made intensionally[10] by defining axioms, namely, logical definitions among the concepts, which makes it possible to perform reasoning and infer new knowledge from data described through ontologies. Ontologies can be used as a model to represent the individuals reflecting such definitions, i.e., instances.[11] Among the available serializations, the Web Ontology Language (OWL)[12] is the standard used to define such concepts and relations in the semantic web domain. A good practice of ontology development is to define explicit

10 In logic, an object can be defined intensionally by specifying the properties that the object should have in order to be defined as such (e.g., a 'living being' should be born). More detail is available at: https://en.wikipedia.org/wiki/Extensional_and_intensional_definitions

11 (e.g., the individual 'John' born in Rome in 1981 can be described as a person, class of an ontology; as people can read, and the ontology establishes the relation 'reads' between people and books, we can describe John as reading the book *Invisible Cities* by Italo Calvino

12 https://www.w3.org/2001/sw/wiki/OWL

alignments of the terms of a newly created ontology with standards or foundational ontologies, namely, ontologies defining high-level, domain-independent concepts and relations that serve as a basis for further ontology developments that guarantee the compliance of the underlying logic adopted. In this way, even if the description of instances in different datasets is made according to different ontologies, the logic and the semantics remain consistent.

For their characteristics, conceptual models and ontologies often constitute standards of a community or domain to allow interoperability of knowledge stored by different institutions. In the domain of cultural heritage, the standard is CIDOC-CRM (Bekiari et al., 2021), an ontological conceptual model defined by the International Council of Museums (ICOM). It has an event-based structure and covers fundamental aspects of the life cycle of a cultural heritage object. Nevertheless, the knowledge expressed is centered on the generic aspects of cultural heritage objects to provide a generic modeling that can be adopted in multiple contexts and foster information aggregation and retrieval in the museum sector. For this reason, further ontologies were proposed to describe specific aspects of cultural heritage objects, their making, interpretation and value over time, which we illustrate in the following sections.

3.3.1 Art history

In the context of art history, several attempts have been made to create models that cover some specific elements related to the description and interpretation (i.e., the recognition of symbolic meanings) of a cultural heritage object.[13] The linked.art project provides application profiles based on CIDOC-CRM which address specific aspects of the artwork description and make the implementation of the standard ontology more usable and comprehensible.[14] Recently, the Art and Architecture Argumentation Ontology, (AAAO),[15] was released, which extends CIDOC-CRM with a focus on expressing the contextual social facts as temporally and socially bounded, as they may vary across time and space and among different social groups (e.g., the reconstruction of the ownership of an object may vary according to the reference social group). The ontology is relevant for integration of diverging knowledge and for describing in an explicit way the unstable, changing nature of contextual art information.

Some ontologies focused on the modeling of the visual content of artworks. Carboni et al. extended CIDOC-CRM with the Visual Representation ontology (VIR) (Carboni & de Luca, 2019). VIR ontology explores the concept of visual representations in artworks. It associates the portion (called *iconographical atom*) of the cultural heritage object to the recognized subject by representing an interpretation act. Thanks to such structure, further specifications of the assertion can be added, such as the person expressing the interpretation and the sources on which it is based. Although funded on a thorough

13 Section 3.3 was published in (Sartini et al., 2023), here presented in an enriched version

14 https://linked.art/about/. See Sanderson (2024) for an example of the implementation of the profile

15 https://www.sari.uzh.ch/en/ordea/aaao.html

consideration of interpretation level theories described in Chapter 2, VIR focuses only on subjects of level 2, considering iconographies and their attributes, consequently lacking a clear distinction between levels. The preliminary study in Baroncini et al. (2021) further extends VIR by adding an iconological interpretation class linking the concepts of the artwork and external cultural phenomena. It is evaluated on 11 real case studies taken from the literature in iconology, which illustrate a wide variety of aspects included in an iconological analysis. In addition, the work is based on a careful theoretical comparison of the main iconological and iconographical interpretation theories. For its comprehensive overview of iconographical and iconological theories, along with the real-base evaluation, it is used as a source for the ontology developed as part of the current Ph.D. project, which has to be seen as its development and refinement. We deepen this study by developing aspects not already considered, such as a more detailed description of level 1 and level 2 subjects and integrating multiple interpretations by different art historians (see section 5.3). Gartner (2020) proposes an ontology, which was not released, to facilitate and automate the identification of subjects (level 2) in works of art through logical inferences.

The ARCO ontology (V. A. Carriero et al., 2019) was developed to model artifacts from Italian cultural heritage by converting the information contained in traditional catalog sheets into Linked Open Data. Among the possible aspects modeled for an artwork, some classes were designed to describe its iconographical apparatus. Apart from this class, the schema does not mention any distinctions between different levels of interpretations. Most of the information about the iconographical and iconological interpretations in ArCo are provided through natural language descriptions with the property `dc:description` or `core:description`,[16] not exploiting the full potential of the Semantic Web (Sartini & Gangemi, 2021). Sartini et al. (2021) modeled symbolic meanings in cultural heritage in the Simulation Ontology. This ontology does not consider the hermeneutic act of interpretation of the association of symbolic meanings with artworks. However, its conceptualization of symbolic meanings using n-ary relationship classes that link a symbol, its symbolic meaning, and the cultural context in which the symbol-symbolic meaning relationship proposes a thorough description of symbols.

3.3.2 Interpretations and meaning

In the context of knowledge organization, several ontologies addressed the concept of interpretation.[17] CIDOC-CRM models assertions with the class `E13_Attribute_Assignment`, which relates the assertion made by one agent to the object considered. Since each assertion reflects the agent's opinion, multiple, contradictory assertions may be represented. The concept of interpretation is applied broadly, including measurements and other types of scientific observations. Similarly, the class `Interpretation`

16 See https://dati.beniculturali.it/lodview-arco/resource/HistoricOrArtisticProperty/0500653281.html dc:description value.

17 the current section was published in (Sartini et al., 2023)

of Arco is intended to describe every piece of information asserted by an agent about an object on the basis of stated sources.[18] The CIDOC-CRM extension CRMinf deepens the concept expressed by `E13_Attribute_Assignment` distinguishing the type of argumentation and if the belief resulting from the argumentation holds true or not. The concept is further explored by the CRMsci, another CIDOC-CRM extension, which integrates CRMinf by formalising the shared scientific process adopted across different domains and the scientific activities involved. In detail, of great interest is the class `crmsci:S4_Observation`, subclass of `crmsci:I1_Argumentation` and of `crm:E13`, expressing the scientific observation of physical events or reality which is done directly or through measurements. It represents the 'transition between reality and propositions´ (Doerr et al., 2014). Furthermore, the VIR ontology adds a domain-definition of `crmsci:S4 Observation` by declaring its subclass `vir:IC12 Representation`, which represents an assignment of a solely iconographical status to a physical object.

The same topic is also addressed by the history domain to represent the frequent case of disagreeing historians' interpretations of the same events. As reported by Frank (2019), several ontologies afford the theme by modeling different views of the same observed events, such as the SEM ontology[19] MIDM (Van Ruymbeke et al., 2017), the ODP Event-Model-F,[20] expanding DOLCE+DnS Ultralite (DUL), in which a distinction between facts and interpretations has already been made (Gangemi & Mika, 2003). The STAR model developed in the Releven project (Andrews, 2023) reuses CIDOC-CRM to describe the provenance of the assertion and its authority. HiCO (Daquino & Tomasi, 2015)[21] goes further by adding contextual information to the interpretation, such as interpretation type and criterion.

