

# A Wikidata-based Tool for the Creation of Narratives

Tesi di Laurea Magistrale in Informatica Umanistica

*Author:*

Daniele Metilli

*Supervisor:*

Prof. Maria Simi

*Co-supervisors:*

Carlo Meghini  
Valentina Bartalesi Lenzi



Università di Pisa

2016



UNIVERSITÀ DI PISA

*Abstract*

Laurea Magistrale in Informatica Umanistica

**A Wikidata-based Tool for the Creation of Narratives**

by Daniele Metilli

This thesis presents the results of a 12-month-long internship in the Digital Libraries group of the NeMIS laboratory of ISTI-CNR. The focus of the internship was the development of a semi-automated tool for representing narratives in a digital library. This work is part of a larger research effort on the representation of narratives through Semantic Web technologies, including the development of an Ontology of Narratives based on the CIDOC CRM reference ontology. The tool which is presented in this thesis is composed of a web interface which allows the user to create a narrative starting from a specific subject, a triplification component to transform the inserted knowledge into the OWL language, and several visualization components. The tool is based on the Ontology of Narratives, of which it is the first application, and also on Wikidata, an open collaborative knowledge base of recent development, whose entities are imported in the tool in order to facilitate the data insertion by the user. This thesis also reports the challenges that were encountered in the development of a mapping between the CIDOC CRM and Wikidata ontologies. The tool is the first application to make use of the Ontology of Narratives, and will allow an evaluation of it through the representation of a case study, the biography of Dante Alighieri, the famous Italian poet of the late Middle Ages.



## *Acknowledgements*

I would like to thank the following people who directly or indirectly supported the development of this work: Prof. Maria Simi of the University of Pisa, for supervising this thesis; Valentina Bartalesi Lenzi and Carlo Meghini of ISTI-CNR, for accepting me as part of their team, creating the Ontology of Narratives and co-supervising this thesis, but most importantly for being all-around the best people I have ever worked with; everyone else at ISTI-CNR who technically, administratively or otherwise supported me during the development of this work, and especially Paola Andriani and Filippo Benedetti; Prof. Mirko Tavoni and Giuseppe Indizio, for their invaluable help and support in the development of the project; Martin Doerr and the CRM Special Interest Group, for creating the CIDOC CRM and FRBRoo ontologies and advising the development of the Ontology of Narratives; Stephen Stead, for developing the CRMinf ontology; the FORTH, for their contribution to the CRM and CRMinf and for maintaining the CRMsci ontology; Bernhard Schiemann, Martin Oischinger, and Günther Görz, for developing the Erlangen implementation of the CRM and FRBRoo; the Wikimedia Foundation, and especially Danny Vrandečić and Lydia Pintscher, for creating and managing the Wikidata knowledge base; all users and administrators of Wikidata who made this work possible by materially adding data to the knowledge base; Wikimedia Italia, for their continued support for open collaborative projects in the Italian language; my former colleagues at Mart, and especially the archivists who taught me so much; the University of Chicago Library and everyone who supported my previous deciphering endeavors; all the professors who taught me and the students I crossed paths with at the University of Pisa, at the Polytechnic of Milan and at the School of the State Archive in Milan; my family, and especially my mother Franca and my sisters Lucia and Elisa, for always being there when I needed help; my nieces Raffaella and Arianna, I can't wait to play with you again; all my friends in Italy and around the world, I have no space to name you here but you know who you are; and finally, my girlfriend Giulia, I love you very much and thank you with all of my heart.



# Contents

<b>Abstract</b>	<b>iii</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xi</b>
<b>List of Abbreviations</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Background</b>	<b>5</b>
2.1 Narratology . . . . .	5
2.2 Computational Narratology . . . . .	7
2.3 The Semantic Web . . . . .	8
2.3.1 RDF . . . . .	9
2.3.2 RDF Schema . . . . .	9
2.3.3 OWL . . . . .	10
<b>3 Related Works</b>	<b>13</b>
3.1 Narrative Representation . . . . .	13
3.2 Knowledge Base Visualization . . . . .	15
<b>4 Methodology and Requirements</b>	<b>17</b>
4.1 Selection of a Case Study . . . . .	17
4.2 Requirements . . . . .	18
4.3 Implementation . . . . .	20
4.4 Visualizations . . . . .	20
<b>5 The Ontology of Narratives</b>	<b>21</b>
5.1 Components of a Narrative . . . . .	21
5.2 Expression in the CIDOC CRM . . . . .	22
5.3 The Semantic Network . . . . .	26
<b>6 A Collaborative Knowledge Base: Wikidata</b>	<b>31</b>
6.1 Wikidata–CRM Mapping . . . . .	34
6.2 Event Types . . . . .	39
6.3 Roles . . . . .	42

<b>7</b>	<b>A Tool for the Creation of Narratives</b>	<b>45</b>
7.1	Subject Selection . . . . .	46
7.2	Narrative Creation . . . . .	47
7.2.1	First Load and Entity List . . . . .	48
7.2.2	Event Creation Form . . . . .	50
7.2.3	Default Events . . . . .	54
7.2.4	Event Timeline . . . . .	55
7.3	Causal and Mereological Relations . . . . .	56
7.4	Subsequent Loadings and Export . . . . .	57
7.5	Triplification . . . . .	60
<b>8</b>	<b>Visualizations</b>	<b>63</b>
8.1	Timeline Visualization . . . . .	63
8.2	Event-Centered Graph . . . . .	67
8.3	Entity-Centered Graph . . . . .	69
8.4	Primary Sources Table . . . . .	70
8.5	Events in a Specific Time Range . . . . .	72
8.6	Relations Between Events . . . . .	73
<b>9</b>	<b>The Dante Alighieri Case Study</b>	<b>75</b>
9.1	Narrative Representation . . . . .	75
9.2	Limited Evaluation . . . . .	76
<b>10</b>	<b>Conclusions</b>	<b>79</b>
10.1	Future Developments . . . . .	80
10.2	Final Remarks . . . . .	82
<b>A</b>	<b>OWL Schema</b>	<b>83</b>
A.1	Schema . . . . .	83
	<b>Bibliography</b>	<b>105</b>



# List of Figures

5.1	A portion of the <i>fabula</i> . . . . .	26
5.2	The <i>narration</i> . . . . .	28
5.3	The structure of proposition p1. . . . .	29
5.4	Representation of the <i>provenance</i> . . . . .	30
6.1	Flowchart showing the mapping process between Wikidata and the CRM. . . . .	37
7.1	View of the subject selection interface. . . . .	46
7.2	View of the narrative creation interface. . . . .	47
7.3	SPARQL query to extract entities from Wikidata. . . . .	48
7.4	View of the entities container. . . . .	49
7.5	An example of entity popover for the city of Pisa, Italy. . . . .	50
7.6	The popover to add new entities. . . . .	51
7.7	View of the event creation form. . . . .	52
7.8	View of the sourcing popover. . . . .	53
7.9	SPARQL query to extract occupations and positions held for entity <i>Q1067 Dante Alighieri</i> . . . . .	54
7.10	SPARQL query to extract birth and death dates for entity <i>Q1067 Dante Alighieri</i> . . . . .	55
7.11	View of the event timeline. . . . .	55
7.12	Causal and mereological relations. . . . .	56
7.13	JSON representation of the event <i>Birth of Dante Alighieri</i> . . . . .	59
7.14	The process we followed to create the OWL graph. . . . .	61
8.1	Visualization of the events on a timeline. . . . .	63
8.2	View of a digital object in the timeline. . . . .	64
8.3	SPARQL query to extract the knowledge shown in the timeline visualization. . . . .	65
8.4	SPARQL query to extract images from Wikidata. . . . .	66
8.5	SPARQL query to extract links to the English Wikipedia pages describing the entities. . . . .	66
8.6	Visualization of the image of the Baptistery of Florence in the timeline event <i>Baptism of Dante</i> . . . . .	67
8.7	Visualization of the Wikipedia page about Henry VII in the timeline event <i>Dante Meets Henry VII</i> . . . . .	67

8.8	SPARQL query to extract all entities related to all events that compose the narrative. . . . .	68
8.9	Visualization of the event <i>Dante's Education</i> and its related entities. . . . .	68
8.10	SPARQL query to extract all events which are linked to the specific entity <i>Florence</i> . . . . .	69
8.11	Visualization of the entity <i>Florence</i> and its related events.	69
8.12	Table view of the events and their primary sources. . .	70
8.13	SPARQL query to extract all events with their dates and their primary sources. . . . .	71
8.14	SPARQL query to extract all events that occurred from 1265 to 1277. . . . .	72
8.15	Visualization of the events that occurred in a selected time range. . . . .	72
8.16	SPARQL query to extract all sub-events that compose the super-event <i>Education of Dante</i> . . . . .	73
8.17	Visualization of the mereological relations and temporal between events. . . . .	73

# List of Tables

6.1	Mapping between Wikidata and the CRM. . . . .	34
6.2	Roles based on event types. . . . .	43
7.1	Mapping between the tool and the CRM. . . . .	49
7.2	Properties of a JSON event. . . . .	57
7.3	Properties of a JSON proposition. . . . .	58
7.4	Properties of a JSON primary source. . . . .	58



