Chapter 7

Integrative Narrative Networks for Art Interpretation

Sofia Baroncini*, Luc Steels†, Remi van Trijp◇[1], [2]

* Leibniz Institute of European History Mainz, baroncini@ieg-mainz.de
† Venice International University, steels@arti.vub.ac.be
◇ Sony Computer Science Laboratories, Paris Lab, remi.vantrijp@sony.com

Abstract

Recent advances in AI have steadily supplied Digital Humanities with new tools to analyse and contextualise artworks using techniques from computer vision, pattern recognition, ontology engineering, natural language processing, and the semantic web. These tools are important because they may unlock more insightful art descriptions and facilitate access to a vast range of online materials like catalogue descriptions, art history texts and thesauri. However, to obtain the full potential of these tools we need to tackle two issues: (i) how to integrate the fragmented and sometimes contradictory information these various tools provide, and (ii) how to make it much easier for art historians, curators, and artists to use and extend these tools.

This paper explores how these issues can be addressed through the use of Integrative Narrative Networks (INN), a data structure that formally represents the process of understanding in terms of narrative questions and answers, and that integrates information from various knowledge sources such as vision, pattern recognition and language. For this purpose, we present a case study on a Late Renaissance Italian painting by Lorenzo Lotto, which is notoriously challenging for the artist's complex symbology. We built two different scenarios, according to the type of source information leveraged. First, we approached art interpretation by extracting entities from the museum's catalogue entry through Natural Language Processing (NLP), and linking them to Wikidata. In the second scenario, the process of understanding starts using as a source information the artwork itself, and progressively deepens the knowledge about it by answering narrative questions using information from available Knowledge Graphs.

Keywords Art interpretation, AI tools, Integrative Narrative networks, NLP, Knowledge Graphs.

[1]S. Baroncini is responsible for Sections 7.1.2 and 7.3. L. Steels is responsible for sections 7.1.1 and 7.2.1 R. van Trijp is responsible for Sections 7.2.2 and 7.4. All the authors are responsible for the remaining sections. The current paper is an extended and revised version of [5]

[2]Supported by the EU Pathfinder project MUHAI (EU grant 951846) and by the international exchange program of Bologna University.

7.1 Narrative-based Art Interpretation

7.1.1 Understanding as question-answering

Recent advances in AI have offered exciting new possibilities for analysing and contextualising artworks using techniques from computer vision, pattern recognition, ontology engineering, natural language processing, and the semantic web. However, these tools also offer fragmented pieces of insight that still need to be combined into coherent interpretations that are easy to access and extend for art historians and other viewers. In this paper, we propose to address this challenge through **narrative-based understanding** [23].

Narratives are an essential tool of the human sensemaking process [1] in which we take otherwise fragmented events and pieces of information, and turn them into a coherent *plot* – a causal chain or network of events [2]. When interpreting an artwork, the viewer tries to construct a narrative that integrates the visual depictions, memory of past experiences, knowledge of artistic styles, the personal history and prior work of the artist, and general world knowledge, in order to answer a whole series of questions, such as: Who is depicted? What are the characters doing? Why are they doing this and why have these actions been shown? What is the background scene? Why has it been chosen? What perspective has been adopted and why? etc. The viewer also tries to put an artwork in context. Who is the painter? Where and when did s/he paint this work? What was the function of the work? What kind of style was common at the time? These are called **narrative questions** that give the viewer a sense of direction; and answering these questions helps them to drive the plot of their narrative forward.

One major challenge is that we cannot conceptualize the process of interpretation and that of constructing a narrative as a linear pipeline of subprocesses in which one feeds into another, as is common in data-driven AI, because of a paradox known as the *hermeneutic circle* [15]: to understand the whole we need to understand the parts but to understand the parts we need to understand the whole [14]. Instead, narrative-based understanding has to be conceived of as a spiraling process. Starting from an initial examination of some input elements with a lot of ambiguity, uncertainty and indeterminacy, hypotheses of the whole are constructed, which then provide top-down expectations to be tested by a more detailed examination of the same or additional elements, leading to a clearer view of the whole, which then leads back to the examination of additional elements, etc., until the narrative 'makes sense' and resonates with the personal episodic memory of the viewer, reaching a state known as **narrative closure** in literature studies [10, 25].

When considering partial automation or digital support for interpretation we need not only work on the individual knowledge sources that can answer or raise questions, such as pattern recognition algorithms that can detect iconic elements or find the meaning of these elements in a database linking icons and their significance. We also need to work on the integration process, and this requires, as a first, fundamental step a data structure that supports integration.

This paper leverages such a data structure, called an **Integrative Narrative Network (INN)**, previously introduced in [25], which acts as a kind of blackboard on which different knowledge sources write partial descriptions and through which they can consult information provided by other knowledge sources to advance understanding.

This data structure potentially supports several functions: (i) share the information between knowledge sources so that there can be both bottom-up and top-down information flow, (ii) orchestrate semi-automatically the process of calling on different knowledge sources based on which information would be of most interest to enrich the interpretation, (iii) drive a user-interface that allows art historians to organise their work by making it possible for them to

actively invoke knowledge sources, and (iv) drive user interfaces to support the progressively deepening experience of a new artwork.

7.1.2 Case Study

We have already applied the methodology and techniques described in this paper to the interpretation of a contemporary artwork 'Secrets' by the painter Luc Tuymans. This experiment was the subject of an exhibition at the BOZAR cultural center in Brussels in 2011.[3] For that project we focused on integrating results coming from pattern recognition and image processing as reported in [1] and from NLP techniques applied to text from the catalogue and other textual resources.

In this paper, we explore historical artworks from the late Renaissance period by the Venetian painter Lorenzo Lotto and we focus on integrating results from semantic web resources, specifically knowledge graphs, with the outputs of NLP processing. More concretely, we report on a use case for the painting 'Venus and Cupid' by Lorenzo Lotto at the Metropolitan Museum in New York[4] (see Figure 7.1). This painting was chosen as a feasible case study for art understanding since the lack of textual documentation and its complex iconography constitute an interpretative challenge for art historians [11].

At first glance, the observer's attention is driven by some visual indicators, such as the brightness contrast, to the body of the main characters, i.e., a putto (little angel) and a woman. If the observer is familiar with the art themes depicted during the Italian Renaissance, s/he will recognise immediately these figures respectively as Cupid and his mother Venus, thanks to the identification of their attributes. Nevertheless, the actions performed (e.g. Cupid's act of urinating through a wreath, with a playful expression) and the presence of other objects (e.g. the snake, the cone shell, etc.) may lead the observer to raise new questions: 'What exactly is depicted?', 'Why are these objects represented?', 'What is the scene telling us?'.