3.4 Automatic extraction of knowledge from artworks and texts

Besides the current work focuses on representing the content of the artwork as descriptive metadata, it is crucial to assess what information can be currently extracted from the artworks and the texts describing them using alternative techniques. As the main object of study, i.e., the artwork, is of a visual nature, we provide a brief overview of the knowledge that can currently be extracted from artworks' pictures through computer vision (CV), machine learning (ML), and Deep Learining (Castellano & Vessio, 2021) techniques. Computer vision is a branch of Artificial Intelligence that focuses on processing visual images to interpret and understand them. Since most CV algorithms are trained on photographs, the task is challenging when applied to

18 https://dati.beniculturali.it/lodview-arco-onto/ontology/context-description/Interpretation.html
19 https://semanticweb.cs.vu.nl/2009/11/sem/
20 http://ontologydesignpatterns.org/wiki/Ontology:Event_Model_F
21 http://purl.org/emmedi/hico

art pieces. Nonetheless, recent research shows growing interest in applying CV techniques to cultural heritage and art analysis, supported by domain-specific datasets and methods. The adoption of ML and DL, has then significantly improved the accuracy and robustness of automated image analysis in this field.

Several attempts have been made to detect objects (Ahmad & Schich, 2023; Gonthier et al., 2019; Y. Kim et al., 2024),[22] subjects, styles, and visual patterns in artworks. Some studies concern the detection of iconography with a focus on Saints' identification in the context of Western Christian Art. Among them, Stork et al. (2021) exploit the art practice of using attributes to identify a saint (i.e., the keys held by a man indicate that the figure is Saint Peter), performing good results in identifying ten saints. Nevertheless, the ambiguity that attributes have (i.e., saints identified by the same object) prevents them from extending the methodology to all the Christian saints. Similarly, Milani and Fraternali (2021) trained a Convolutional Neural Network (CNN) to automatically classify a corpus of artworks according to 19 iconographical classes, mainly identifying Christian characters. Madhu et al. (2019) focused on the facial and body recognition of Mary and Gabriel in the *Annunciation* scene by applying a gender detection algorithm. Another approach, taken by Marinescu et al. (2020) to improve the precision of object detection in paintings, introduces semantic information about the time in which objects were used to correct CV output. Further studies have addressed the central theme of visual pattern detection in artworks. Ufer et al. (2021) developed style-independent research, which detects visual patterns in the overall image or a portion of it, offering the opportunity to search for a combination of patterns. The retrieval of images having similar compositional structure is addressed by Madhu et al. (2023), based on the iconic language of the images afforded by Max Imdhal. Impett (2020) trained an algorithm to detect recurring poses in artworks based on Warburg's attention to gestures and body position in art. The results are shown in the *Gesture Atlas* project,[23] in which figures depicted in artworks are clustered according to gesture similarity. Further research on the the recognition of gestures in artworks was conducted by Zinnen et al. (2024) and Bernasconi et al. (2023). Brandhorst (2022) suggests the use of mixed Iconclass-based retrieval and Computer Vision search to retrieve similar subjects that have the same visual arrangement, an approach adopted in the upgraded version of Iconclass. Other studies focused on the detection and analysis of materials depicted in paintings,[24] or of punches patterns in Medieval panel paintings to support their attribution (Bruegger et al., 2025). Current developments leverage the use of Multimodal Large Language Models to interpret images with the aid of metadata or textual descriptions of art.[25] Martinez Pandiani et al. (2023) focuses

22 We refer to Bengamra et al. (2023) for a comprehensive overview of object detection in art history

23 https://biblhertz.github.io/atlas/

24 Examples are the project 'How Material Came into the Picture (KIKI)", available at https://www.imareal.sbg.ac.at/en/imareal-projects/how-material-came-into-the-picture-kiki/, which investigates the rendering and function of materials depicted in western art of the 14th and 15th Century, and the study by Zuijlen et al. (2021), in which a dataset of 19k paintings was annotated to study the rendering of materials by artists.

25 See Fanelli et al. (2025) and the cited references

on the detection of abstract concepts in visual images (e.g., the different ways in which death is represented), underlying the challenges that the task raises due to the wide variety of ways in which the same concept can be conveyed, influenced by subjective interpretation and cultural variations.

Further research is ongoing on the use of AI for extracting structured data from textual sources, with the aim of extracting domain-specific, structured data which would enhance further data analysis or the retrieval of artworks depicting the same subjects. Jain et al. (2022) performed an open information extraction on a dataset including a wide variety of art historical texts (e.g., art books, exhibition catalogs), showing the challenges that the variety of entities cited in the document raises. The Viewsari project focuses on the creation of a KG involving the extraction of entities from the key art historical text of Vasari's *Vite* (Ondraszek et al., 2024; Santini et al., 2022). Further ongoing studies leverage the combination of information extraction from domain-specific textual sources and CV models to improve a precise and trustworthy content-based artwork retrieval (Stalter et al., 2024). To the best of the author's knowledge, scholarly texts still present characteristics that make the automatic extraction of information challenging.

From the current state-of-the-art, it emerges that the application of AI methods for the retrieval, understanding, and analysis of artworks is getting wide attention, with a growing use of combined techniques. The progress made in identifying icons and subjects through current AI techniques is promising. However, some aspects remain challenging, such as resolving ambiguities in recognizing saints, historically feasible object detection, and the identification of abstract concepts. Furthermore, it seems highly challenging to detect with such an approach the complexity of cultural meanings unveiled by iconological studies. Ongoing research that proposes mixed approaches that integrate ML, CV, information extraction and knowledge graphs, and experiments with Multimodal Large Language Models (MLLMs) show promising results in providing more refined subject identifications. In this landscape, the availability of domain-specific structured data on iconographies, historical objects, and deeper meanings from authoritative sources still plays a key role to advance the state-of-the-art of CV performances for the recognition of iconographical and iconological aspects in works of art.

3.5 Data analysis for art history

Museums are increasingly publishing open source data on their collections. Although art historical studies exploiting computational methods for art analysis remain a minority, the state of the art gradually shows an increase in their adoption. Joyeux-Prunel (2010) collected essays in which quantitative methods were used to study art-related aspects, spanning from the art market to the analysis of taste or canon construction. A decade later, Dressen (2025) underlines the analytical potential that the new availability of data brings in an introductory essay of a volume dedicated to CH data for research.

The major areas in which computational or quantitative approaches are cultural economics and cultural and social art history. In fact, as interdisciplinary fields, they combine quantitative analyses on large datasets of sociology and economy with art topics. The artworks-related inquiries of cultural economics usually involve art market studies, generally focusing on a specific period. Among them, we cite the study by Greenwald (2021) in which the art circulated in France, England, and the United States during the XIX century is quantitatively analyzed. The study provides a robust methodological approach in which domain-specific research questions drive quantitative inquiries. On a similar research line, the Artl@s project collects, from 2009 onward, geolocalized information about art exhibitions provided in catalogs from the 19th and 20th centuries, which allow us to examine the art displayed and the actors involved.[26] For its familiarity with data analysis, cultural economics is also the domain in which many studies exploit semantic web technologies and artificial intelligence. Schich et al. (2014, 2017) perform a network analysis of the movement of notable historical figures, quantitatively seeing some characteristics of the art market, such as the importance of certain centers (e.g., Paris) (Kienle, 2017). Filipiak et al. (2016) and Rother et. al. (2022, 2023) use the Natural Language Processing technique to extract semantic information from provenance datasets and represent them in RDF format.

A popular method applied to art history is network analysis, used in a wide range of studies to detect, among other things, thematic and stylistic borrowings, artwork and artists' movements, and power dynamics in the development of art (Kienle, 2017; Tuscher et al., 2025). Furthermore, image analytics allows us to gain further insight in artworks and their relations. Klinke (2020) underlines the potential that distant viewing has for seeing similarities and patterns in artwork images. Similarly, Manovich proposes cultural analytics applied to images (Manovich, 2015, 2020). Similar interest is witnessed by D. Kim et al. (2014), who quantitatively compare artworks of different periods on some parameters, such as brightness and use of colors, and by Castellano and Vessio (2022), who proposed a deep learning approach capable of clustering artworks according to their content. More studies on images focused on the detection of similarities between printed illustrations to track their cross-cultural evolution and dissemination (Dondi et al., 2020; Malaspina & Zhong, 2017).