# List of Abbreviations

<b>AI</b>	<b>A</b> rtificial <b>I</b> ntelligence
<b>API</b>	<b>A</b> pplication <b>P</b> rogramming <b>I</b> nterface
<b>ASCII</b>	<b>A</b> merican <b>S</b> tandard <b>C</b> ode for <b>I</b> nformation <b>I</b> nterchange
<b>CNR</b>	<b>C</b> onsiglio <b>N</b> azionale delle <b>R</b> icerche
<b>DL</b>	<b>D</b> escription <b>L</b> ogic
<b>CRM</b>	<b>C</b> onceptual <b>R</b> eference <b>M</b> odel
<b>CSS</b>	<b>C</b> ascading <b>S</b> tyle <b>S</b> heets
<b>ECMA</b>	<b>E</b> uropean <b>C</b> omputer <b>M</b> anufacturers <b>A</b> ssociation
<b>FORTH</b>	<b>F</b> oundation for <b>R</b> esearch and <b>T</b> echnology – <b>H</b> ellas
<b>FRBR</b>	<b>F</b> unctional <b>R</b> equirements for <b>B</b> ibliographic <b>R</b> ecords
<b>FRBRoo</b>	<b>F</b> unctional <b>R</b> equirements for <b>B</b> ibliographic <b>R</b> ecords <b>O</b> bject-oriented
<b>HTML</b>	<b>H</b> yper <b>T</b> ext <b>M</b> arkup <b>L</b> anguage
<b>IRI</b>	<b>I</b> nternational <b>R</b> esource <b>I</b> dentifier
<b>IFLA</b>	<b>I</b> nternational <b>F</b> ederation of <b>L</b> ibrary <b>A</b> ssociations and <b>I</b> nstitutions
<b>ISO</b>	<b>I</b> nternational <b>O</b> rganization for <b>S</b> tandardization
<b>ISTI</b>	<b>I</b> stituto di <b>S</b> cienza e <b>T</b> echnologie dell' <b>I</b> nformazione
<b>JSON</b>	<b>J</b> ava <b>S</b> cript <b>O</b> bject <b>N</b> otation
<b>MIME</b>	<b>M</b> ultipurpose <b>I</b> nternet <b>M</b> ail <b>E</b> xtensions
<b>NeMIS</b>	<b>N</b> etworked <b>M</b> ultimedia <b>I</b> nformation <b>S</b> ystems
<b>OWL</b>	<b>W</b> eb <b>O</b> ntology <b>L</b> anguage
<b>PATHS</b>	<b>P</b> ersonalised <b>A</b> ccess to <b>C</b> ultural <b>H</b> eritage <b>S</b> paces
<b>RDF</b>	<b>R</b> esource <b>D</b> escription <b>F</b> ramework
<b>SPARQL</b>	<b>S</b> PARQL <b>P</b> rotocol and <b>R</b> DF <b>Q</b> uery <b>L</b> anguage
<b>UI</b>	<b>U</b> ser <b>I</b> nterface
<b>URI</b>	<b>U</b> niform <b>R</b> esource <b>I</b> dentifier
<b>URL</b>	<b>U</b> niform <b>R</b> esource <b>L</b> ocator
<b>WD</b>	<b>W</b> ikidata
<b>WQS</b>	<b>W</b> ikidata <b>Q</b> uery <b>S</b> ervice
<b>XML</b>	<b>E</b> xtensible <b>M</b> arkup <b>L</b> anguage
<b>XSD</b>	<b>X</b> ML <b>S</b> chema <b>D</b> efinition



*To my father, Gianfranco (1950–1994)*





# Chapter 1

## Introduction

Since ancient times, humans have always been interested in stories. Listening to stories, telling stories, this is what makes up much of our communication. From myth to parables to the birth of romance, stories have been at the center of human culture. As time progressed, we started to interrogate ourselves on the structure and meaning of these stories. How are they constructed, what are their components, what patterns they follow, all these questions have been asked several times in the history of human culture.

The answers to them are varied, from classical philosophy to Russian formalism, from structuralism to contemporary narratology. In the last few decades, the rise of computer science and the Internet has brought new answers, but also new questions. Today, *storytelling* and *narrative* are popular buzzwords in the Digital Humanities, reaching fields as diverse as journalism, education, libraries, and business communication. However, there is still no universally accepted way to represent a narrative from a computational point of view.

The rise of the Semantic Web<sup>1</sup> and the developments in the Artificial Intelligence field in the last decade have not yet provided clear solutions to this problem. The application of Semantic Web technologies to the field of narratology is still relatively new and unexplored. A standard ontology for narratives does not exist, and there have been few comprehensive approaches to tackle this issue.

This thesis is part of a larger research effort started by Carlo Meghini and Valentina Bartalesi Lenzi at the NeMIS<sup>2</sup> laboratory of ISTI-CNR<sup>3</sup>. The project aims to establish an Ontology of Narratives based

---

<sup>1</sup><https://w3.org/standards/semanticweb/>

<sup>2</sup><http://nemis.isti.cnr.it>

<sup>3</sup><http://isti.cnr.it>

on existing standards such as the CIDOC CRM<sup>4</sup> ontology. Starting from the basic questions (e.g. what is a narrative? What are its components?), a conceptualization of the structure of narratives, and a specification of it based on the CRM, were created. The first results of this work are presented in (Bartalesi, Meghini, and Metilli, 2016)<sup>5</sup>.

Following the creation of the ontology, a semi-automated tool was developed with the purpose of validating it through the representation of a case study: the biography of Dante Alighieri, the Italian poet from the late Middle Ages. However, the tool is by no means limited to this case study. Potentially, it could be applied to any biography or historical narrative, and, with minor modifications, even to fictional stories. The tool is the prime product of a 12-month-long internship in the Digital Libraries group of NeMIS, with the aim of producing applied research based on the Ontology of Narratives.

The tool provides an easy way for users to create and share a narrative using the ontology, through simple drag-and-drop and data insert. Most of the technicalities of the ontology are hidden from view, so as to facilitate the work of a user who might not be particularly accustomed to Semantic Web technologies. The other fundamental characteristic of the tool is the import of data from an existing knowledge base. The knowledge base of our choosing is relatively new and, to the best of our knowledge, it has never been applied in the field of computational narratology.

This knowledge base is Wikidata<sup>6</sup>, a free and open collaborative knowledge base managed by the Wikimedia Foundation. Born as a sister project to Wikipedia, Wikidata has in time become one of the biggest general-purpose knowledge bases in history. Its merger with Google's Freebase<sup>7</sup> has made it even larger and more comprehensive. Wikidata is now mature and stable enough that we feel confident to use it for our purposes. Our tool leverages Wikidata to import knowledge and identifiers, thus freeing the user from the most tedious aspects of the data entry.

---

<sup>4</sup><http://cidoc-crm.org>

<sup>5</sup>A full expression of the ontology in the CRM will be presented in a further publication. The ontology and its development process will also be presented in detail in the soon-to-be-published doctoral thesis of Valentina Bartalesi Lenzi.

<sup>6</sup><https://wikidata.org>

<sup>7</sup><https://developers.google.com/freebase>

After the data entry is completed, a triplification component transforms the knowledge that was entered by the user into an OWL<sup>8</sup> graph, which can subsequently be fed to several visualization modules. The possibilities are endless: for instance, this structured knowledge could be used in digital libraries and archives, in educational settings, in journalism, and in general wherever there is a need to formally represent and visualize a story.

We view the process of building a narrative primarily as a creative work, hence the term *creation* in the title of this thesis. In parallel, we also view it as a *construction*, similar to an engineering work, and, especially when the narrative is built on historical events, also as a *curation* process involving the selection and ordering of events on a timeline. We think our tool provides the means to fulfill all three of these aspects of narrative representation.

The functionalities that our tool offers are by no means unique – several tools have been developed in the past to create or represent narratives – but, to the best of our knowledge, none of the previous attempts empower the user to start from a pre-loaded set of entities, manipulate them in a visual and friendly way, and at the same time produce a proper semantic representation of the inserted knowledge. This is why we think our tool represents a step forward.

In Chapter 2 we discuss the history of narratology, computational narratology, and Semantic Web technologies. In Chapter 3 we describe the current existing tools for the representation of narratives. In Chapter 4 we present the methodology we adopted for the development of the ontology, the tool, and the visualization components, including the requirements that we identified. Chapter 5 presents the Ontology of Narratives. In Chapter 6 we discuss the knowledge base the tool is built on, Wikidata, and the mapping between it and our ontology. In Chapter 7 we provide a full description of the interface and functionalities of the tool, including the triplification phase. Chapter 8 presents a few possible visualizations built on the resulting knowledge base. Chapter 9 describes the representation of the Dante Alighieri case study and briefly touches the subject of the evaluation of the tool, which is still

---

<sup>8</sup><https://w3.org/OWL/>

ongoing. Chapter 10 presents our conclusive remarks and future developments. Finally, Appendix A contains the provisional OWL schema of the Ontology of Narratives.

## Chapter 2

# Background

In order to develop the tool that is the main subject of this thesis, and the Ontology of Narratives it is based on, we first had to analyze past attempts at modeling the structure of a narrative. We will briefly present the history of narratology and its computational applications, and provide a brief introduction to Semantic Web technologies.

### 2.1 Narratology

The discipline that studies the logic, principles, and practices of narrative representation is called “narratology” (Meister, 2012). Antecedents to it can be found as early as Aristotle’s theory of aesthetics (Halliwell, 1987), where a narrative is defined as the imitation of real actions (*praxis*) that forms an argument (*logos*) whose fundamental units, or events, can be arranged in a plot (*mythos*) (Kenny, 2013).

Modern narratology has its origin in the Russian formalism of the early 20th Century. According to this school of thought, a universal pattern of codes, or common literary language, can be found in any narrative work, independently of the medium. The fundamental work of Propp (1973) proposed a model to represent folktales as combinations of basic building blocks, including thirty-one “narrative functions” and seven roles, or “spheres of action”, of the characters.

Russian formalism identifies two elements of a narrative: the *fabula*, defined as a series of events taking place at a certain time at a specific location, and the *syuzhet*, which is the particular way the story is narrated. While the order of the fabula is chronological, the order of the syuzhet corresponds to the way the events are presented in the narrative by its author (Propp, 1973; Shklovsky, 1965).

The other movement that greatly influenced the field of narratology in the 20th Century is structuralism. Drawing on Russian Formalism and on the work of De Saussure (1989), Levi-Strauss (1963) outlined a grammar of mythology. Todorov (1969) coined the term *narratologie* and argued that narratives usually follow a path from an initial state of equilibrium to a different state of equilibrium, through a disruption phase.