These questions may be further enhanced by the observer's past experience with other paintings by Lotto, who often depicted symbols or visual anagrams in his artworks. At this point, the observer, particularly if s/he is an art historian, will start to look up the possible meanings of the iconic objects through visual and textual sources and verify if the newly discovered knowledge about it is coherent with the context.

Narrative closure is reached when answers to all these questions are coherently integrated with each other and grounded in external and past experiences. For example, the art historian Keith Christiansen recognises that the curious action made by Cupid and the majority of the objects (e.g. the red cloth, the snake, the myrtle wreath, the incense burner) belongs to the sphere of marital love, concluding that the artwork expresses a wedding wish.[5]

Note that the focus of the interpreter has gone beyond the main focal point of the painting (i.e., Venus and Cupid), to explore secondary details, which are now seen as signs, triggering a deeper interpretation of the scene.

An online user or museum visitor has usually a wealth of information available which s/he can use to understand the artwork, without going through the full hermeneutic process described above. For example, s/he can read the textual apparatus curated by museum's practitioners and art historians to progressively understand the painting. Following the narrative of the text,

[3] https://readymag.com/u3083945729/secrets-guide/

[4] Catalogue entry available at `https://www.metmuseum.org/en/art/collection/search/436918`

[5] A more thorough explanation of this interpretation can be found in the artwork catalogue entry of the Metropolitan museum.

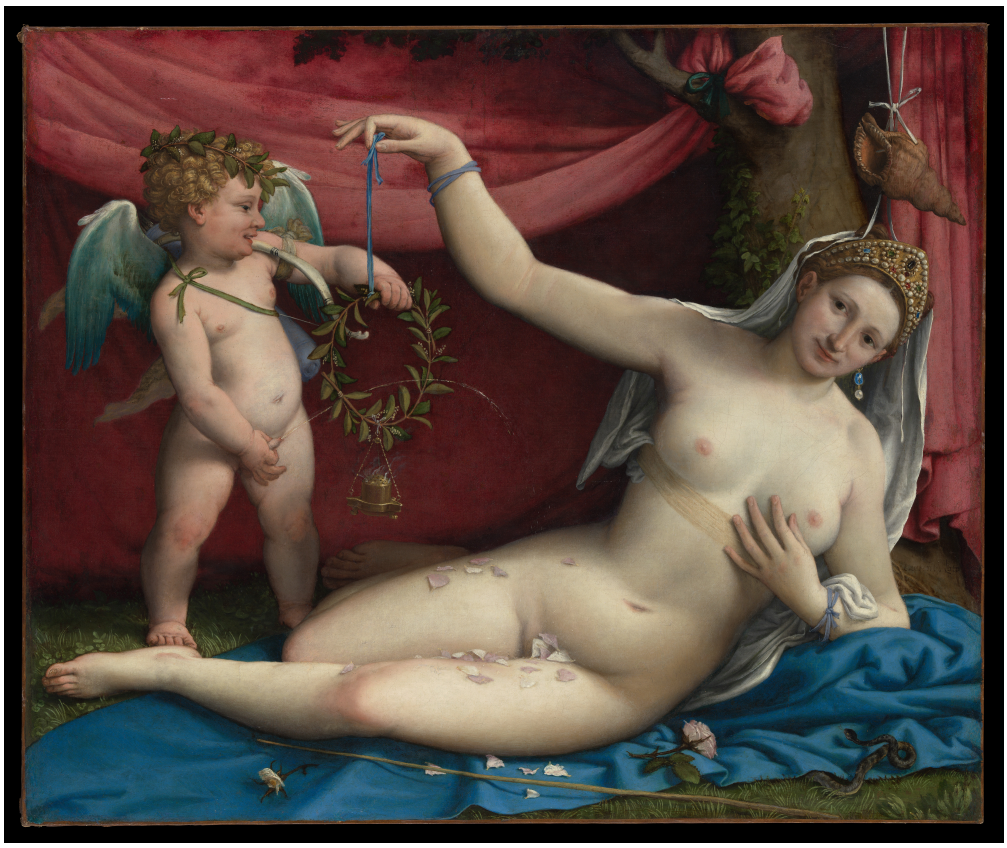

Figure 7.1: Lorenzo Lotto, *Venus and Cupid*, 1520s, Metropolitan Museum of Art, New York. Public domain

s/he will recall the personal knowledge to understand the concepts explained, and observe the painting to relate the textual description with the depicted items. For example, if the text says that the painting shows Venus, the observer will leverage his/her background knowledge to search for a female figure in the artwork, if s/he is familiar with such a subject. In this process, the observer may not know all the knowledge treated by the text, and retrieve more information from external sources, for example through a web source, to answer his/her open questions (e.g., 'what is an epithalamium?'). In this case, the understanding process is guided by the narrative that the text's author adopted in order to effectively communicate to the museum's visitors the knowledge previously acquired. The observer follows the text guide, and occasionally distances him/her from it to relate such information to the observed features of the painting through the aid of background knowledge. Open questions for which the answer is not provided by the text may be answered through access to external sources. The process ends when the observer reaches a satisfactory understanding of the painting and the text.

In the remainder of this paper, we first discuss the technical solutions that we adopted for providing support to the interpretation process sketched out above. Then we illustrate how this can be applied to the Lotto's painting, in both cases in which 1) the knowledge is acquired from the catalogue entry or 2) when no previous knowledge of the painting is available.

7.2 Integrative datastructure (INN)

7.2.1 Narrative questions and answers

As explained, we frame art interpretation as a process whereby the agent raises narrative questions, finds answers to those questions, and interlinks narrative questions and answers. We focus in particular on which knowledge can be automatically extracted from the text and better understood through the aid of semantic web sources, and how these can help to raise questions and find answers to them.

Whereas we refer to the previous work by the authors in which the INN structure was extensively defined [25], [6] we here resume the main features of the data structure to allow the understanding of its application to the art history domain.