Little attention has been paid to the quantitative analysis of iconographical and iconological topics. Li (2021) examines artistic innovations in the Dutch Golden Age by creating a network of artists related through shared subject matters that they depicted in their works, showing how the choice of visual subject is related to the artists' ability, career, and marketing rather than their location and personal network. Sartini (2024) addressed the potential Christian symbolism that still life artworks in the openly available LOD datasets have. Wansink et al. (2016) conducted a cultural analysis of the food depicted in art in 750 paintings, discovering that the food depicted rarely corresponded to the food effectively eaten and available.

26 More details available athttps://artlas.huma-num.fr/en/about/

Despite the growing interest in conducting quantitative analyses for art history, to the best of the author's knowledge, a quantitative investigation of the variance of iconography and its deep meanings had been so far addressed, nor the representation in structured data and a quantitative analysis of the individual art historians' perspectives was addressed.

3.6 Iconography and Iconology statements on the Semantic Web: a survey

Semantic web technologies, which enrich structured data with machine-readable semantics provided by vocabularies, schemas and ontologies, offer an opportunity to express semantically complex information formally. For this reason, they are a suitable means to express fields of study as complex as art-related studies, such as iconography and iconology. Recent years have witnessed a growing interest in linked open data describing Cultural Heritage (Davis & Heravi, 2021).[27] Despite many Cultural Institutions releasing their data only in a simple tabular form, several Knowledge Graphs (KG) address the description of artworks in a more structured, logical form. Some of them, e.g. Wikidata (Vrandečić & Krötzsch, 2014), have a general scope and are created collaboratively, while others, (e.g. ArCo [V. A. Carriero et al., 2019], Zeri & Lode [Daquino et al., 2017]) are generated by conversion of authoritative data from cultural institutions. Other art-related projects exposed RDF datasets to provide data on the socio-cultural context in which artworks were created. An example is the Golden Agent project,[28] which provides a set of LOD datasets containing information about art creation and market of the Duch Golden Age, a particularly prosperous period for art making. To this aim, they extract resources concerning the production and consumption of creative goods in that period from biographical records (Brouwer & Nijboer, 2017), offering extremely valuable research for domain studies.

In this diversified setting, it is crucial to assess the coverage, accuracy, and reliability of the available data to allow their reuse for domain-specific purposes. Therefore, the work in this section provides an evaluation of the coverage of the content represented in visual works over existing KGs, focusing on iconographical and iconological aspects (i.e., artistic subjects and their symbolic and cultural meanings). We survey KG evaluation

27 Section 3.6 reports a collaborative survey conducted in 2022 with Bruno Sartini under the guidance and assistance of our respective supervisors. The research was published in Baroncini, Sartini, et al. (2023) and is reported here in its entirety, as it constitutes a core study of the state of the art related to the doctoral research project. The wording of the text slightly diverges from that of the paper for subsequent revisions. In particular, the current section is based on the introduction of the paper. Section 3.6.2 corresponds to article Section 4, for which S. Baroncini is responsible. Sections 3.6.1 and 3.6.3 refer to the article's Sections 2 and 5, for which B. Sartini is responsible. Section 3.6.4 corresponds to article's Section 6, for which B. Sartini and S. Baroncini are responsible. Both authors contributed equally to the research. © S. Baroncini, B. Sartini, M. van Erp, F. Tomasi, A. Gangemi. Published in Journal of Documentation. Published by Emerald Publishing Limited. The full terms of this license can be seen at http://creativecommons.org/licences/by/4.0/legalcode Creative Commons Attribution (CC BY 4.0)

28 https://www.goldenagents.org/

methodologies and adapt some of their metrics to the domain of knowledge considered. Furthermore, theories concerning the iconographical and iconological (*icon*) domain are reviewed to assess how much KGs cover information about visual items' subject and content description.

The contents of an artwork should be analyzed both isolated, i.e., by identifying relevant characteristics and associating them to features of other artworks (e.g., the study of patterns recurring in different subjects [Warburg, 1999; Wittkower, 1977]). Therefore, the knowledge emerging from an analytic approach is mostly missed when an artwork's content is described just by a general subject term.

The traditional sources of knowledge are natural language descriptions of artworks found in texts. However, texts need knowledge extraction methods to enable further analysis and interlinking, limiting computational reuse of that knowledge (Sartini & Gangemi, 2021). Another problem is the lack of advanced ontologies that provide a detailed semantic form for artwork description data (see Section 3.3). In addition, since iconographical-iconological analysis can potentially involve very different types of cultural objects, often stored by different institutions, the major benefits of storing information about this domain in knowledge graphs include at least:

- The opportunity to answer domain-specific questions through quantitative analysis (e.g., which attributes and meanings were related to the mythological character of Mercury across the centuries?);
- Accessing and querying interlinked information about worldwide objects that could not otherwise be experienced together (e.g., all artworks with political implications stored in different museums worldwide);
- Formally expressing the semantic complexity of the topic (e.g., the levels of meanings of an artwork and its relations to external resources, such as other artworks, texts, etc.).

In this Section, we assess the available data accuracy, reliability, and interoperability in the iconographical and iconological domain of knowledge. Therefore, the main benefit is to provide a clear state of the quality of semantic domain-specific data available online. Other benefits include improving current data reuse following LOD principles and fostering the creation of a shared semantic description framework for iconology and iconography. With this analysis, we show the reusability potential of existing KGs based on defined *icon* requirements. Finally, the main findings of this survey are shown in a landscape (Figure 3.2) in which KGs are positioned according to their performance in the chosen metrics.

3.6.1 Knowledge graph evaluation methods

The reference standard for data quality is ISO 25012,[29] in which 15 data quality characteristics are identified and classified as inherent and system-dependent data quality. Even if only a few data quality assessment methodologies explicitly refer to the standard (Radulovic et al., 2018), its metrics are generally included in all the proposed methods. Although the standard constitutes a reference point, knowledge graphs require specific evaluation methods and metrics. They show differences from traditional relational databases in their structure (graph versus table), the reasoning possibilities that can be applied to them, and their interoperability and interconnections (Janev et al., 2020). Ji et al. (2022) surveys evaluation metrics and methodologies for representation learning, knowledge acquisition, and completion, with additional analyses over temporal KGs and applications developed from them. Paulheim (2017) provides a series of refinements methods to increase the quality of knowledge graphs. Pellegrino et al. (2023) evaluates Cultural Heritage knowledge graphs regarding their suitability for question-answering tasks. Zaveri et al. (2016) proposes a conceptual framework for quantitative and qualitative metrics in evaluating knowledge graphs from a study of more than 100 scholarly publications. Various general metrics for knowledge graph quality evaluation and applications thereof are provided in (Färber et al., 2018). We reuse parts of these metrics, adapting some to focus on the fields of iconography and iconology (see Section 3.6.3). Behkamal et al. (2014) presents a similar study but uses the Goal-question metric paradigm to assess the quality of knowledge graphs. Ringler and Paulheim (2017) also compare several general domain knowledge graphs in their content coverage. It contains interesting reflections, particularly regarding coverage of artistic fields, in which YAGO and DBpedia seem to be the most detailed. Heist et al. (2020) uses coverage and a metric for evaluation, although this work does not mention cultural heritage-related findings. Shenoy et al. (2022) evaluates Wikidata on schema violations and deprecated entities, looking at its history of updates. Freire and Isaac (2020) also evaluate Wikidata's completeness in the description of data related to cultural heritage. To do so, the information contained in it is compared with the information available on Europeana (Isaac & Haslhofer, 2013), which is used as a gold standard for completeness. This study does not mention specific aspects related to iconography and iconology. Issa et al. (2021) offers a thorough study on the completeness metric when evaluating knowledge graphs. Finally, Ruan et al. (2016) introduces the concepts of queriability to knowledge graphs, developing a framework for evaluating *quality in use*, applying it to DBpedia and YAGO. Queriability is a very interesting concept when extracting relatively complex sets of information from knowledge graphs. Such complex relationships might be present in knowledge graphs that describe artworks with high granularity. However, to verify the queriability of the *icon* content, a first assessment of what is currently included in a knowledge graph is needed.