Greimas, McDowell, and Velie (1983) proposed a system of six basic structural elements of narratives called *actants*; Genette and Lewin (1983) codified a system of analysis that studied both the narration and the act of narrating, considering them separately from the story and content of the text. They also identified five concepts that characterize the syntax of narratives: order, frequency, duration, voice, and mood. Labov and Waletzky (1997) developed a model for the analysis of narratives in face-to-face interactions.

It should be noted that a universally accepted definition of narrative has not yet been found (Ryan et al., 2007). For instance, in addition to the *fabula* and the *syuzhet*, Bal (1997) defines a third level that constitutes the concrete representation of the content that is conveyed to the audience (e.g., the text in a novel). On the other hand, Crawford (2012) posits that a narrative is a high-level structure based on causality, not on temporal or spatial relations.

Since the 1980s, post-structuralist perspectives on narratology have emerged. In particular, cognitive narratology considers the narrative as a psychological phenomenon, which should be studied from a cognitive perspective (Herman, 2000). A precursor to this approach can be found in (Barthes, 1966), who posited that the textual nature of a narrative should be studied together with its cognitive nature, i.e. the interpretation of it by the reader.

According to Georgakopoulou (2007), the study of narrative activity should shift focus to small stories, which “are presented as part of a trajectory of interactions rather than as a free standing, finished and self-contained unit”. Ochs and Capps (2009) developed a dimensional approach to the definition of narrative, positing that a narrative can be “examined in terms of a set of dimensions that a narrative displays to different degrees and in different ways”.

Herman (2011) proposes a definition of narrative based on four elements: (i) situatedness (the narrative is situated in a context), (ii) event sequencing, (iii) world disruption (the events introduce a disequilibrium in the world), and (iv) the experience of living through the story. Hyvärinen (2006) describes the spread of narrative across several disciplines, starting from the 1960s to the present day. This has been called by Kreiswirth (2000) the “narrative turn”.

## 2.2 Computational Narratology

Computational narratology studies narratives from a computational perspective. According to Mani (2012), its focus is on “the algorithmic processes involved in creating and interpreting narratives, modelling narrative structure in terms of formal computable representations”. Its aim is to develop Artificial Intelligence systems for reproducing human-like narrative behaviour and intelligent interfaces and game environments for interacting with narratives (Mani, 2014).

Harrell Jr (2007) states that computational narratology provides “techniques from computer science to provide a language to describe cognitive insights and to implement narrative effects”. According to Cavazza and Pizzi (2006), computational narratology is focused on story generation systems, i.e. any computer application that creates a written, spoken, or visual presentation of a narrative. Indeed, the aim of this thesis is the development of a semi-automatic tool that allows users to construct a narrative, on top of the ontology we developed.

From a cognitive psychology point of view, Frixione and Lieto (2012) and Lieto and Damiano (2014) highlight the fact that many common-sense concepts are difficult to characterize in terms of necessary/sufficient conditions, and propose a hybrid model that includes prototypical knowledge. However, commonly-used monotonic description logics are not sufficient to represent knowledge in prototypical terms.

## 2.3 The Semantic Web

Since the early 21st Century, the advent of the Semantic Web has profoundly impacted the field of narratology. The aim of the Semantic Web is to advance the state of the Web through the use of semantics (Berners-Lee, Hendler, Lassila, et al., 2001). Specifically, it proposes the usage of semantic annotations to describe the meaning of the information found on the Web, making it fully accessible to machines (and not just to humans).

The Semantic Web also allows the formal specification of entire application domains, and the sharing of their specification on different systems, through the use of *ontologies*. According to Guarino (1998), “an ontology is a logical theory accounting for the intended meaning of a formal vocabulary, that is, its ontological commitment to a particular conceptualization of the world”.

According to Gruber (1993), the main components of ontologies are:

- Individuals: basic objects
- Classes: sets, collections, or types of object
- Attributes: features or parameters that objects can have
- Relations: logical bonds that relate objects to one another
- Events: the changing of attributes or relations

Other components include: (i) axioms, i.e. assertions in a logical form, (ii) rules, i.e. if-then statements that define the logical inferences that can be derived from assertions, and (iii) restrictions, i.e. limitations or ranges of valid values (Sikos, 2015).

Once an ontology has been created, it can be populated, i.e. a knowledge base containing instances of classes and relations and encoded in a Semantic Web language can be built on top of it.

Ontologies are commonly encoded using ontology languages, i.e. formal languages that allow the encoding of knowledge about specific domains and often include reasoning rules that support the processing



of that knowledge. The current web standards for semantic annotations are RDF<sup>1</sup>, RDF Schema<sup>2</sup>, and OWL<sup>3</sup> (Antoniou and Doerr, 2007).

### 2.3.1 RDF

RDF (Resource Description Framework) is the basic language of the Semantic Web. It expresses information about resources by using triples, i.e. statements in the format *subject-predicate-object*.

In order to uniquely identify resources, each of them is assigned an International Resource Identifiers (IRI)<sup>4</sup>. For instance, the IRI for the resource representing the city of Florence, Italy in the Wikidata knowledge base is <https://wikidata.org/entity/Q2044>. Resources with IRIs can occupy both the subject and object positions in a triple.

Basic values that are not assigned IRIs are called *literals*, and they are associated with a datatype that enables their correct parsing and interpretation, and an optional language tag. For instance, a property could be used to state that the name of Florence in the Italian language is “Firenze”, with datatype `xsd:string`<sup>5</sup>. Literals can only occupy the object position in a triple.

The predicate position in a triple is always occupied by a *property*, i.e. a resource (identified by an IRI) that expresses the relationship existing between the subject and the object of the triple.

### 2.3.2 RDF Schema

RDF Schema is an extension of RDF which introduces the concept of *class*, i.e. “a group of individuals which belong together because they share some characteristics” (Doerr, 2003). A class is thus a collection of objects, and each of these objects is declared as an instance of the class through the *instanceOf* property, e.g. *Julia* is an instance of *Human*.

---

<sup>1</sup><https://w3.org/TR/owl2-overview/>

<sup>2</sup><https://w3.org/TR/rdf-schema/>

<sup>3</sup><https://w3.org/OWL/>

<sup>4</sup>IRIs are a generalizations of Uniform Resource Identifiers (URIs), allowing non-ASCII characters in the character string. In the following, we will refer to IRIs and URIs indistinctly.

<sup>5</sup><https://w3.org/TR/xmlschema11-2/>

Classes can be organized in hierarchies through the *subClassOf* property. For example, *Human* could be a subclass of *Mammal*. Properties can also be organized in hierarchies through the *subPropertyOf* property. For example, *hasBestFriend* could be a subproperty of *hasFriend*.

RDFS also allows the definition of a domain and a range for each properties. For example, we could define a property *hasRead* that connects an instance of class *Human* to an instance of class *Literary Work* that the human has read during her life, e.g. *Julia hasRead "Alice in Wonderland"*.

### 2.3.3 OWL

OWL (Web Ontology Language)<sup>6</sup> is the current standard ontology language of the Semantic Web. OWL is built on top of RDF and RDFS and is fully compatible with them. It comes in three different profiles: OWL Lite, OWL DL, and OWL Full.

**OWL Lite.** OWL Lite introduces the possibility to represent equality, i.e. an individual is the same as another individual, through the *sameAs* axiom. This can also be applied to the equivalence of classes and properties. Complementary to this feature is the faculty to describe differences through the *differentFrom* property.

OWL Lite also allows the definition of: (i) the inverse relationship, e.g. property *eats* is the inverse of property *isEaten*; (ii) the transitive property, e.g. if the morning *occursBefore* the afternoon and the afternoon *occursBefore* the evening, then the morning *occursBefore* the evening; (iii) the symmetric property, e.g. if *Julia hasSibling Mary*, then *Mary hasSibling Julia*.

Another important feature of OWL Lite is the possibility to define restrictions on the cardinality of properties, i.e. how many values should a property have for each instance (limited to 0/1 in OWL Lite), and also on the domain and range of properties depending on the class of the individual the property is applied to.

---

<sup>6</sup><https://w3.org/TR/owl-ref/>

**OWL DL.** OWL DL contains all the features of OWL Lite, and additionally it allows the representation of: (i) disjoint properties, (ii) arbitrary Boolean algebraic expression, (iii) negation, (iv) the *oneOf* construct, (v) restrictions on values a property can have. Furthermore, it also allows the definition of arbitrary cardinality constraints.

**OWL Full.** OWL Full is more liberal than OWL Lite and OWL DL in that it allows equivalence between classes and individuals, e.g. *Essay* could be both a subclass of *Type of Text* and an instance of the same class. Furthermore, OWL Full does not include any of the limitations on cardinalities and constraints that are present in the other two profiles. For instance, cardinality restrictions can be applied to transitive properties, while in the other profiles this is not allowed.

**OWL 2.** The second edition<sup>7</sup> of the OWL language adds a few missing features<sup>8</sup>, such as keys (i.e. properties that uniquely identify a specific named instance of a class), property chains (i.e. properties can be composed of other properties), reflexive properties (e.g. an object is always *part of* itself), asymmetric properties (e.g. if object A *is smaller than* object B then it is impossible that object B *is smaller than* object A), and disjoint properties (e.g. the *mother* of a person must be different from her *child*). OWL 2 maintains the three previous profiles Lite, DL, and Full, but it also defines three new profiles<sup>9</sup> which are more specialized in nature.

---

<sup>7</sup><https://w3.org/TR/owl2-overview>

<sup>8</sup><https://w3.org/TR/2012/REC-owl2-new-features-20121211/>

<sup>9</sup>The description of these profiles is beyond the scope of this thesis, but a full description of them can be found here: <https://w3.org/TR/2012/REC-owl2-profiles-20121211>



## Chapter 3

# Related Works

In the last two decades, several tools have been developed in order to computationally represent narratives. Most of them are centered on the semantic annotation of texts, allowing a user to upload and annotate a text in order to give it a semantic representation. In this Chapter we present a brief overview of the main works that were of interest for the development of our tool. We also include a few which create historical visualizations based on existing knowledge bases, which were of interest for the development of the visualization components presented in Section 8.