What are the questions? Questions are computationally operationalized as variables. Following AI tradition, the name of the question is written as a symbol with a question mark in front, as in ?question-name, where the question-name is a symbol that is chosen on the basis of the nature of the answer to the question. The name is meant to be meaningful for us. For example, we write `?Where` instead of `?Question-540` to be suggestive that this question is about where the artwork was produced. Note however that question-names (and also names for entities) are only formal symbols that obtain their function solely from their connections to other symbols, not from the name they happen to have.

What are the answers? The answers to a question are entities in the domain of discourse. Entities are objects, events or (reified) concepts. They are designated with a symbol, called the entity-id, within angular brackets. The entity-id can be a proper name for a constant (e.g. <LorenzoLotto>) or an instance of a certain type of entity. Entities either refer to real-world observational data (for example a physical painting, a region in an image), to virtual entities (which may or may not exist in reality), or to entities in a knowledge graph in which case we use the URI (Universal Resource Identifier) as unique identifier.

INN makes the relation between a question and an answer computationally operational in terms of a binding between questions (which are technically variables) and identifiers of entities or constants. Bindings are a familiar concept in computer science. Bindings can be set explicitly or indirectly in a pattern matching process (as in logic programming languages like Prolog for example).

Where do questions from? We use a frame-based approach which was pioneered in research on knowledge representation and object-oriented programming, starting in the mid-seventies with the proposal by [18] and technical realisations such as KRL [7], KRS [22] or CLOS [13]. In the present INN implementation we use CLOS. In Minsky's original conception, a frame is a bundle of important questions to be asked about a particular type of object. A frame has a set of slots with values, which are in effect the questions that can be asked about a particular entity represented by a frame and respectively the answers to these questions. An entity which is described with a certain frame is called a frame-instance of that frame.

Where do answers come from? Answers, more concretely default answers, may already be contained in the frame itself. For example, the default value for the ?Period slot in the [renaissance-painter] frame would already be filled with the value <Renaissance>. In the current case study, answers consist of the information that it is possible to retrieve through

[6] The definition of INN is also supported by an open-source software implementation of integrative narrative networks (Apache 2.0 license) as part of the Babel cognitive software suite [17, 24] which can be downloaded at `https://emergent-languages.org/`

queries to knowledge graphs or from NLP information extraction. In the network, there is no specific order in which information must be available to perform a computation, and in the examples the user is free to guide the development of the network.

7.2.2 Interactive Visualization

The open-source software implementation of Integrative Narrative Networks interfaces with the browser-based visualization library vis.js (community edition; `https://visjs.org`) for displaying an INN in a web browser.[7] The implementation also allows users to make these visualizations interactive, which makes it possible to interact with and modify the networks directly in the web browser (e.g. by removing or adding nodes and links).

By default, narrative questions are represented as diamond shapes that are coloured red for unanswered questions, and green for answered questions. Instances, which may be the answers to questions, are represented with squares, representing frame-instances or constants, such as a number or a boolean value.

The user is also free to extend the types of nodes in the network. For instance, directed edges can be used for indicating a particular path in the network (e.g. for representing the plot of the narrative); or different colours can be used for differentiating the type of knowledge source.

Finally, for users who prefer other visualization libraries, networks can be exported so they can be used in other programming environments and network tools.

7.3 Case study for the Lotto painting

This section illustrates the INN and its use in interpretation for a concrete case study on the painting *Venus and Cupid* by late Renaissance painter Lorenzo Lotto introduced earlier (Section 7.1.2). The case study has been implemented in two scenarios of understanding. The first one reconstructs the case in which an observer tries to better understand the painting through the aid of the catalogue entry provided by the museum cutators. In this case, we experiment whether NLP tools can help in extracting relevant knowledge from the text, and connect it to external sources to allow the user further navigation and discovery. The building of the network follows the text narrative. In the second case, the scenario presupposes the sole knowledge of the artwork, and the user conducts an exploratory search about it on available semantic knowledge graphs, progressively extending the network according to the questions answered and information retrieved.

7.3.1 Leveraging the catalogue entry

Recent advances in text processing with various AI techniques (e.g., Natural Language Processing (NLP) and Large Language Models (LLMs)) have shown an increasing potential in text processing to structure information from free text. In the art history domain, a wealth of information about artworks is provided in the form of texts, such as the museums' catalogue entries, scholars' writings, auction books, and archival records. Some studies have shown promising, but still challenging attempts in structuring information provided by art historical texts. Among them, it is worth considering Viewsari, the Knowledge Graph generated from Vasari's *Vite*, the main historiographical source for Early Modern painters in Italy, by extracting main entities

[7] Our paper is supported by an open-source software implementation of Integrative Narrative Networks for Art Interpretation (Apache 2.0 license) as part of the Babel cognitive software suite [17, 24] which is available at `https://github.com/SonyCSLParis/art-network`

such as places, people and their biographies [19]. Further studies extracted knowledge from multiple types of art historical texts with an open information extraction approach [16] and entity linking resolutions between the artwork's description provided by the dataset Artpedia[8] and Wikidata [21]. Although relevant, all these studies highlight the challenges raised by correctly identifying and disambiguating highly domain-specific and context-dependent knowledge and describing it with accurate modeling. Given such current challenges, it is relevant to verify the extent to which AI text processing techniques can be combined with other sources, such as semantic ones, to enhance a meaningful structured representation of the information extracted from the textual sources.

In our study, we experimented with how the knowledge provided by the catalogue entry of Lotto's painting[9] can help in advancing human and machine understanding through the extraction of information from the text and its linkage to existing knowledge bases. The retrieval of more information from knowledge graphs would help users to better understand unknown concepts, and experts in quickly retrieving more information about the same topic. Nevertheless, the quantity and quality of the information that can be collected highly depends on the performance of text-processing tools and on the quality and (specialistic) detail provided by the knowledge sources.

In our study, the text of the catalogue entry was processed through Textrazor[10], a tool combining NLP and AI methods not only to parse and analyze text but also to create semantic metadata, generating mappings with Wikidata entities [26]. The approach adopted consisted of launching the text on the Textrazor interface and examining the results. Textrazor automatically subdivides the text into phrases, and, for each phrase, recognizes several entities, among which are Wikidata identifiers. We first manually corrected the subdivision of the phrases. Following, we checked the results of the Wikidata entities identified, evaluating which alignments were correct.

Each phrase was represented as a node in a graph. The NLP agent seeks to answer, for each phrase, the question 'what is this sentence about?'. Therefore, we relate `?What` nodes to each phrase, which are marked as solved (i.e., visualized in geen) when correct entities are retrieved. In this way, the network represents the narrative with which the discourse is conducted by the art historian who wrote the catalogue entry, and it is possible to identify the important entities repeatedly cited.