29 https://iso25000.com/index.php/en/iso-25000-standards/iso-25012

In summary, previous work evaluates knowledge graph suitability for some automatic tasks or their content in terms of various metrics that go from completeness to accuracy to quality in use. Some of them focus on specific fields (such as cultural heritage). There is no study yet that evaluates specific aspects related to iconology and iconography in knowledge graphs, which would require a specific evaluation due to the complexity of the information expressed in this domain of knowledge (Baroncini et al., 2021; Sartini et al., 2021). Therefore, the contribution of the current survey is to adapt a selection of general metrics from the literature to the domain-specific needs, adding a newly created metric. As a result, this contribution attempts to provide a domain-specific overview of the quality of available data according to the domain focus of interest and research questions.

3.6.2 Selection of the knowledge graphs

To collect the most representative RDF data about the description of the artwork, we need to consider which kind of cultural objects can represent a visual subject and can have a cultural meaning. Potentially, every image representing a subject that can be invested with a cultural meaning can be considered by an iconographical-iconological interpretation. To narrow down the research in the art history field, we focus our selection on paintings, sculptures, frescoes, visual subjects on coins (numismatics), and illuminations. Therefore, in this survey, we considered graphs containing data on Cultural Heritage, Museums, Libraries (manuscripts' drawings and decorations), and numismatics. In addition, we included general-purpose knowledge graphs likely containing information about artworks such as Wikidata, DBpedia (Auer et al., 2007) and YAGO (Rebele et al., 2016).

We used the following methodology. We first define our object of interest, namely artworks and information about their subject and meaning. Then, we collect the KGs through 1) the analysis of literature concerning a survey or evaluation of CH KGs (Bikakis et al., 2021; Pellegrino et al., 2023; Savnik et al., 2021) and 2) direct search on the web, through a manual keyword search on Google Database Index[30] and other main databases search engines.[31] [32] [33] This led to 56 graphs. These graphs were further pruned according to the criteria of their online availability through a SPARQL endpoint.[34] We considered these criteria fundamental to assessing data that follows the principle of availability and reusability of the Semantic Web (Wilkinson et al., 2016), according to its shared standards.[35]

30 https://datasetsearch.research.google.com/
31 https://datahub.io/
32 https://triplydb.com/
33 https://www.kaggle.com/
34 SPARQL is the query language used to retrieve information from RDF data. SPARQL endpoints are online services linked to specific knowledge graphs that let users query knowledge graphs through SPARQL queries. For additional information about SPARQL and SPARQL endpoints, we refer to https://www.w3.org/TR/rdf-sparql-query/
35 https://www.w3.org/standards/semanticweb/query

Only 27 of 56 graphs were active online, 18 with a SPARQL endpoint. The KGs for which the SPARQL endpoint was not responsive and those without information about subjects were discarded. Consequently, we obtained nine graphs. Table 3.2 gives an overview of the number of artworks having a subject, distinguishing between URIs and literals.[36] [37] This analysis was conducted through SPARQL queries and by consulting the KGs' documentation. The selection process of our analysis highlights how information about cultural heritage is very scarce when considering data that follows Semantic Web principles, as few domain-specific knowledge graphs[38] are available under those conditions. This makes the inclusion of general domain knowledge graphs essential to assess how *icon* aspects are described in the Semantic Web, as the majority of *icon* data is stored in them. From a structural perspective, we would expect the ontological schemas[39] of domain-specific knowledge graphs to describe *icon* information with a higher degree of granularity compared to general ones. This assumption is proved wrong by our results (see Section 3.6.4), as Wikidata performs better than domain-specific KGs.

One critical aspect we encountered while doing this analysis is the proper identification of what is a work of art. While some graphs use a specific class or property to express it (e.g., `fabio:ArtisticWork` in Zeri & Lode), others do not have a unique way to identify it. In some cases, e.g., Wikidata, many specific classes are used, subclasses of a general 'visual work'. In others, e.g., SARI's RDS platform, the class 'Work' corresponds to many different types of cultural objects specified by a controlled vocabulary. Although this granularity in the artwork description is appreciable, it may generate a few issues when approaching the data quantitatively. First, the selection of what is considered an artwork is left to the user, whose subjective decisions may influence this definition. Second, the high number of entities to be included in a SPARQL query can influence the server response.

36 Literals represent basic data types, such as strings, boolean values, and integers. They are not assigned a URI, and therefore, they can only be referred to as the object of a triple and never as the subject. Literals contain unstructured information (such as natural language descriptions) that might require additional processing before being machine-readable. For additional information about literals, we refer to https://www.w3.org/TR/rdf11-concepts/

37 For an overview of the relations considered to identify subjects and other *icon* information, see Section 3.6.3.2

38 With domain-specific knowledge graphs, we refer to a data source modeled according to a domain ontology 'established to capture the domain of interest' that defines them as 'semantically interrelated entities and relations' (Abu-Salih, 2021).

39 We consider the ontological schema as the portions of those ontologies that knowledge graphs actually use as a data model. There exist several general domain schemas, such as Dublin Core https://www.dublincore.org/specifications/dublin-core/, Simple Knowledge Organization System (SKOS) https://www.w3.org/TR/skos-reference/, or Friend of a Friend (FOAF) http://xmlns.com/foaf/0.1/, that are reused in many different knowledge graphs. Domain-specific knowledge graph schema might include specifically developed ontologies in their schema. See the ArCo Ontology (https://w3id.org/arco/ontology/arco) for the ArCo knowledge graph, or the Nomisma ontology (https://nomisma.org/ontology) for the Nomisma knowledge graph

Table 3.1: Classes and properties related to the recognition of artworks (sculptures and paintings if available) in selected knowledge graphs

Name (Abbreviation)	**Artwork (paintings and sculptures if possible)**
ArCo	<artwork> a arco:HistoricOrArtisticProperty
Zeri&Lode (Zeri)	<artwork> a fabio:ArtisticWork
Nomisma	<artwork> nmo:hasObverse <something> <artwork> nmo:hasReverse <something>
Wikidata	<artwork> wdt:P31 wd:Q3305213 (Painting) <artwork> wdt:P31 wd:Q860861 (Sculpture)
RDS Platform (SARI)	<artwork> a gndo:Work <artwork> a gndo:formOfWorkAndExpression
Europeana	<artwork> a http://vocab.getty.edu/aat/300033618 <artwork> a http://vocab.getty.edu/aat/300047090
National Digital Data Archive of Hungary (ND_Hungary)	<artwork> a dcmitype:Image
DBpedia	<artwork> a dbo:Artwork
YAGO	<artwork> a schema:Painting <artwork> a schema:Sculpture

In the context of this study, we selected which classes could be considered artworks from the analysis of the documentation or from data retrieval. We decided to focus our attention on paintings and sculptures when available (if the information present in the knowledge graphs made them distinguishable from other artworks), as they are universally considered as artworks with at least a subject. When paintings and sculptures were unavailable in the studied knowledge graph, we shifted our attention to the most prominent class in the schema that could represent an artwork (as the numismatic items in Nomisma). On the other hand, when the total number of sculptures and paintings was too little for conducting an evaluation (e.g., in SARI's RDS platform), we included in the analysis broader terms, such as prints, illustrations, and graphics. Table 3.1 summarizes classes that define artworks from the selected KGs, along with properties that link information relevant to iconography and iconology.