### 3.1 Narrative Representation

Franzosi (1998) presents a set of linguistic and statistical tools, based on semantic grammars, that can be used by historians to study events (such as strikes, demonstrations and other types of collective conflict). These tools are built on top of relational database systems, and they are used to perform complex data analysis such as network models<sup>1</sup>.

Alani et al. (2003) describe the development of an ontology for the domain of artists and paintings. They developed information extraction tools and techniques to automatically populate the ontology with information from online documents, based on the given ontology and WordNet lexicons. They also developed narrative construction tools to query the knowledge base for relevant facts or textual paragraphs, in order to generate biographies.

---

<sup>1</sup>An outstanding application of the tools developed by Franzosi in the field of Medieval Studies has been carried out by Fernández-Aceves (2013).

Uhlir and Falc (2004) present a generic semi-automatic language-independent framework for annotating content using ontologies as the controlled vocabulary for annotations and Conceptual Graphs<sup>2</sup> as the annotation language. The Dynamic Narrative Annotation Tool (DNA-t) allows the creation of semantic annotations. When some words are selected in the annotated text, the tool can automatically draw concepts as nodes of the annotation graph. The user then finds an appropriate conceptual type in the ontology browser of the tool and assigns it to the newly created concept. The resulting annotation can then be exported into several formats.

Zarri (2009) presents a novel approach to the representation and management of narrative information, through the use of a newly defined language called Narrative Knowledge Representation Language. This language is able represent the narratives themselves and also the templates, or story schemes, that constitute the overarching structures of narratives.

Elson and McKeown (2009) present an annotation tool that allows users to define formal propositions to represent spans of text, as well as temporal relations and other aspects of narratives. A built-in natural language generation component regenerates text from the formal structures, which eases the annotation process.

Lombardo and Damiano (2010) describe an annotation schema for the narrative features of media objects which relies on a formal theory of story and characters, along with the Cinematic tool for the annotation of video objects and the automatic editing of the annotated objects.

The Story Workbench (Finlayson, 2011) is an annotation tool that provides a generic platform for text annotation. It provides a number of common text annotation operations, including representations (e.g., tokens, sentences, parts of speech), functions (e.g., generation of initial annotations by algorithm, checking annotation validity by rule, fully manual manipulation of annotations) and tools (e.g., distributing texts to annotators via version control, merging doubly-annotated texts into a single file).

---

<sup>2</sup><http://jfsowa.com/cg/>

PATHS (Personalised Access to Cultural Heritage Spaces)<sup>3</sup> is a system that acts as an interactive tour guide through existing digital library collections. Navigation is based around the metaphor of a path, which guides the user through the collection and can be based around themes, time periods, people, or places. The system allows users to construct their own paths or follow predefined ones.

Semantator (Song, Chute, and Tao, 2012) is a semi-automated tool for document annotation with Semantic Web ontologies. Starting from a text document and an ontology, Semantator supports the creation and deletion of ontology instances for any document fragment, and the linking and disconnection of instances with the properties in the ontology.

Agosti et al. (2013) present CULTURA, a tool to enrich cultural heritage collections through community-created guided paths called *narratives*.

Storyspace (Wolff, Mulholland, and Collins, 2012) is a software which allows the description of event-based stories from objects preserved in museums. It can be used to create multiple narratives from a single museum exhibition.

Niehaus et al. (2014) introduce SONNET, a flexible narrative framework to integrate, select, and reuse narrative models, thereby lowering development costs and improving benefits from each model. The framework includes a lightweight ontology language for the definition of key terms and relationships among them. The framework also specifies model metadata in order to allow developers to discover and understand models more readily.

## 3.2 Knowledge Base Visualization

Hirsch, Hosking, and Grundy (2009) present two applications, Thinkbase and Thinkpedia, which produce visualizations of the semantic knowledge contained in Freebase and Wikipedia, respectively, that allow the user to explore the semantic graphs of the two knowledge bases in an accessible and interactive way.

---

<sup>3</sup><https://sheffield.ac.uk/is/research/projects/paths>

Histropedia<sup>4</sup> allows users to create or view timelines on topics of their choice by importing statements from Wikidata. Links to related Wikipedia articles and Wikimedia Commons images are automatically added, resulting in rich spatio-temporal visualisations. The scope of the project includes research, education, academic research, tourism, and proprietary applications (Mietchen et al., 2015).

Chronas<sup>5</sup> is a chronological and cartographical history application, with a special focus on visualizing maps of the world across human history. The knowledge visualized in the application is automatically imported from Wikipedia, but manual additions and modifications by the users are allowed.

Histogramy<sup>6</sup> is a web application that visualizes events imported from Wikipedia, spanning the entire history of the universe. The project is updated daily with new information from Wikipedia. The user can focus on a particular period of history or even a specific event.

---

<sup>4</sup><http://histropedia.com>

<sup>5</sup><http://chronas.org>

<sup>6</sup><http://histogramy.io>



## Chapter 4

# Methodology and Requirements

We can now present the methodology that we adopted during the development of our project. First of all, the project can be divided into three main phases:

1. creation of an ontology for representing the knowledge about narratives, encoded in the OWL language;
2. development of a tool to populate the ontology in an easy and user-friendly way;
3. development of a web interface for visualizing narratives.

As explained in the previous chapters, this thesis is focused especially on the second phase (development of a tool to populate the ontology). The results of the first phase (creation of the ontology) are presented in Chapter 5, while the initial results of the third phase (visualizations) are briefly described in Chapter 8.

### 4.1 Selection of a Case Study

We initially selected a case study in order to validate the results of each phase of our work. The narrative that we selected is a biography that describes the life of Dante Alighieri, the Italian poet of the late Middle Ages. This biography was written by an authoritative scholar, who collaborated with us during the course of the project and was available to provide useful suggestions and evaluate our work.

## 4.2 Requirements

After selecting a case study, we decided to analyze the requirements of our tentative users. As specified in the ISO 9241-210 standard (Human-centred design for interactive systems)<sup>1</sup>, the understanding of the needs and requirements of the users is the first step in order to develop successful systems and products. The results of this analysis can include increased productivity, enhanced quality of work, reductions in support and training costs, and improved user satisfaction.

The analysis of requirements was accomplished through semi-structured interviews with historians, experts on Dante's life and works, and students. During the course of these interviews, we covered the needs and aspirations regarding the representation and creation of narratives about the life of Dante, how well the current systems meet those needs, and possible improvements that could be made.

After the interviews, an initial set of user requirements was developed for the three phases of our project. For each of the three phases, we reported the related user requirements we elaborated based on the analysis of requirements and a review of prior literature.

1. Creation of an ontology for representing the knowledge on narratives, based on Semantic Web technologies.
  - Representing the fabula of the narrative;
  - Reconstructing the plot of the narrative;
  - Representing the provenance, i.e. the inferential process of a narrator who composes a narrative from primary sources describing the events in the fabula.
  
2. Development of a web tool for the creation of narratives, allowing the following actions:
  - Creating the events that compose the narratives in a semi-automated way;

---

<sup>1</sup>[http://iso.org/iso/home/store/catalogue\\_ics/catalogue\\_detail\\_ics.htm?csnumber=52075](http://iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=52075)

- 
- Defining the factual aspects that characterize the events (such as time, persons, locations, physical objects, and conceptual objects);
  - Identifying the role that a specific person performed in an event;
  - Linking a digital object with an event (e.g. objects from digital libraries);
  - Defining the type of each event, choosing from a basic list of pre-defined options;
  - Linking the events to each other using appropriate semantic relations;
  - Storing the textual fragment representing an event in the narrative;
  - Relating the textual fragment to its narrator;
  - Storing the primary sources from which the events are defined (provenance information);
  - Saving the narrative as a digital object in JSON format.
3. Development of a visualization component with the following functionalities:
- Extracting the events along with their primary sources;
  - Extracting the events with a specific primary source;
  - Extracting the events featuring a particular related entity (location, person etc.);
  - Extracting the events that happened in a defined range of time;
  - Extracting the events that are linked with a causal or mereological relation.

During each phase of the project, we developed a prototype to illustrate the user requirements that we had defined. User feedback was then obtained on the prototype to validate and refine the requirements.

### 4.3 Implementation

After selecting a case study and identifying the requirements, we started the implementation phase. This phase depended on the prior creation of an ontology for representing narratives, which is described in the following chapter. The tool would allow us to apply the ontology in order to validate it. The implementation phase can be divided in the following steps:

1. Selection of a knowledge base from which to import data;
2. Design of the user interface of the tool;
3. Selection of a programming language and libraries;
4. Development of the tool for the creation of narratives;
5. Representation of the case study by using the tool.

The first step, selection of a knowledge base, is described in detail in Chapter 6. The other steps are described in Chapter 7.

### 4.4 Visualizations

Finally, the last phase is the development of several visualization components to show a graphical representations of the output of the tool. This phase is ongoing and it is not the main focus of this thesis, however we have decided to present a few of these components, in order to provide a better understanding of the output of the tool. These are shown in Chapter 8.

# Chapter 5

## The Ontology of Narratives

In this Chapter we present the Ontology of Narratives that serves as a basis for the tool and motivated its development. An initial mathematical specification and conceptualization of the ontology can be found in (Bartalesi, Meghini, and Metilli, 2016)<sup>1</sup>. The provisional OWL schema for the ontology is reported in Appendix A.

### 5.1 Components of a Narrative

Following the narratology theory that we briefly presented in Section 2.1, we identified two main components of a narrative: the *fabula* and one or more *narrations* of the fabula.