The fact that the identified entities are Wikidata ones allows the connection of multiple sources and the exploration of further information in the knowledge base. In particular, we queried the type of Wikidata entities through the relation 'instance of' (`wdt:P31`) or 'subclass of' (`wdt:P279`). Such information is the answer to `?What` narrative questions evoked by the Wikidata entities retrieved. Therefore, we visually represented them as related to question nodes liked to the respective entities.

Results

Out of the 51 text sections identified, 46 could be related to at least one Wikidata entity (see Table 7.1).[11] The 5 phrases for which an entity couldn't be retrieved are related to a red narrative question, as shown in Figure 7.2. The error percentage of the entities retrieved through TextRazor is 9.7% of the entity occurrences. The majority of the errors lie in the

[8] Available at `https://doi.org/10.48610/8a1ccdf` [21]

[9] available at `https://www.metmuseum.org/art/collection/search/436918` in the 'calatog entry' section

[10] `https://www.textrazor.com/`

[11] The complete list of results is made available at `https://inn-painting-art-interpretation.streamlit.app/Catalogue_Entry_Narrative_Network`

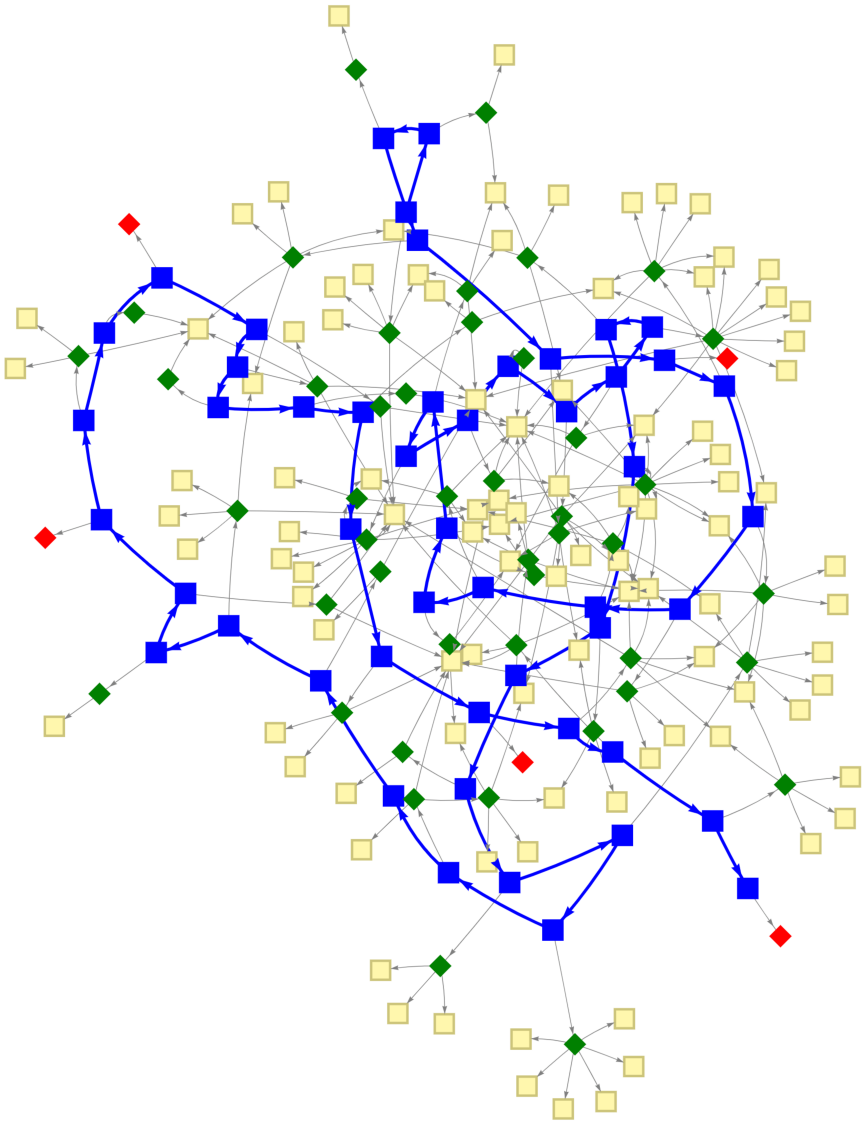

Figure 7.2: Network generated by processing the painting's catalogue entry. Each blue node is a sentence linked to the `?What` questions raised when the entity extraction with Textrazor is performed. The blue edges link one sentence to the next one, highlighting the narrative of the text. When a Wikidata correct entity is retrieved (yellow nodes), the entity is bound to the question node, which color changes from red to green. The figure shows the network after the NLP processing