3.6.3 Evaluation criteria

Following the approach presented by Wang and Strong (1996), we define metrics that go beyond accuracy, as we are interested in i) the coverage of the KGs schemas and their data, ii) the references and interlinking with existing taxonomies that identify subjects in art (Iconclass, Getty), iii) alignments, and iv) linking to external knowledge graphs to foster poly-vocality in art interpretations. These general metrics were adapted to evaluate the specific knowledge domain to obtain a specific quality assessment of domain data. In addition, these metrics acquire a particular relevance for the domain studies, which analyze the relations between cultural objects, their sources, and multiple interpretations. Following the Panofsky theory explained in Section 2.2.2,

Table 3.2: Overview of the artwork subject presence in the selected graphs (as of 01/12/2022)

Short Name	Artwork	Percentage of artworks having a subject (URI)	Average of subjects (URI) defined per artwork	Percentage of artworks having a subject (literal)	Average of subject (literal) defined per artwork
ArCo	2111726	45.86%	1.01	100	1.22
Fondazione Zeri	20082	99.99%	1.19	0	0
Nomisma	566732	21.16%	1.1	0	0
Wikidata	669857	26.76%	3.37	0	0
SARI	339	72.57%	1.11	0	0
Europeana	13861	9.32%	2.38	33.8	1.64
ND_Hungary	11655	0%	0	55.97	6.04
DBpedia	12250	93.93%	5.7	4.8	1.09
YAGO	29324	12.75%	1.02	0	0

we are interested in analyzing whether the current knowledge graphs distinguish between elements that belong to the first, second, and third levels of interpretation. We are therefore looking for clear distinctions when it comes to the description of natural elements depicted in a painting, the recognition of subjects and symbols, and the reflections of the influence of the cultural period in which the artwork was created on the artwork itself and vice versa.

Taking this into consideration, we applied parts of the framework formulated in Färber et al. (2018) in the evaluation of the chosen KGs. This study proposes the possibility of a weighting system applied to each metric according to the importance of the task in the context of the evaluation. In our case, we give more weight to the evaluation criteria referring to the elements addressed the most in the literature of *icon* studies. Specifically, we assign the maximum weight (1) to those criteria that we consider wholly related to iconography and iconology evaluation, 0.8 to those that we consider closely related, and 0.6 to those criteria that we consider partially related. All other criteria are excluded; considering their weight would be 0, they were not computed. Therefore, of all the categories described by Färber et al. (2018), we focus only on column completeness, schema completeness, semantic validity, reference to external vocabularies, and interlinking via owl:sameAs.[40] We adapted all metrics cited above to address the specific tasks of evaluating the *icon* content. As a result of the adaptation, we decided to rename them to address their new specific purpose. Column completeness was changed into Iconographical and Iconological column completeness (IICC), semantic validity became Semantic validity of Iconographical and Iconological triples (SVIIT) schema completeness became Iconographical and Iconological schema granularity (IISG), reference to external vocabularies became References to external taxonomies of art and culture (RETAC) and Interlinking via owl:sameAs became Interlinking of artworks (IA). The differences and specific changes applied to these metrics will be explained in the sub-paragraphs of this section. Finally, we added a new metric to measure intralinking potential for subject comparisons (IPSC).

Table 3.3 summarizes (i) the reused metrics plus the newly created one, (ii) their adaptation to the *icon* field, and (iii) the weight assigned to the metric. We applied these measurements to the knowledge graphs listed in Section 3.6.2. We then grouped these metrics into 2 macro-categories, namely i) structure of the knowledge graphs, which includes IISG, IA, RETAC, IPSC, and ii) content of the knowledge graphs, which includes SVIIT and IICC. The analysis results and the formulas used to calculate the overall score will be discussed in Section 3.6.4.

40 The cited categories will be thoroughly explained in the following part of this section

Table 3.3: Evaluation metrics, the first five criteria are adapted from (Färber et al. 2018), the last criterion is newly developed

Area	Criterium	Adaptation	Weight [0–1]
Content	Semantic Validity	Semantic Validity of Iconographicaland Iconological Triples (SVIIT)	1
Content	Column Completeness	Iconographical and Iconological Column Completeness (IICC)	1
Structure	Schema Completeness	Iconographical and IconologicalSchema Granularity (IISG)	1
Structure	Using External Vocabulary	References to External Taxonomiesof Art and Culture (RETAC)	0.8
Structure	Interlinking via owl:sameAs	Interlinking of Artworks (via variousproperties) (IA)	0.6
Structure	Intralinking Potential for Subject Comparisons (IPSC)		0.6

3.6.3.1 Evaluation Methodology

Of the chosen metrics, three (interlinking of artworks, references to external taxonomies of art and culture, and intralinking potential for subject comparisons) could be processed automatically by analyzing the data, one through an analysis of the schemas of the various KGs (iconographical and iconological schema granularity), and two required qualitative evaluations (semantic validity of iconographical and iconological triples and iconographical and iconological column completeness). For all automatic evaluations, a series of SPARQL queries were launched on the analyzed graph, and some will be listed as examples in the following subsections. For the metrics that required a qualitative evaluation of the content, we extracted random representative samples of the knowledge graphs and evaluated the graphs manually on those samples through annotations.

All annotations were performed by two annotators.[41] In the annotation process, they could express their inability to evaluate the veracity of some of the triples if the information contained in the knowledge graph was unreachable (broken links) or too scarce to fully assess its quality. We used Cohen's Kappa (using quadratic weights) (Cohen, 1960) to measure the agreement score between the annotators. The triples considered invalid by annotators were mutually excluded when computing these agreement metrics.[42] Given the general agreements of the two annotators for all the different samples annotated, as shown in Table 3.4, we decided to average the evaluation scores of the two annotators for both the qualitative categories.

41 One of the annotators was an Art History domain expert.

42 Only 3,3% of total evaluated triples was considered invalid

In the following paragraphs, the metrics and our computations to obtain them are described in natural language and their mathematical formulas.

Table 3.4: Inter-annotator agreement scores as measured by quadratically weighted Cohen's Kappa for semantic validity of iconographical and iconological triples and column completeness per knowledge graph

Knowledge Graph	Semantic Validity	Column Completeness
Yago	1.00	0.65
Nd hungary	0.82	0.62
ArCo	0.77	0.77
Zeri	0.66	0.78
Nomisma	1.00	1.00
Sari	0.78	0.68
Europeana	0.82	0.79
DBpedia	0.89	0.66
Wikidata	0.76	0.90

3.6.3.2 Iconographical and Iconological Schema Granularity

This metric is a re-elaboration of the 'Schema completeness' metric in Färber et al. (2018). Schema granularity aims to verify to what extent the ontologies and vocabularies, and corresponding classes and properties instantiated in the knowledge graphs, cover the domain of interest. In this work, we verify to what extent the schema of the knowledge graph is suited for the complete description of *icon* elements. Based on the comparison of theories of art interpretation discussed in Section 2.2.2, we formulate the following competency questions, which should be possible to apply to a work of art or a part of it (e.g., a detail or a scene of a pictorial cycle) (Uschold & Grüninger, 1996):

- What are the pre-iconographical elements that appear in a work of art?
- What actions are depicted in a work of art?
- What are the subjects of a work of art?
- What are the symbols represented in a work of art?
- What are the stories represented in a work of art?
- What are the allegories represented in a work of art?
- What are the intrinsic meanings associated with a work of art?
- What cultural phenomena are reflected in a work of art?
- What are the corresponding external taxonomies for the identified iconographical terms?