The fabula is composed of events, which can be related to each other through:

- a *temporal occurrence* relation which associates each event with a time interval during which the event occurs; an event occurs before (or during, or after) another event if and only if the period of occurrence of the former event is before (or during, or after) the period of occurrence of the latter event;
- a *mereological* relation which connects events to other events that include them as parts, e.g., the event *Marriage of Dante Alighieri* is part of the event *Life of Dante Alighieri*;

---

<sup>1</sup>A more comprehensive presentation of the ontology, including its expression in the CIDOC CRM vocabulary, will be published in a forthcoming paper. The ontology will also be described in detail in the soon-to-be-published doctoral thesis of Valentina Bartalesi Lenzi.

- a *causality* relation which links events that have a *cause-effect* relation. This relation may be indirect or far away in time, e.g. the event *Industrial Revolution* caused the event *Global Warming*. The existence of a causality link is established by the narrator.

In addition to the three relations listed above, an event is contextualized in terms of other properties, which describe *Who* (persons) and *What* (things) were involved in it, and *Where* and *When* the event happened (Shaw, Troncy, and Hardman, 2009; Van Hage et al., 2011).

Each narration of a fabula consists of one or more narrators and a text, which is authored by the narrator(s) and constitutes the narration proper. The text may be structured in parts, connected to each other through *textual mereology* relations; each part is also connected by an *event reference* relation to the event of the fabula that it narrates. This relation can be used to obtain the *plot* of the narrative, i.e. the sequence of events as chosen by the narrator.

## 5.2 Expression in the CIDOC CRM

We decided to express our ontology by re-using classes and properties from an existing event-based vocabulary. The chosen vocabulary is part of a well-known standard ontology, the CIDOC CRM<sup>2</sup> (CRM for short). The CRM is a high-level ontology and an ISO standard<sup>3</sup> that allows the representation of information relating to the cultural heritage domain and its correlation with knowledge stored in libraries and archives (Doerr, 2003). The CIDOC CRM ontology is available both in RDF<sup>4</sup> and OWL<sup>5</sup> format.

We chose to represent events from the fabula as instances of the CRM class *E5 Event*, which “comprises changes of states in cultural, social or physical systems, regardless of scale, brought about by a series

---

<sup>2</sup><http://cidoc-crm.org>

<sup>3</sup>[http://iso.org/iso/catalogue\\_detail?csnumber=34424](http://iso.org/iso/catalogue_detail?csnumber=34424)

<sup>4</sup>[http://cidoc-crm.org/official\\_release\\_cidoc.html](http://cidoc-crm.org/official_release_cidoc.html)

<sup>5</sup><http://erlangen-crm.org>

or group of coherent physical, cultural, technological or legal phenomena”<sup>6</sup>. It is possible to also express actions as instances of class *E7 Activity*, which comprises “actions intentionally carried out by instances of *E39 Actor* that result in changes of state in the cultural, social, or physical systems documented”. *E7 Activity* is a subclass of *E5 Event*.

Time intervals are represented in the CRM as instances of class *E52 Time-Span*, which “comprises abstract temporal extents, in the sense of Galilean physics, having a beginning, an end and a duration”.

The three relations we defined on the events of the *fabula* are expressed through the following CRM properties:

- the *event occurrence* relation is represented through the CRM property *P4 has time-span*, which “describes the temporal confinement of an instance of *E2 Temporal Entity*”, and therefore of an event. Because the period of occurrence of an event may not be known, the CRM allows to directly relate events based on their time of occurrence. To this end, it introduces seven properties (*P114* to *P120*) mirroring the temporal relations formalized in Allen’s temporal logic (Allen, 1984). For instance, the CRM property *P120 occurs before* “implies that a temporal gap exists between the end of A and the start of B”, and mirrors the *before* relationship of Allen’s temporal logic.
- the *mereological* relation is represented through the property *P9 consists of*, which “associates an instance of *E4 Period* with another instance of *E4 Period* that is defined by a subset of the phenomena that define the former”. Since *E5 Event* is a subclass of *E4 Period*, *P9* can also be applied to events.
- the *causality* relation is represented by introducing a new property of *causal dependency*. This is due to the fact that the only causal property from the CRM, *P17 was motivated by*, applies exclusively to instances of *E7 Activity*, and we would like to represent causality among general events. An extension of CRM, CRMsci<sup>7</sup>, defines a direct causality relation via the property *O13 triggers*,

<sup>6</sup>All quotations in this Chapter are from the CRM specification version 6.2.2, available at the following URL: [http://www.cidoc-crm.org/docs/cidoc\\_crm\\_version\\_6.2.2%20\(WorkingDoc\).pdf](http://www.cidoc-crm.org/docs/cidoc_crm_version_6.2.2%20(WorkingDoc).pdf)

<sup>7</sup><http://ics.forth.gr/isl/CRMext/CRMsci/docs/CRMsci1.2.3.pdf>

which “associates an instance of *E5 Event* that triggers another instance of *E5* with the latter; an event can activate (trigger) other event/s; in that sense it is interpreted as the cause”. This property can be useful in the context of a scientific experiment, but it is inadequate to the needs of narratives, where causality may not be immediately effective.

The narrator is represented by an instance of the CRM class *E21 Person*. The authoring relation between the narrator and the text of the narration is represented through a creation event, defined as an instance of the CRM class *E65 Creation*. The CRM property *P14 carried out by* connects the creation event to its narrator. The CRM property *P94 has created* connects the creation event to the created narration. The narration itself is an instance of the CRM class *E73 Information Object*, which identifies “immaterial items, such as poems, data sets, images, texts, multimedia objects [...] that have an objectively recognizable structure and are documented as single units”.

The mereology of the text is represented through the CRM property *P106 is composed of*, which connects a structural whole to its parts. However, *P106* represents the structure of the text as defined by its author, and the units of this structure do not necessarily coincide with the units that narrate a single event. A single structural unit may narrate more than one event. For instance, the sentence “Dante took refuge first in the city of Lucca and then in the Casentino region“ narrates at least two events: (i) Dante’s residence in Lucca, and (ii) Dante’s residence in the Casentino region. Each of these events is described by a fragment of the full sentence. It is also possible that a single event is narrated by two or more structural units, which may even be far apart in the actual text of the narration. For instance, the narrator might state “We are fairly certain that Dante also lived in Pisa” in an initial chapter of her work, and then reprise the description of the event in a more specific way in a later chapter (e.g. “Dante moved to Pisa where Emperor Henry VII was temporarily staying”).

In order to identify the chunks of text that narrate a single event we use a specific class, *F23 Expression Fragment*, from the FRBRoo<sup>8</sup> ontology. FRBRoo (Doerr et al., 2008) is a bibliographic ontology

<sup>8</sup><http://new.cidoc-crm.org/frbroo/>



resulting from the harmonization of the FRBR<sup>9</sup> ontology and the CRM. Class *F23* comprises “parts of Expressions and these parts are not Self-Contained Expressions themselves” and is a subclass of the CRM class *E73 Information Object*. Expression fragments are connected to the structural units of text they belong to via the *P106 is composed of* property.

Finally, in order to express the relation between a fragment of text and the event narrated in the fragment, we assume that a fragment is related to a number of propositions, which collectively formalize the content of the fragment. Such relation is expressed through the CRM property *P129 is about*. Each proposition is an instance of the CRM class *E73 Information Object* and is structured according to the Resource Description Framework (RDF) (Manola and Miller, 2004) as consisting of a subject, a predicate and an object. The relations between a proposition and its constituents are expressed by borrowing three properties from the RDF vocabulary: *rdf:subject*, *rdf:predicate* and *rdf:object*. Overall, the relation between a fragment of text and the narrated event is obtained as follows: an expression fragment *is about* a proposition having the event as a *subject* (see Figure 5.3).

Since one of our requirements was to express the primary sources supporting the plot of the narrative, we had to enrich our ontology in order to document the process through which a certain narrative came into existence. To this end, we modeled the construction of a narrative by an historian as an inferential process, using evidence collected from sources to infer propositions that were subsequently narrated in a text. The primitives for modelling inferences are borrowed from an extension of the CRM, the CRMinf, an ontology for capturing argumentation and inference making in descriptive and empirical sciences. The CRMinf is still a proposal for approval by the Special Interest Group of the CIDOC CRM<sup>10</sup>. Generally speaking, the adoption of the CRMinf allowed us to describe the knowledge *provenance* as the process of tracing the origins of knowledge (Committee et al., 2008).

---

<sup>9</sup>The FRBR ontology was developed by the International Federation of Library Associations and Institutions (IFLA), and it is described here: <http://www.ifla.org/publications/functional-requirements-for-bibliographic-records>

<sup>10</sup><http://ics.forth.gr/isl/CRMext/CRMinf/docs/CRMinf-0.7.pdf>

### 5.3 The Semantic Network

The semantic network that we constructed to model Dante’s biography and its provenance includes the *fabula* and the *narration*, as defined above; in addition, it also includes the *provenance*, which represents the inferential process of the narrator who composes a narrative from primary sources whose contents are propositions about the events in the *fabula*. In the following examples, CRM classes and properties are named using their CRM names, without prefix, while terms re-used from other vocabularies are named using prefixed qnames, e.g. *rdf:type*. Class instances are named in lower-case strings ended by a number in Sans Serif type, e.g. e1.

**Fabula.** Figure 5.1 shows a graphical view of a small part of the *fabula*, including two events: e1, the death of Dante, and e2, the exile of Dante from the city of Florence. Both of these events are instances of *E5 Event*, in addition e1 is also an instance of *E63 Beginning of Existence*, which is a subclass of *E5 Event*. Note that we use the *rdf:type* property from the RDF vocabulary to model the instance-of relation. Dante himself is represented by an instance of *E21 Person*. The property *P120 occurs before* links the event e2 to the event e1, meaning that the exile of Dante from Florence happened before his death.