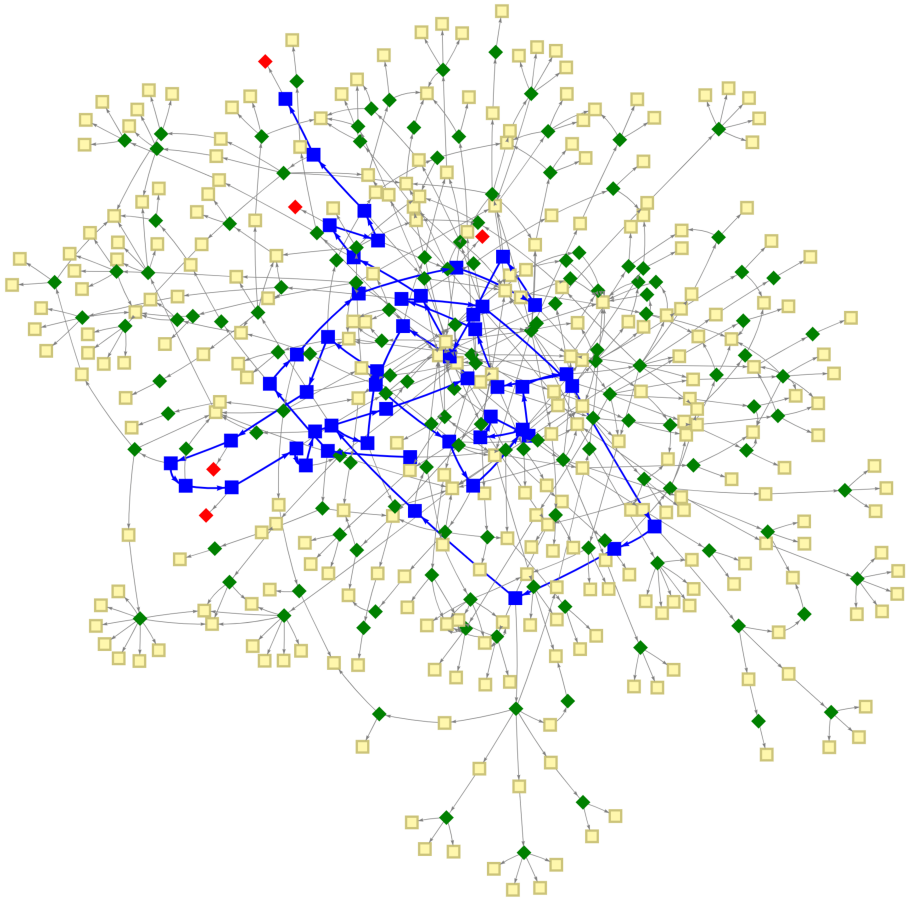

Figure 7.3: Network generated by processing the painting's catalogue entry network after the type of each Wikidata entity is retrieved through the relations `wdt:P31` or `wdt:P279` to answer the further raised `?What` questions

Table 7.1: Overview of the results of entity linking to Wikidata through Textrazor

Type	**Results**
Percentage of correct entity occurrences	90.3%
Total of unique correct entities	100
Total of unique entities (correct and wrong)	109
Phrases without a Wikidata entity	5
Total of phrases	51

wrong identification of the string 'fig.', referring to the image, with the plant 'ficus', or to entities that were too general for the context in which they were used (e.g., the object 'paining' was associated with the concept 'art of painting').

The Wikidata entities majorly linked, as they present the highest number of incoming edges, are 'Venus' (10), 'Cupid' (8), 'Lorenzo Lotto' (8), Venice (6), and 'wreath' (6).

A second iteration retrieved from Wikidata the type of the extracted entities. As all of the entities had at least one type specified, all the `?What` nodes related to them were marked as satisfied (see Figure 7.3). This information gave further hints concerning the nature of the entities treated in the text, showing common types among 35 entities out of the 100 retrieved. The most related one is the node 'human', with 10 incoming edges, showing that at least 10 different people were cited in the text.

Results of the second iteration show that the retrieval of further information from Wikidata is useful not only to better understand the nature of the concepts introduced in the text, but also to identify the ones having a shared type, providing more insights of a quantitative overview of the subjects treated in the text.

In the current scenario, as the main source of information for the understanding process is the catalogue entry, the INN structure is also used as a means of visualizing the narrative of the text, linking each phrase, represented by a blue node, with the next one and with the identified Wikidata entities. In this way, it is possible to understand the central entities of the text, and to reconstruct which textual passages evoke them. The visualization is further enhanced by a dynamic visualization to observe how the network expands phrase by phrase.[12]

The experiment has shown that the integration of NLP tools with semantic web sources allows to deepen the understanding of the entities cited by the text and helps in establishing links among them. Nevertheless, the retrieval of further information was limited to the `?What` question as other information would require a costumed retrieval system according to the type of entity. Furthermore, the experiment would benefit from further testing of other NLP tools and inquiries in how the knowledge retrieved from Wikidata could help to establish further semantic relations among the retrieved entities.

7.3.2 Queries to knowledge sources

In this scenario, the only source of information initially provided is the sole artwork. Consequently, the case study has been implemented in the sense that all the access to diverse knowledge sources is done through exploratory queries to knowledge bases and then integrated into the INN.

We formalised a set of exploratory narrative questions raised during art interpretation with a

[12] Available at `https://inn-painting-art-interpretation.streamlit.app/Catalogue_Entry_Narrative_Network`

Table 7.2: Knowledge bases included in the case study.

Knowledge Base	Description
Wikidata [26]	Open, collaborative KB hosting structured data from other Wikimedia projects
Arco [9]	Italian Cultural Heritage
Zeri&Lode [12]	The photographs about XVI Century art of Federico Zeri's photographic collection
The Iconology Dataset [3]	Iconological interpretations about ca 400 artworks, mainly by the art historian Erwin Panofsky
HyperReal [20]	Symbols and symbolic meanings in different contexts

top-down approach from the didactic art literature[13] [6]. The suggested questions were grouped as sub-questions of five wh-questions, namely *Who? What? Where? When? Why?*. In this way, we proposed a definition of the possible questions that a frame 'visual artwork' may raise when the intention is to interpret the piece. The questions were then expressed as SPARQL queries to be performed over an initial selection of knowledge bases (illustrated in table 7.2). For example, to answer the question about `?Who` (i.e., who is the painter?) the following SPARQL query can be launched to the Wikidata knowledge base[14]:

```
SELECT ?author ?authorLabel
   WHERE {
   wd:Q4009580 wdt:P170 ?author.
   ?author rdfs:label ?authorLabel.
   }
```

Table 7.3 summarizes the questions in natural language that could be expressed as SPARQL queries. The answer from a query (if successful) is then integrated as an answer bound to the question `?Who` in the INN. The criteria for selecting knowledge bases are the possible presence of data about artworks, their content, and/or possible meanings, in compliance with the results obtained by the study in [4]. Furthermore, results obtained from computer vision analysis were integrated into the Narrative Network through a test on Google API Cloud Vision[15].

More concretely, we searched for the painting ID on the KBs through SPARQL queries looking for the artist's name and title or looking for alignments with the ID previously found (e.g. Wikidata's ID). We then performed the queries expressing the 5W questions on each KB. During the process, results were represented according to the INN data structure, showing a graph which progressively expanded from the initial artwork node. Nodes retrieved from different KBs, which were possibly referring to the same concept (e.g., the representation of the character 'Venus' in Wikidata and in Zeri&Lode), were detected through a label similarity fuzzy ratio, and through Wordnet synonyms, and through the presence of common URI to which they were already aligned.