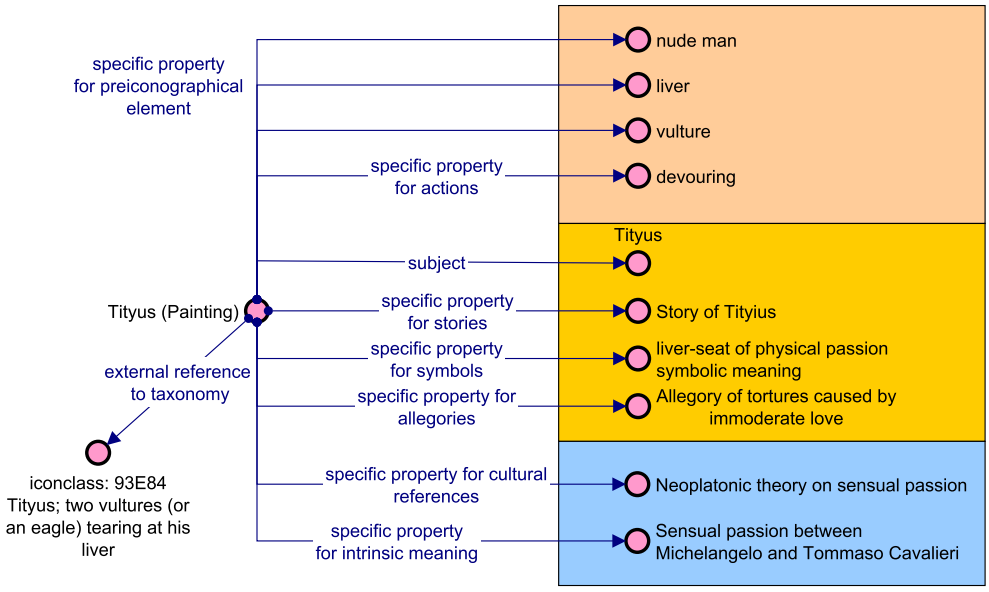

Figure 3.1: The gold standard schema created by the application of CQs to the golden example from the literature, namely, Panofsky's interpretation of *The Punishment of Tityus* by Michelangelo (see Section 2.1)

We then created a gold standard interpretation on the example from Michelangelo's work illustrated in Section 2.1, able to answer those competency questions, as shown in Figure 3.1. We first aligned the properties used in each KG to our example and computed schema granularity as the division between the number of properties of the example that have been aligned and the total number of properties in the example. Given N as the number of properties of the gold standard and N_{akg} as the number of properties of the same gold standard aligned to the properties of the schema of the knowledge graph, we measure the Iconographical and Iconological Schema Granularity ($IISG$) of a knowledge graph as

$$IISG(kg) = \frac{N_{akg}}{N}$$

Table 3.5 shows those properties that were recognized as expressing *icon* content and were aligned to the gold standard.

We weigh this metric as 1 because a schema that permits to express *icon* statements, respecting the required granularity given by the complexity of their field, is essential to correctly and completely store information on this matter.

Table 3.5: Properties identifying iconographical and iconological content for each selected knowledge graph

Name	Iconographic and Iconologic Properties
ARCO	arco-cd:hasSubject arco-dd:hasIconographicOrDecorativeApparatus arco-cd:iconclassCode arco-cd:subject dc:subject
Zeri	fabio:hasSubjectTerm
Nomisma	nmo:hasPortrait nmo:hasIconography nmo:hasControlMark
Wikidata	wdt:P180 (depicts) wdt:P921 (main subject) wdt:P1257 (depicts iconclass notation) wdt:P4878 (symbolizes) (qualifier of wdt:P180) wdt:P6022 (expression, gesture or body pose) (qualifier of wdt:P180)
SARI	gndo:topic gndo:gndSubjectCategory
Europeana	dc:subject
ND Hungary	dc:subject
DBpedia	dc:subject dbp:subject dbp:symbol dbp:symbols
YAGO	schema:about

3.6.3.3 Semantic Validity of Iconographical and Iconological Triples

This metric was modified from the 'Semantic Validity' of Färber et al. (2018), whose purpose is to define whether all the triples statements in knowledge graphs hold true or not. In our study, we consider the semantic validity of *icon* triples only: we evaluate whether triples that refer to a subject, depicted element, or symbol associated with a painting hold true. To evaluate this, we take a subset of the *icon* statements in each KG. Those statements link the artwork to one of the elements relative to the three layers of interpretation explained in Section 2.2.2, agnostic to the property used. We compute this metric by taking a random sample of 100 iconographical/iconological triples from each knowledge graph, evaluating whether the triple is correct (1), partially correct (0.5), or wrong (0). Given S_{ictkg} as the random set of iconographical triples extracted from a knowledge graph, and $S_{evictkg}$ as the evaluation scores set given for each triple $\{sc_1, sc_2 \dots sc_x\}$, and x as the sample size[43] to be extracted from the knowledge graph, the Semantic Validity of Iconographical and Iconological Triples ($SVIIT$) is measured as follows

$$SVIIT(kg) = \frac{\sum_{i \in S_{evictkg}} i}{x}$$

As this metric offers key insights on the quality of the *icon* content of knowledge graphs, which is fundamental for their reuse, we give it a weight of 1.

3.6.3.4 Iconographical and Iconological Column Completeness

This metric, in Färber et al. (2018), considers the general column completeness of knowledge graphs. In our work, we focus only on the column completeness of *icon* statements. Considering the potentiality expressed in a knowledge graph through the iconographical and iconological schema granularity, we evaluate the column completeness as the schema in use. We extract subgraphs from the analyzed KGs that contain all the *icon* triples associated with 100 randomly selected artworks per KG. This evaluation considers two aspects:

1. the expected number of layers of an artwork. Generally, a landscape only contains elements belonging to the first layer, a portrait contains the first layer and then the identification of the subject (second layer), and more complex artworks that represent cultural and religious themes can also be analyzed at a third, iconological level. Despite the potential for every visual image to have a deeper level of interpretation (van Straten, 2012), we decided to expect a third layer only in artworks presenting an explicit cultural subject. This is meant to not affect the artworks' evaluation with the bias of over-interpretation, criticized by some scholars (Gombrich, 1948)

2. the number of layers covered by the current description in the knowledge graph.

43 In our case set as 100

We then divide the covered layers by the expected layers for each artwork in the subset. Having a maximum of three layers, the possible scores for each artwork can be 0 (0 covered layers out of 3 expected, 0/2, 0/1), 0.33 (1/3), 0.5 (1/2), 0.66 (2/3), 1 (1/1, 2/2, 3/3). We do not expect artworks to be described meticulously by indicating every single element of level 1, every single recognizable subject, allegory, and symbol of level 2, and every single intrinsic meaning and culturally related meaning of level 3;[44] for this evaluation, having at least one element for every expected level was considered enough. Given A as the set of the randomly sampled artworks in the knowledge graph of size x[45] $\{a_1 \dots a_x\}$, EL as the array of expected layers (a number from one to three) for each artwork

$$EL = \begin{bmatrix} el_1 & el_2 & el_x \end{bmatrix}$$

in A, and CL as the array of covered layers for each artwork

$$CL = \begin{bmatrix} cl_1 & cl_2 & cl_x \end{bmatrix}$$

we create the array SL that contains the divisions between covered and expected layers

$$SL = \begin{bmatrix} \frac{cl_1}{el_1} & \frac{cl_2}{el_2} & \frac{cl_x}{el_3} \end{bmatrix}$$

and then we measure the Iconographical and Iconological Column Completeness (IICC) of a knowledge graph as follows

$$IICC(kg) = \frac{\sum_{i \in SL} i}{x}$$

We consider this metric as important as having a schema that permits a certain degree of granularity in artwork descriptions; therefore, we give it a weighing of 1.