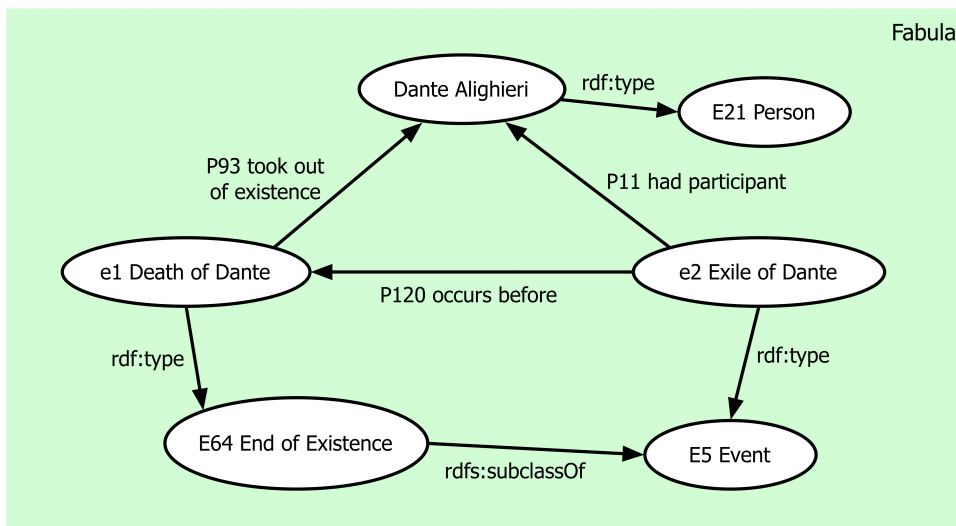


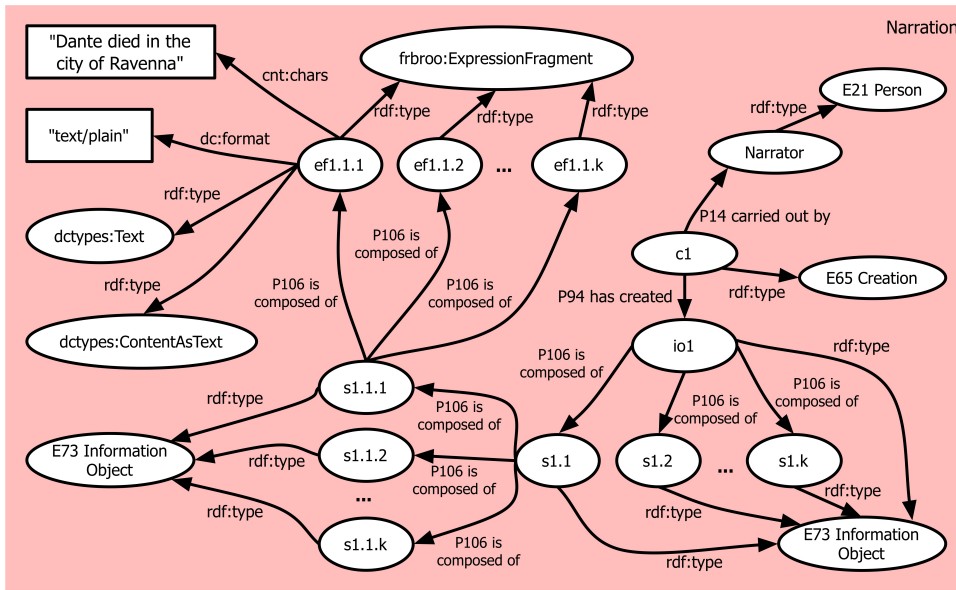
FIGURE 5.1: A portion of the *fabula*.

In our ontology, an event is contextualized through the properties that describe *Who* (persons) and *What* (things) were involved in it, and *Where* and *When* the event happened (Shaw, Troncy, and Hardman, 2009; Van Hage et al., 2011). In particular, in the specific example shown in Figure 5.1:

- *Who* is represented by property *P11 had participant*, which links an event with one or more instances of *E39 Actor*, representing the persons or groups of people who participated in the event<sup>11</sup>. If the event is an action, property *P14 carried out by* is used as predicate in the proposition to link the action with the corresponding actor. In this case, the property *P14.1 in the role of* is added to the proposition in order to specify the nature of the actor’s participation. Since *P14.1* stands for a ternary relation, reification is required for its representation in RDF. In order to also give a role to persons who were present at the event but did not actively participate in it, we decided to extend this property in order to apply it to *P11*.
- *What* is represented by property *P12 occurred in the presence of*, which relates an instance of *E5 Event* with an instance of *E77 Persistent Item*. This property allows us to link events with a variety of objects, including concepts (e.g. “General Theory of Relativity” in the sentence “Albert Einstein developed the General Theory of Relativity”) as instances of *E89 Propositional Object*.
- *When* is represented by property *P4 has time-span*, which links an event with the instance of *E52 Time-Span* that identifies the period of occurrence of the event.
- *Where* is represented by property *P7 took place at*, which links an event with the instance of *E53 Place* giving the location of the event, or by property *P8 took place on or within*, in case the location of an event is described in terms of a physical object (such as the crowning of Emperor Henry VII taking place in the Basilica of St. John Lateran in Rome).

The event may also be related with: (i) a textual annotation (e.g. a comment in natural language that the narrator can write in order

<sup>11</sup> *E21 Person* and *E74 Group* are both subclasses of *E39 Actor*.

FIGURE 5.2: The *narration*.

to add some additional explanation to the event) through the property *P3 has note*; (ii) a digital object, such as an image or video that can be used as a representation of the event, through the property *P128 represents*<sup>12</sup>.

**Narration.** Figure 5.2 shows a portion of the *narration* of the biography of Dante. At the bottom of the Figure, the biography is represented by object *io1*, an instance of the class *E73 Information Object*, resulting from the event *c1* of its creation carried out by the biographer. At the top of the figure, the biography is structured in chapters, paragraphs and textual fragments. We used the class *E73 Information Object* for representing chapters and paragraphs and *ExpressionFragment* of FR-BRoo (Doerr et al., 2008) for textual fragments.

Each instance of *ExpressionFragment* has a textual content which is modelled according to the recommendations of the W3C's Content in RDF<sup>13</sup>, as shown in the top left corner of Figure 5.2. Specifically:

1. the property *format* of the Dublin Core ontology<sup>14</sup>, giving the MIME media type of the instance. This allows distinguishing,

<sup>12</sup>For instance, the event *Global Warming* could be linked to the digital object found at <http://xkcd.com/1732>, which provides a simple visual description of it.

<sup>13</sup><http://w3.org/TR/Content-in-RDF10>

<sup>14</sup><http://dublincore.org/>

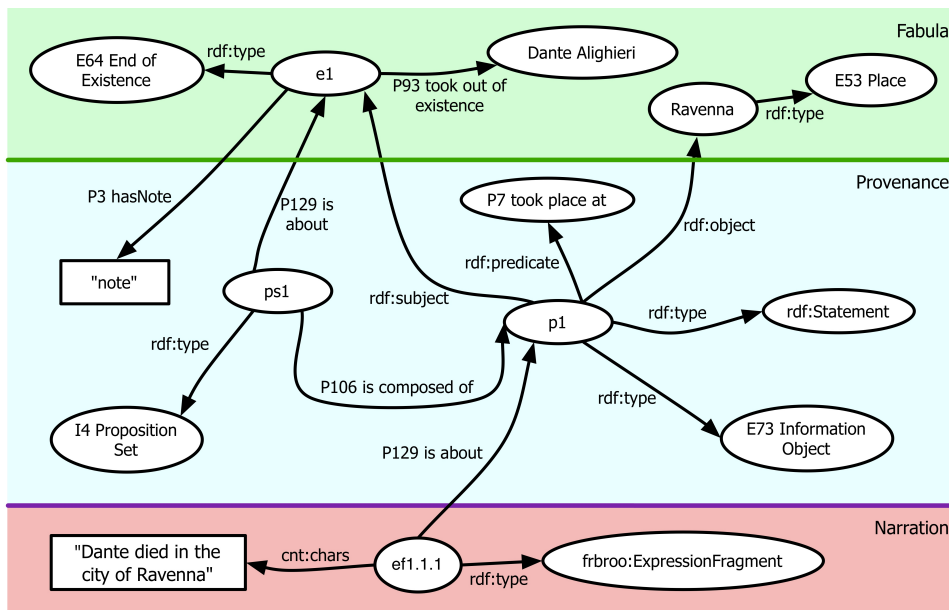


FIGURE 5.3: The structure of proposition p1.

e.g. embedded content in plain text from content encoded in HTML;

2. also from Dublin Core, the *dctypes:Text* class, to indicate that the instance represents a resource primarily intended to be read;
3. the property *chars* in the RDF Content namespace<sup>15</sup>, giving the sequence of the characters of the content;
4. also from the RDF Content namespace, the *cnt:ContentAsText* class, to indicate that the instance represents textual content.

**Provenance.** Figure 5.3 shows a portion of the provenance, centered around proposition p1. This proposition expresses that its subject, the death of Dante (event e1), stands according to its predicate *P7 took place at* with its object, the city of Ravenna. Therefore, p1 says that Dante died in Ravenna. Note that p1 is part of ps1, an instance of *I4 Proposition Set* from the CRMinf ontology.