Results

In the initial phase, we retrieved the artwork ID, its title and its eventual image from the selected knowledge bases, starting from Wikidata. The artwork was described only in 3 of them, namely Wikidata, the Iconology Dataset, and Zeri&Lode. The first node, representing the artwork,

[13] University of Manchester, Art historical research. Url: `https://library.wcupa.edu/_art_history`

[14] The SPARQL queries performed are available at `https://inn-painting-art-interpretation.streamlit.app/Narrative_Questions_and_Queries`

[15] `https://cloud.google.com/vision`

Table 7.3: Exploratory narrative questions to interpret a visual work of art

Question	**Sub-questions**
?Who	What is the artist's name?
	What is the artist's movement?
	When did he live?
	Where did he live?
	What other artworks did he do?
?What	What is the type of the artwork?
	What is the genre of the artwork?
	What is the material of the artwork?
	What is the subject of the artwork?
	What is the source of the artwork?
	Are there symbolic meanings?
?Where	Where was the artifact created?
?When	Which is the exact date of creation of the artwork?
	In what cultural tradition or era was it made?
	Are there other historical and cultural events that are possibly related to the artwork?
?Why	Who commissioned the artwork?

was added to the graph as an INN `Artwork` node, which constituted the starting point of the information retrieval. The 5W narrative questions and their sub-questions were added[16]. Figure 7.4 illustrates how the network visualisation appears at this stage. The following sections illustrate in detail the network expansion according to each main question, as information is retrieved from the cited KBs.

?Who This question concerned the author since no information about other involved people (.e.g the patron) was available on the KBs. It was possible to retrieve the author's date of birth and death, places and art movement. Since this information partially answered other open narrative questions, such as `?When` and `?Art movement` (sub-question of `?What`), a relation between them was added. It was also possible to retrieve other paintings by the same author, which is relevant information for art historical research, e.g. for style comparison, or analysing if and how the painter depicted similar subjects.

?What This question considered the retrieval of metadata, such as the title, type, genre, material and art movement of the painting. All the sub-questions found an answer on Wikidata and Zeri. We then answered the narrative question `?Subject`, sub-question in the `?What` variable, which retrieved a great amount of information. In Wikidata, two ways for registering subjects are available, namely specific objects (`wdt:P180` 'depicts') or the main subject (`wdt:P921` 'main subject'). The graph is enriched with the subjects depicted (`wdt:P180` 'depicts'). Whereas Zeri&Lode included only one entity for the subject ('Venere e Cupido'), related to the Wikidata entity thanks to their common alignment to the Iconclass code 92C454, the Iconology dataset provided a thorough description. The identified subjects are here described according to three layers of understanding, showing how more concrete objects (e.g. 'woman') relate to iconographies (e.g. 'Venus') and to deeper meanings (e.g. 'wedding wish'). For answering this

[16]The visualisation of the graph at each stage is available at `https://inn-painting-art-interpretation.streamlit.app/Network_Evolution`

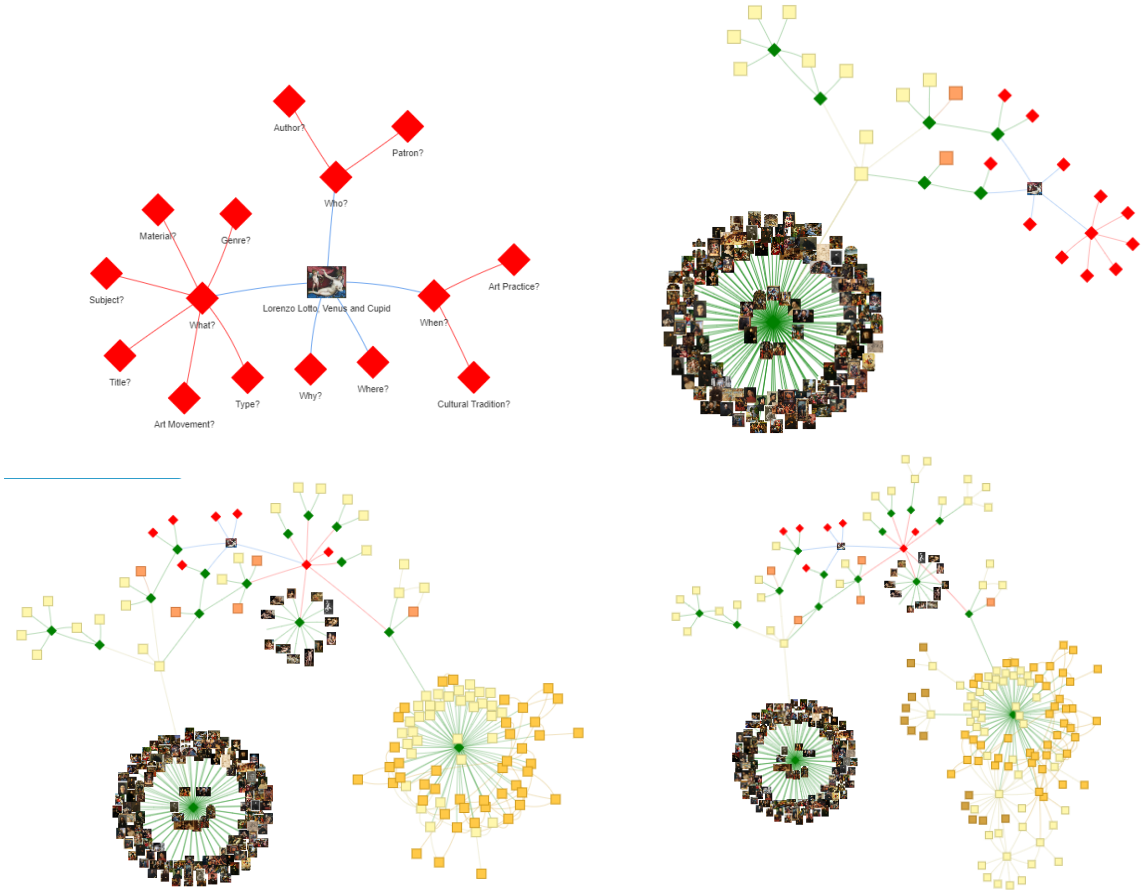

Figure 7.4: How the network appears (a) with the 5WH questions and sub-questions still unbound, (b) after answering the `?Who` question, (c) the `?What` question, and (d) at the end of the information retrieval.

question, also the results obtained by CV analysis were integrated, despite only the concept of 'person' was recognized, and the algorithm erroneously recognized Venus's diadem ad a hat. Following, the sources of the depicted theme were retrieved. In detail, the painting is based on an epithalamium, i.e., a classical genre of poems written to celebrate weddings (source: Wikidata), and various objects and their meanings have specific sources, e.g. the incense burner as a proper bridal chamber decoration has as source the epithalamium of the roman poet Sidonius (source: Iconology Dataset).