3.6.3.5 Interlinking of artworks

We adapted the metric 'Interlinking via owl:sameAs' described by Färber et al. (2018) to only apply to artworks. 'Interlinking' is considered as the connection between entities belonging to different knowledge graphs. Although less central than the other used metrics (weight = 0.6), we decided to include it because aligning artworks across different knowledge graphs fosters poly-vocality in art interpretation, especially if these knowledge graphs have been manually curated.[46] We measure this metric by dividing the number of artworks in a knowledge graph that are connected to their

44 Especially considering that in the field of iconography and iconology, there could be potentially endless different interpretations of a painting, and it is not possible list them all

45 In our case set as 100

46 We acknowledge that poly-vocality can be achieved also by giving iconographical and iconological assertions a provenance (even in the same knowledge graphs), although for this work we only focus on statements agnostic to the provenance of the interpretation, which would require another specific study

corresponding versions in external knowledge graphs by the total number of artworks present in a knowledge graph. The main property used to align artwork across different KGs is `owl:sameAs`, but we also looked at other possible alignments from the analyzed KGs.[47]
Given KG as the set of triples $\{t_1 \dots t_n\}$ in a knowledge graph (a triple being a sequence of subject, predicate, object $\{s_i, p_j, o_k\}$), A as the set of artworks $\{a_1 \dots a_m\}$ denoted by s_i or o_k, and R_a as the set of relationships $\{r_1 \dots r_z\}$ that are used to align an artwork in a knowledge graph to the same artwork in other knowledge graphs, we consider $A_a = \{a_1 \dots a_w\}$ as a subset of A if

$$\forall a_i \in A_a : a_i \in A \wedge (\exists p_j \exists o_k : (a_i, p_j, o_k) \in KG \wedge p_j \in R_a)$$

and we measure Interlinking of Artworks (IA) as

$$IA(kg) = \frac{n(A_a)}{n(A)}$$

Two example queries launched on DBpedia to count the number of artworks and the number of artworks aligned to different KGs can be seen in listings 3.1 and 3.2 respectively.

Listing 3.1: SPARQL query launched on DBpedia to count the number of artworks

```
SELECT (COUNT(DISTINCT ?artwork) as ?tot)
WHERE { ?artwork rdf:type/rdfs:subClassOf* dbo:Artwork }
```

Listing 3.2: SPARQL query launched on DBpedia to count the number of artworks aligned to external knowledge graphs

```
SELECT (COUNT(DISTINCT ?artwork) as ?tot)
    WHERE { ?artwork a dbo:Artwork; owl:sameAs ?x
}
```

3.6.3.6 References to external taxonomies of Art and Culture

This metric is a re-elaboration of the 'Using external vocabulary' metric of Färber et al. (2018). In our work, we focus on the use of vocabulary that belongs to taxonomies of art and culture, which play an important role in artwork descriptions as they provide permanent URIs for specific subjects, scenes, and other *icon* elements represented in artworks. Moreover, they are curated by domain experts, and referring to them gives more authoritativeness to the interpretations. For this analysis, we selected four core taxonomies: Iconclass,[48] the Getty Art & Architecture thesaurus,[49] the Getty

47 The link to external artworks is expressed 1) in Europeana through the relations `dc:relation` or `edm:relatedTo`, 2) in Wikidata through different `wikibase:identifier`, 3) in ARCO and Zeri&Lode through `rdfs:seeAlso`, beyond `owl:sameAs`

48 http://www.iconclass.org/help/outline

49 https://www.getty.edu/research/tools/vocabularies/

Iconography Authority vocabulary,[50] and the Getty Cultural Object Name Authority vocabulary[51] (see Section 3.2). We measure references to external taxonomies of art and culture by dividing the number of artworks in a knowledge graph that are associated with at least one of them by the total number of artworks present. Given A as the set of artworks and KG as the set of triples $\{t_1, \ldots, t_k\}$ in a knowledge graph (a triple being a sequence of subject, predicate, object $\{s_i, p_j, o_k\}$) and T as the set of nodes in a knowledge graph representing a particular subject expressed using a taxonomy of art and culture, we consider an artwork part of the subset A_t that contains artworks with a taxonomy reference if

$$\forall a_i \in A_t \colon a_i \in A \wedge (\exists p_j \exists o_k \colon (a_i, p_j, o_k) \in KG \wedge o_k \in T)$$

and we measure the References to external taxonomies of Art and Culture (RETAC) of a knowledge graph as

$$RETAC(kg) = \frac{n(A_t)}{n(A)}$$

The list of taxonomies of art and cultures used for this analysis contains only those that are referenced at least in one of the analyzed knowledge graphs. Increasing the number of taxonomies referenced would not change the evaluation methodology (and its formula). We welcome potential changes to this list to address *icon* aspects of more specific artworks, such as the reference to the Chinese Iconography Thesaurus[52] for a potential analysis on Chinese *icon* statements in the Semantic Web. References to external taxonomies are strictly related to iconography and iconology but are not essential to give a complete artwork description. For this reason, we weigh this metric 0.8.

The query shown in Listing 3.3 was used to count all the artworks in ArCo, referring to a taxonomy of art and culture (Iconclass).

Listing 3.3: SPARQL query launched on ArCo to count the artworks that have a reference to a taxonomy of art and culture (iconclass)

```
SELECT (COUNT (DISTINCT ?s) as ?tot)
    WHERE {
    ?s a arco:HistoricOrArtisticProperty ;
        arco-cd:iconclassCode ?res
}
```

50 https://www.getty.edu/research/tools/vocabularies/guidelines/cona_3_6_3_subject_authority.html
51 https://www.getty.edu/research/tools/vocabularies/cona/
52 https://chineseiconography.org/

3.6.3.7 Intralinking potential for subject comparisons

We introduce this metric to highlight the importance of intralinking subjects in the same knowledge graph. We consider 'intralinking' as the connection between entities belonging to the same knowledge graph. Having a URI as a subject of an artwork allows grouping artworks per subject and comparing them in respect to having a subject as a literal. Moreover, the same subject can then be aligned to other subjects in different knowledge graphs to foster interlinking in the digital art history LOD field. We measure the Intralinking potential for subject comparison by dividing the number of subjects that are linked to more than one artwork by the number of total subjects. Given S as the artistic subjects (expressed as URIs) in a knowledge graph and S_2 as the artistic subjects that are linked to more than two artworks, we measure the Intralinking potential for subject comparison (IPSC) of a knowledge graph as

$$IPSC(kg) = \frac{n(S_2)}{n(S)}$$

As this aspect is relevant but not fundamental for iconographical content representation, we weight it 0.6. Two example queries that count the number of subjects (URIs) in Europeana, and the number of subjects that are linked to more than one artwork can be seen respectively in listing 3.4 and 3.5.

Listing 3.4: SPARQL query launched on Europeana to count all the subjects that are URIs

```
SELECT (count(distinct ?sub) as ?tot)
WHERE {?s a skos:Concept ; skos:broader* aat:300033618 .
       ?s2 a skos:Concept ; skos:broader* aat:300047090 .
    {?CHO dc:type ?s ; dc:subject ?sub .
     FILTER (isURI(?sub))}
    UNION
    {?CHO dc:type ?s2 ; dc:subject ?sub .
     FILTER(isURI(?sub))}
}
```

Listing 3.5: SPARQL query launched on Europeana to count all the subjects that are linked to more than one artwork

```
SELECT (COUNT(?sub) as ?tot) WHERE { FILTER (?tot > 1){
SELECT ?sub (COUNT(DISTINCT ?CHO) as ?tot) WHERE {
    ?s a skos:Concept ; skos:broader* aat:300033618 .
    ?s2 a skos:Concept ; skos:broader* aat:300047090 .

    {?CHO dc:type ?s; dc:subject ?sub .}
    UNION
    {?CHO dc:type ?s2; dc:subject ?sub .}

    GROUP BY ?sub
}
```