Figure 5.4 shows another portion of the *provenance*, representing the inferential process of the biographer. Everything starts from the reading of a source, modelled as an event of observation, o1, that is an instance of class *S4 Observation*. The text of the source oe1, and

<sup>15</sup><https://w3.org/TR/Content-in-RDF10/>

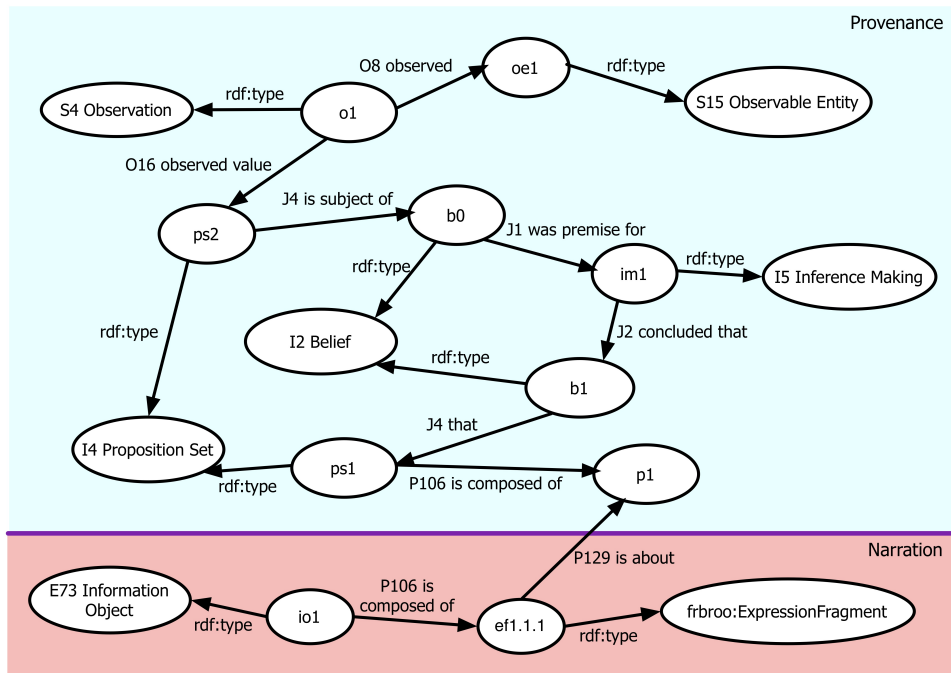


FIGURE 5.4: Representation of the *provenance*.

instance of class *Observable Entity*, has been *O8 observed* by *o1*, resulting in the *O16 Observed Value* *ps2*, a set of propositions. After performing the observation, the narrator holds belief *b1* *J4 that* proposition set *ps2* is true. The narrator performs an inference making *im1* in which she *J1 used as premise* the belief *b1*, and *J2 concluded that* the belief *b2* is true. *b2* holds a particular set of propositions *ps1* to be true, and among these is the proposition *p1* asserting that Dante died in Ravenna. Therefore, the connection between the reading of a specific primary source and the proposition that Dante died in Ravenna is established through an inference making event, between the proposition set resulting from the reading of the source and the proposition set which includes the conclusion.

This concludes our description of the Ontology of Narratives for the purposes of this thesis.

## Chapter 6

# A Collaborative Knowledge Base: Wikidata

To simplify the insertion of data by the user and subsequent population of the ontology, we decided to build our tool on top of an existing knowledge base, in order to provide the user with a vast and detailed amount of resources with which to build the narrative. For this purpose we initially considered two knowledge bases: DBpedia<sup>1</sup> and Wikidata<sup>2,3</sup>.

DBpedia (Auer et al., 2007) is a free knowledge base developed by the Free University of Berlin, the University of Leipzig and OpenLink Software through automatic extraction of structured knowledge from multiple language versions of the Wikipedia collaborative encyclopedia. It currently contains 4.5 million entities<sup>4</sup> and allows exporting of the data through the DBpedia Lookup Service and a SPARQL endpoint. DBpedia is especially popular as a research tool (Lehmann et al., 2015).

Wikidata (Vrandečić and Krötzsch, 2014) is a free collaborative knowledge base developed and hosted by the Wikimedia Foundation, built by extracting structured knowledge from several open collaborative projects including Wikipedia, Wikisource, Wikispecies, Wikibooks, Wikinews, Wikiquote, and more. It currently contains more than 20.1

---

<sup>1</sup><http://dbpedia.org>

<sup>2</sup><http://wikidata.org>

<sup>3</sup>Another knowledge base that we considered is Freebase, owned by Google, but we decided to exclude it when we learned that it had been discontinued in favor of Wikidata (Pellissier Tanon et al., 2016).

<sup>4</sup><http://wiki.dbpedia.org/about/facts-figures>

million entities<sup>5</sup> and allows exporting of the data through the Wikidata API and a SPARQL endpoint. Wikidata encourages collaborative addition and editing of the data through manual and automated means.

We identified the following differences between the two knowledge bases:

- Wikidata contains more than four times as many entities as DBpedia, and it features automatic importing of entities not just from Wikipedia but from many more projects managed by the Wikimedia Foundation. It also accepts imports of knowledge from non-Wikimedia external sources, provided their contents are in the public domain, and it has recently been chosen as the successor to the discontinued Freebase project. DBpedia is more limited in scope, focusing on importing knowledge from Wikipedia.
- DBpedia has many international versions but these are often disconnected from each other (Lehmann et al., 2015), while Wikidata acts as a central hub for the almost 300 language versions of Wikipedia. Wikidata is also more fully multilingual than DBpedia, with 39% of the entities having labels in multiple languages.
- Wikidata aims to collect not just statements about entities, but also the primary sources behind those statements. Referenced statements are currently more than 62% of the total<sup>6</sup>.
- DBpedia has a stable ontology that is carefully managed. Wikidata is more open to collaborative user input, allowing the free addition, editing and correction of knowledge from users and also allowing them to refine and improve the class hierarchy. This openness increases the number of participants, and thus the quantity of the imported knowledge, but it also allows users to modify the ontology in unpredictable ways, as shown in Section 6.1.
- Wikidata is fully integrated into Wikipedia and other Wikimedia projects, allowing instant automatic updating of the knowledge base following each change in the connected projects. DBpedia

---

<sup>5</sup><https://wikidata.org/wiki/Wikidata:Statistics>

<sup>6</sup><https://tools.wmflabs.org/wikidata-todo/stats.php>



also imports data from Wikipedia in real time through the DBpedia Live system, but is not integrated into Wikipedia itself (Lehmann et al., 2015).

- DBpedia adopts a Creative Commons Attribution-ShareAlike<sup>7</sup>. Wikidata adopts a Creative Commons Zero<sup>8</sup> license, which is equivalent to the public domain, thus making it easier to re-use the knowledge contained in it.

Some of the differences we listed are among those that were identified by (Färber et al., 2015) in their recent review of five knowledge bases (DBpedia, Wikidata, Freebase, OpenCyc<sup>9</sup>, and YAGO<sup>10</sup>), where Wikidata reported the highest scores in most of the metrics considered by the authors. Based on this review and on the comparison presented above, it would be easy to dismiss DBpedia and embrace Wikidata for the purposes of our project.

However, Wikidata also presents some potential downsides. Chief among these is its collaborative nature, which can sometimes be a double-edged sword. Just like its sister project Wikipedia, Wikidata’s openness makes it prone to vandalism from malicious users. Furthermore, it makes the structure of its ontology less stable than that of DBpedia. This issue has recently been investigated by Spitz et al. (2016). The authors lamented the “complicated hierarchy, in which classes appear at somewhat arbitrary levels”, and the “perpetually changing content that is subject to constant updates”. However, they also recognized the appeal of Wikidata for research and identified most of the positive aspects that we reported above.

Another potential downside of Wikidata could be the fact that it was not initially conceived as a Semantic Web project. Its internal structure is a document-based database, and compatibility with Semantic Web technologies was added later. In some cases, the structure of a Wikidata statements requires reification to obtain a full conversion to RDF, as discussed by Hernández, Hogan, and Krötzsch (2015). Despite these challenges, a translation of the Wikidata model to RDF has now been completed (Erxleben et al., 2014).

<sup>7</sup><https://creativecommons.org/licenses/by-sa/4.0/>

<sup>8</sup><https://creativecommons.org/publicdomain/zero/1.0/>

<sup>9</sup><http://sw.opencyc.org>

<sup>10</sup><http://yago-knowledge.org>

At the end of our analysis of Wikidata and DBpedia, while being wary of the issues listed above, we decided that the advantages of Wikidata outnumbered its shortcomings. Given that Wikidata had not yet been applied in the Computational Narratology field, a secondary purpose of our research thus became the exploration of those shortcomings, and how they could affect a project such as ours.

As explained in Chapter 7, our tool leverages the knowledge stored in Wikidata for the following purposes:

- providing the user with a list of entities to use in the narrative;
- importing basic events, such as the birth and death of a person;
- referencing primary sources and their authors.

In order to extract RDF knowledge to import in our tool, we made use of the Wikidata Query Service (WQS)<sup>11</sup>, a SPARQL endpoint that provides full access to the knowledge stored in Wikidata and is constantly updated following each change in the knowledge base.

## 6.1 Wikidata–CRM Mapping

In order to re-use Wikidata entities, we developed a mapping that links Wikidata classes to the CIDOC CRM classes that we adopted in our ontology. The mapping is reported in Table 6.1<sup>12</sup>.

TABLE 6.1: Mapping between Wikidata and the CRM.

Wikidata Class	CRM Class
<i>Q5 Human</i>	<i>E21 Person</i>
<i>Q16334295 Group of Humans</i>	<i>E74 Group</i>
<i>Q223557 Physical Object</i>	<i>E19 Physical Object</i>
<i>Q7184903 Abstract Object</i>	<i>E89 Propositional Object</i>
<i>Q17334923 Location</i>	<i>E53 Place</i>
<i>Q234460 Text</i>	<i>E73 Information Object</i>
<i>Q107715 Time</i>	<i>E52 Time-Span</i>
<i>Q1190554 Event</i>	<i>E5 Event</i>

<sup>11</sup><https://query.wikidata.org>

<sup>12</sup>Wikidata classes are generally written lowercased, but for clarity we will use the title case throughout this section.