?When As anticipated, the retrieval of information about the artist partially answered the question, restricting the range of the possible date of the artwork creation to the one of the author's life. The information about a more precise date of creation was provided by Wikidata (inception 1530) and Zeri (1520-1556).

?Where For this section, the main questions raised for a better understanding of the artwork are 1) the location where the artwork was created, and 2) the place where the artwork was intended to be displayed. Whereas the first question can be potentially answered through the modelling adopted by Zeri, the second one can be retrieved on Wikidata. Nevertheless, no further information was available for Lotto's painting. Therefore, the question is partially answered by the information about the author's work location previously retrieved in the `?Who` question. We consequently added a relation between the narrative questions `?Location` and `?Where`

?Why The narrative question `?Why` had as a main sub-question the retrieval of the patron, namely the person who commissioned the artwork. Although this sub-question can be formally expressed with Wikidata ontology with the relation `wdt:P88`, such information is not available for the chosen case study. Therefore, the narrative question remains unbound.

Second iteration After the first binding of the main information to the narrative questions, we performed a second iteration for answering potential questions raised by the user. For example, a user may be interested in continuing the exploration of a certain KB starting from a retrieved node, or retrieving more possible meanings through the exploration of new KBs. For example, we retrieved more information about Venus, Cupid and their Greek counterparts (Aphrodite and Eros), discovering that they are gods of love and that the latter is the son of the former. Secondly, we explored the potentially embedded symbolism not yet discovered by querying HyperReal. As a result, we included in the INN representation potential symbolical meanings of Venus, the tree and the cone shell, finding, for example, that the conch is not only one of the attributes of Venus but is also a symbol of fertility and gestation, concepts supporting the meaning of the artwork as a wedding wish.

By implementing the questions raised by the literature, the major part of the questions could be expressed as SPARQL queries, performed and answered with currently available information on knowledge graphs. Some answers were bound to multiple questions, as they answered both. Despite multiple information could be aligned through the strategies of reconciliation with a common vocabulary, word similarity, and synonyms detection, challenges in information integration remain open (e.g. when pieces of information slightly differ, as happened with the date of creation). As a solution, we make the INN available as an interactive tool as described in the following section, in which the user can manually select the desired data and perform reconciliations.

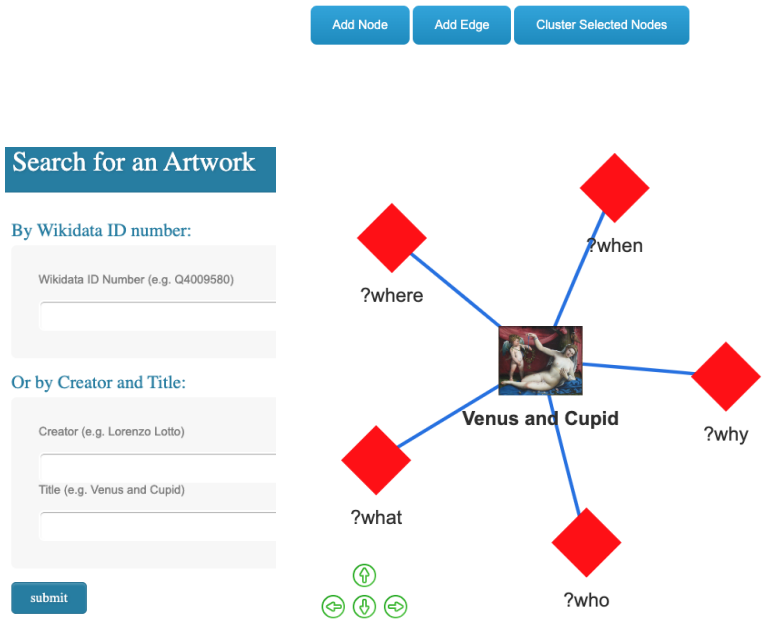

Figure 7.5: This Figure shows an interface for art historians to steer the INN expansion process. The historian can click on certain questions to expand them further and zoom in and out.

7.4 Utilisation of the INN by art historians

The INN is a central data structure around which various applications can be developed. One application, which we have operationalized, is intended for art historians. It allows the historian to direct the expansion of the INN by indicating which questions should be preferentially explored and by choosing which knowledge sources should be utilised preferentially.

Figure 7.5 shows a snapshot of this interface, which is linked to Wikidata. The user starts by searching for an artwork based on its Wikidata ID number; or alternatively, its creator or title. Once an artwork is selected, it becomes the initial node from which the five major WH-questions spring out: `?Who`, `?What`, `?Where`, `?When`, and `?Why`. Based on this initial network, the art historian can develop a narrative that is supported by online knowledge sources by introducing and answering new narrative questions until a satisfactory ending (narrative closure) is reached. By giving the human user full control over the network, other potential issues – such as conflicting or erroneous information – can be solved manually.

Future implementations of this first prototype will consider the inclusion of feedback *to* the knowledge sources. For instance, a specific piece of information about the artwork may be missing. If the user adds a node for that information to their network, it should also be possible to transmit that information to collaborative graphs (e.g. Wikidata) and expand or update that knowledge source. In this way, while users develop their private network, they can simultaneously foster the improvement of collective data quality and the enrichment of knowledge graphs with their expert knowledge.

7.5 Conclusions

This paper introduced an application of Integrative Narrative Networks (INN), a data structure for representing progress in understanding, to the interpretation of artworks in the form of a graph that connects questions and answers for a specific work. Narrative closure occurs when the main questions of relevance to the human interpreter have been answered. The paper

defined the INN and focused on how NLP tools and semantic web resources about art history and interpretation, which are becoming more and more available, can be marshalled to push the interpretation forward. We used a case study of a Late Renaissance painting by Lorenzo Lotto to illustrate the proposed methods and techniques to represent the progressive understanding that can be obtained through 1) the painting's catalogue entry processed through an NLP tool and related to further knowledge available on Wikidata and 2) by retrieving and integrating explorative information about the painting from multiple knowledge bases. Future works include further refinements of the experiments. In particular, we foresee further testing of the capabilities of other NLP algorithms to retrieve further entities and relations from the catalogue entry, and to understand how semantic web sources can be helpful in establishing relations among the extracted entities. Furthermore, we are planning to integrate additional resources, conduct more case studies, expand the capabilities of the user interface for art historians, develop another interface for viewers embedded in an augmented reality device so that viewers can interactively explore an artwork in situ and explore gradually how an 'artificial interpreter' can be built that semi-autonomously expands the Integrative Narrative Network.