3.6.4 Results and discussion

Results obtained from applying the metrics over the KGs are summarized in Table 3.6 and visualized in Figure 3.2. To give a better overview of the results of the metric evaluation, they were then used to place the knowledge graphs inside a two-dimensional landscape. The landscape coordinates are determined by the two macro-aspects, namely content and structure, described in Section 3.6.3. We averaged the metrics relative to these two macro-categories to obtain a score for content and structure. These averages are computed taking into consideration the weights of each metric. Given M_s and M_c as the sets of scores of a knowledge graph relative to its structure and content respectively {IISG, IA, RETAC, IPSC} and {SVIIT, IICC}, WM_s and WM_c as the sets of weights given to M_s and M_c respectively $\{w_{iisg}, w_{ia}, w_{retac}, w_{ipsc}\}$ and $\{w_{sviit}, w_{iicc}\}$ we computed the structure score (SS) of a knowledge graph as follows

$$SS(kg) = \frac{IISG \cdot w_{iisg} + IA \cdot w_{ia} + RETAC \cdot w_{retac} + IPSC \cdot w_{ipsc}}{\sum_{i \in WM_s} i}$$

and the content score (CS) of a knowledge graph as follows

$$CS(kg) = \frac{SVIIT \cdot w_{sviit} + IICC \cdot w_{iicc}}{\sum_{i \in WM_c} i}$$

Table 3.6: Results for each metric over the selected Knowledge Graphs (as of 01/12/2022)

Short Name	SVIIT (weight 1)	IICC (weight 1)	IISG (weight 1)	IA (weight 0.6)	IPSC (weight 0.6)	RETAC (weight 0.8)
ArCo	0.8278	0.74	0.3333	0.0026	0.172	0.1238
Fondazione Zeri	0.9925	0.5117	0.1111	0.0005	0.266	0.5449
Nomisma	0.995	0.5	0.2222	0	0.749	0.0001
Wikidata	0.9768	0.74	0.6667	0.699	0.367	0.157
SARI	0.849	0.3783	0.1111	0.997	0.5	0
Europeana	0.4688	0.236	0.1111	0.0073	0.6122	1
ND_Hungary	0.13	0.5392	0.1111	0	0	0
DBpedia	0.655	0.7242	0.2222	0.994	0.41	0
Yago	0.99	0.4825	0.1111	1	0.1675	0

We divided the graphs into four categories that represent the four quadrants of the landscape, according to their averaged scores, namely: high in content and structure (both scores ≥ 0.5), low in content and high in structure (content < 0.5 and structure ≥ 0.5), high in content and low in structure (content ≥ 0.5 and structure < 0.5), low in content and in structure (both scores < 0.5).

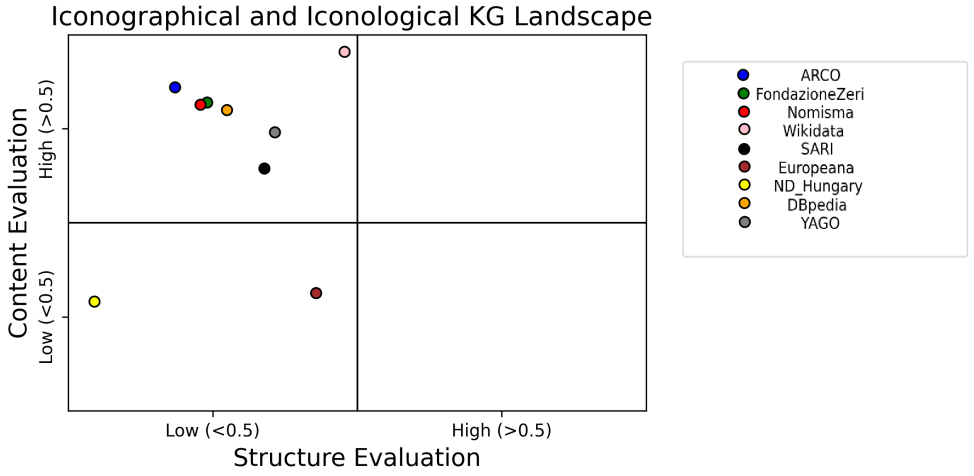

Figure 3.2: Landscape of the knowledge graphs in terms of their iconographical and iconological statements, and the underlying schema structures that describe them

Figure 3.2 shows a clear scenario: the data content is generally correct but not thoroughly described. In fact, none of the graphs has acceptable results in the structure quadrants, and most (7 out of 9) present high scores in content. Nevertheless, this result is given by higher rates in semantic validity (six KGs score more than 0.8) rather than in column completeness (only 3 KGs score more than 0.7). Among them, despite being a general-purpose graph, Wikidata performs the best results. It has the best schema granularity, as several properties can be aligned to the prototype schema of Figure 3.1. In addition, its column completeness scores are higher than some Art History graphs. This is because, in contrast with the approach adopted in the other graphs, the first level of interpretation is often described even when a second or third-level subject is identified.

The granularity in the levels' description may influence the intralinking metric since the description of simpler and more generalizable elements of the first level of description can positively affect the capability of comparing artworks that share them. This assumption is evidenced by the fact that graphs such as SARI's platform,[53] where the subjects considered are broad concepts (e.g. 'persons related to art'), perform better results in intralinking. However, it is important to underline that the general purpose of the graph and the restricted number of subjects described can affect this evaluation. For example, Nomisma,[54] having as subjects only deities, personifications, or Roman emperors, scored the highest in this metric.

53 https://rds.swissartresearch.net/resource/rdsPages:Start
54 http://nomisma.org/

Other relevant qualitative observations can be made over the results obtained. Firstly, we envision that Art History KGs such as Zeri&Lode, which precisely identify second-level subjects with an acceptable percentage of interlinking to vocabularies, could foster subject retrieval and semantic computational capabilities by adding information on more levels of interpretation. Additionally, ArCo, created by automatic conversion of cultural heritage catalogs, despite having a high result in column completeness, has low rates in subject intralinking (0.172) and in relation to external taxonomies (0.123). This may be due to the highly automatic process through which the knowledge graph was created (V. A. Carriero et al., 2019). The automatic creation of URIs for subjects from strings extracted from catalog data could be improved to avoid duplicates of URIs referring to the same entities, therefore increasing the intralinking potential of the KG. For what concerns references to external taxonomies, Europeana shows the best results. In fact, it is possible to retrieve different types of artworks according to the Getty vocabulary category, allowing feasible reusability and retrieval of information for people knowledgeable about them. Moreover, by defining artwork types in this way, it is also possible to retrieve information without having to know specific classes for types of artworks, shifting from the necessity to know the specific schema of the knowledge graphs to the knowledge of general taxonomies applicable to different linked open data datasets. It is interesting to note that, despite having a perfect score in references to taxonomies of art and culture, Europeana does not have any specific property that links an artwork to a taxonomy (it uses `dc:subject`), which decreased the score obtained in the schema granularity metric. Finally, the National Data Archive of Hungary (Fülöp et al., 2005) scores worst in the general categories, given the absence of subjects expressed as URIs, the only use of `dc:subject` to describe *icon* statements, and the complete absence of references to taxonomies.

To exploit the capabilities of interlinking, inference, and analysis of the semantic technologies applied to *icon* study of artworks, reliable, complete, and well-structured data are required. We assess the data quality of current CH KGs that are openly available, online queryable, and have data on artwork subject descriptions. Our results indicate that only a few KGs describe the artwork's iconography and iconology (Section 3.6.2).

This survey gives a critical overview of the complexity involved in the correct and exhaustive creation of domain-specific data. Since the artwork *icon* descriptions are generally correct, the current data can be reliable for data reuse and analysis. Nevertheless, a deeper, more accurate description and a better schema are required to enhance all the expressivity that may lay in them. Whereas *icon* descriptions exist, they are not sufficiently interlinked, searchable, and exhaustively described, showing the need of i) a more extended reuse of existing domain-specific controlled vocabularies; ii) the development of domain-specific ontologies that thoroughly cover iconography and iconology; iii) either the creation of new domain data, formally expressed at a finer granularity or the re-engineering of current data following newly developed

ontologies. As shown in this survey, the quantity and correctness of the data cover only one side of the coin. It is also important to express the newly generated knowledge with the correct schema that respects the granularity and complexity of iconography and iconology.