Chapter 7: References

[DA1] Sinem Aslan and Luc Steels. Aligning Figurative Paintings With Their Sources for Semantic Interpretation. *IJIMAI*, In Press(In Press), 2023.

[DA2] Mieke Bal. *Narratology: Introduction to the Theory of Narrative*. Toronto Univeristy Press, Toronto, 1985.

[DA3] Sofia Baroncini. *Diving into Artworks Interpretations through the Lenses of Semantic Data: An Application on Panofsky's Iconological Studies*. Doctoral Thesis, Alma Mater Studiorum - Università di Bologna, April 2024.

[DA4] Sofia Baroncini, Bruno Sartini, Marieke van Erp, Francesca Tomasi, and Aldo Gangemi. Is dc:subject enough? a landscape on iconography and iconology statements of knowledge graphs in the semantic web. *Journal of Documentation*, 79(7):115–136, 2023.

[DA5] Sofia Baroncini, Luc Steels, and Remi van Trijp. Semantic Data Retrieval and Integration for Supporting Artworks Interpretation Through Integrative Narrative Networks. In *SWODCH'23: International Workshop on Semantic Web and Ontology Design for Cultural Heritage*, Athens, November 2023. CEUR Workshop Proceedings.

[DA6] Terry Barrett. *Interpreting art: Reflecting, wondering, and responding*. McGraw-Hill New York, 2003.

[DA7] Daniel G. Bobrow and Terry Winograd. An overview of KRL, a knowledge representation language. *Cognitive Science*, 1(1):3–46, 1977.

[DA8] J. Bruner. The narrative construction of reality. *Critical Inquiry*, 18(1):1–21, 1991.

[DA9] Valentina Anita Carriero, Aldo Gangemi, Maria Letizia Mancinelli, Ludovica Marinucci, Andrea Giovanni Nuzzolese, Valentina Presutti, and Chiara Veninata. Arco: The italian cultural heritage knowledge graph. In Chiara Ghidini, Olaf Hartig, Maria Maleshkova, Vojtěch Svátek, Isabel Cruz, Aidan Hogan, Jie Song, Maxime Lefrançois, and Fabien Gandon, editors, *The Semantic Web – ISWC 2019*, pages 36–52, Cham, 2019. Springer International Publishing.

[DA10] N. Carroll. Narrative closure. *Philosophical Studies.*, 135:1–15, 2007.

[DA11] Keith Christiansen. Lorenzo lotto and the tradition of epithalamic paintings. *Apollo (London. 1925)*, 124(295):166–173, 1986.

[DA12] Marilena Daquino, Francesca Mambelli, Silvio Peroni, Francesca Tomasi, and Fabio Vitali. Enhancing semantic expressivity in the cultural heritage domain: Exposing the Zeri Photo Archive as Linked Open Data. *J. Comput. Cult. Herit.*, 10(4), jul 2017.

[DA13] Richard P. Gabriel, Jon L. White, and Daniel G. Bobrow. CLOS: Integrating object-oriented and functional programming. *Commun. ACM*, 34:29–38, 09 1991.

[DA14] H-G.. Gadamer. Hermeneutics and social science. *Cultural Hermeneutics*, 2(4):207–316, 1975.

[DA15] Martin Heidegger. *Being and Time*. SCM Press, London, 1962. Transl. by John Macquarrie and Edward Robinson.

[DA16] Nitisha Jain, Alejandro Sierra-Múnera, M. Lomaeva, Julius Streit, Simon Thormeyer, Philipp Schmidt, and Ralf Krestel. Generating domain-specific knowledge graphs: Challenges with open information extraction. In *TEXT2KG/MK@ ESWC*, pages 52–69. Hersonissos, 2022.

[DA17] Martin Loetzsch, Pieter Wellens, Joachim De Beule, Joris Bleys, and Remi van Trijp. The Babel2 Manual. Technical Report AI-Memo 01-08, AI-Lab VUB, Brussels, 2008.

[DA18] Marvin Minsky. A framework for representing knowledge. In P. Winston, editor, *The Psychology of Computer Vision*, pages 211–277. McGraw-Hill, New York, 1975.

[DA19] Sarah Rebecca Ondraszek, Grischka Petri, Ulrike Blumenthal, Lisa Dieckmann, Etienne Posthumus, and Harald Sack. eXtreme design for ontological engineering in the digital humanities with Viewsari, a knowledge graph of Giorgio Vasari's The Lives. In *SemDH@ESWC*, 2024.

[DA20] Bruno Sartini, Marieke van Erp, and Aldo Gangemi. Marriage is a peach and a chalice: Modelling cultural symbolism on the semantic web. In *K-CAP '21: Proceedings of the 11th on Knowledge Capture Conference*, pages 201–208, New York, 2021. ACM.

[DA21] Alejandro Sierra-Múnera, Linh Le, Gianluca Demartini, and Ralf Krestel. MELArt: A multimodal entity linking dataset for art. *Transactions on Graph Data and Knowledge*, 2(2):8–1, 2024.

[DA22] Luc Steels. Design requirements for knowledge representation systems. In Joachim Laubsch, editor, *GWAI-84*, pages 1–19, Berlin, Heidelberg, 1985. Springer Berlin Heidelberg.

[DA23] Luc Steels, editor. *Foundations for Meaning and Understanding in Human-Centric AI*. Venice International University, Venice, 2022.

[DA24] Luc Steels and Martin Loetzsch. Babel: A tool for running experiments on the evolution of language. In Stefano Nolfi and Marco Mirolli, editors, *Evolution of Communication and Language in Embodied Agents*, pages 307–313. Springer, Berlin, Germany, 2010.

[DA25] Lara Verheyen and Luc Steels. Integrative narrative networks. In Luc Steels and Robert Porzel, editors, *Narrative-Based Understanding of Everyday Activities: An AI Cookbook*, pages 173–198. Venice International University, 2024.

[DA26] Denny Vrandečić and Markus Krötzsch. Wikidata: A free collaborative knowledgebase. *Commun. ACM*, 57(10):78–85, sep 2014.