In theory, this mapping appears sufficient to make the two ontologies compatible. In practice, however, we had to face some significant issues:

- The first issue we noticed when applying the mapping is the fact that most of the entities in Wikidata are instances of more than one of our basic classes. For instance, a *Q5 Person* is also a *Q223557 Physical Object*, a building can be both a *Q223557 Physical Object* and a *Q17334923 Location*, and until recently an *Q5061 Administrative Territorial Entity* was both a place and an organization. So it would not suffice to develop the mapping, we also had to create a classification algorithm to sort out the entities and apply the proper class to each of them.
- The second issue we faced is that sometimes Wikidata and the CRM are misaligned. For instance, the CRM class *E21 Person* and the CRM class *E74 Group*, representing groups of people, are both subclasses of *E39 Actor*. In Wikidata, the classes *Q16334295 Group of Humans* and *Q5 Human* are solely connected through the part-whole property *P527 part of*, therefore if we want to group the two classes together we have to select them separately.
- A third issue we encountered is that sometimes a Wikidata user inadvertently makes an edit that changes the structure of the ontology in a major way. This happened to us three times during the development of the tool: (i) when a user changed the definition of a subclass of *Q386724 Work*, (ii) when another user changed the definition of *Q515 City*, and (iii) when two users in sequence changed the definition of *Q5 Human*<sup>13</sup>. The first case affected only a few literary works, with relatively minor consequences. The second one indirectly made every city a subclass of event, which was a major mistake. The third one was the most significant, as it affected all Wikidata entities representing humans. In the first case we resolved the issue by simply pointing

---

<sup>13</sup>A user of Wikidata incorrectly marked *Q5 Human* as a subclass of *Q5891007 Anatomically Modern Human* (the opposite is true), and another marked *Q5891007* as a subclass of *Q83944 Human Evolution*. These two mistakes effectively made *Q5 Human* a subclass of *Q1190554 Event*.

it out to Wikidata users, in the second and third cases we decided to intervene immediately and correct the errors ourselves. As we said before, the open and collaborative nature of Wikidata can provide both advantages and disadvantages. However, we think that it would be wise to protect important classes such as *Q5 Human* from indiscriminate editing.

- A fourth issue is the fact that a small number Wikidata classes have been wrongly applied by users on a mass scale. The prime example of this is the class *Q571 Book*, which is often carelessly applied to entities which should instead be marked as instances of *Q7725634 Literary Work*. The problem is that *Q571 Book* is currently a subclass of both *Q386724 Work* and *Q223557 Physical Object*, making it difficult to understand whether the Wikidata entity represents the work in general or a specific manifestation of it. This issue has been the focus of ongoing discussions by the users<sup>14</sup>, however a solution has not yet been agreed upon.

Despite these issues, we were able to develop a mapping algorithm for use in our tool. For reference, in Figure 6.1 we show the process that we currently follow to classify the entities. In the interest of simplicity, we have omitted the Wikidata class IDs from the image.

Currently, the process we apply in order to classify Wikidata entities is as follows:

1. If the entity is an instance of class *Q5 Human*, it is assigned the CRM class *E21 Person*.
2. If the entity is an instance of class *Q234460 Text*, it is assigned the CRM class *E73 Information Object*.
3. If the entity is an instance of class *Q8205328 Artifact* or class *Q41176 Building*, it is classified as an instance of the CRM class *E19 Physical Object*. This serves to avoid classifying buildings as locations, and also to excludes physical artifacts from the further classification step for works.

---

<sup>14</sup>An example of a recent discussion about this subject can be found at the following URL: [https://wikidata.org/wiki/Wikidata\\_talk:WikiProject\\_Books#Wikiproject\\_Books\\_2.0](https://wikidata.org/wiki/Wikidata_talk:WikiProject_Books#Wikiproject_Books_2.0)

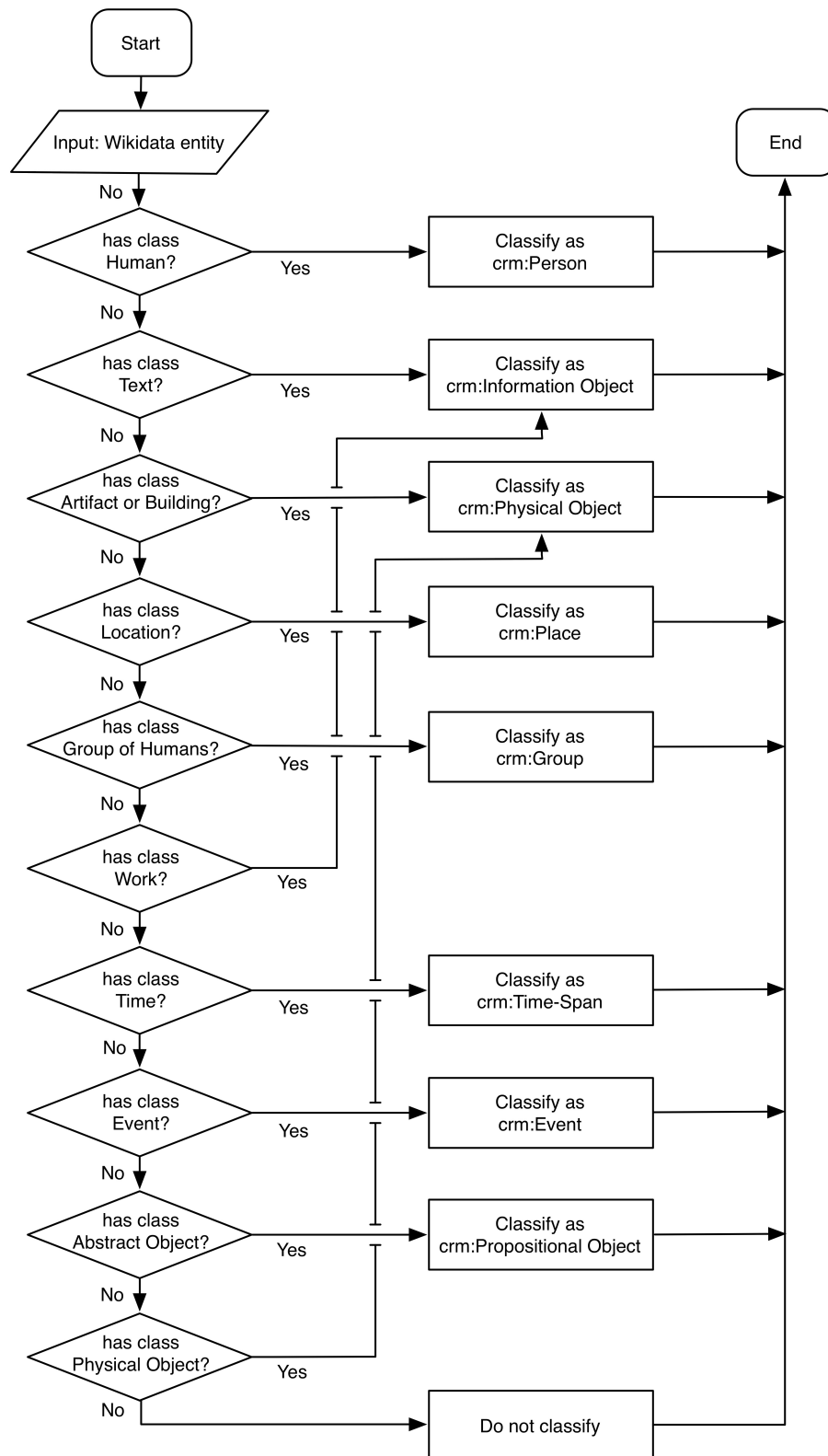


FIGURE 6.1: Flowchart showing the mapping process between Wikidata and the CRM.

4. If the entity is an instance of class *Q17334923 Location*, it is assigned the CRM class *E53 Place* (buildings have been excluded in the previous step).
5. If the entity is an instance of class *Q16334295 Group of Humans*, it is assigned the CRM class *E74 Group*. Given that locations have been classified in step 6, an entity with a dual organizational and territorial nature (e.g. a city) will be classified as a location, not a group of people.
6. If the entity is an instance of class *Q386724 Work*, it is classified as an instance of the CRM class *E73 Information Object*. Buildings and artifacts, which are physical in nature, have been already classified as *Q223557 Physical Object* in step 5.
7. If the entity is an instance of class *Q186081 Time*, it is classified as an instance of the CRM class *E52 Time-Span*.
8. If the entity is an instance of class *Q15474042 Event*, it is classified as an instance of the CRM class *E5 Event*.
9. If the entity is an instance of class *Q7184903 Abstract Object*, it is assigned the CRM class *E89 Propositional Object*. Works and texts are excluded because they have been classified as instances of *E73 Information Object* in steps 4 and 8.
10. Finally, all remaining instances of *Q223557 Physical Object* are assigned the CRM class *E19 Physical Object*.

It should be noted that the steps we described must be executed in this specific order, otherwise items with multiple classes would be wrongly categorized. For instance, if steps 4 and 5 were inverted, cities could be categorized as *E74 Group*<sup>15</sup>. This would not be acceptable in the CRM, where cities should be instances of *E53 Place*. Furthermore, ensuring that the classification step of *E73 Information Object* is made before that of *Q223557 Physical Object* solves the aforementioned problem with the ambiguous Wikidata class *Q571 Book*.

<sup>15</sup>As discussed above, this dual nature of cities in Wikidata has been very recently changed and now they are no longer considered groups of people. Still, we decided not to remove this example because it explains very well the general issue.

## 6.2 Event Types

A list of event types was identified in order to classify the events that compose the narratives. This is useful to: (i) define recurring patterns in similar narratives, (ii) reuse these recurrent patterns in different narratives, simply by changing the instances of the involved classes, (iii) automatically extract the same types of events from different narratives on the same topic. In order to maximize the interoperability of the system, we mapped the event types defined in the CIDOC CRM standard ontology to the equivalent Wikidata classes.

The following list reports the event types extracted from the CRM and, when available, the equivalent Wikidata class:

- *E63 Beginning of Existence*, which may be used “for temporal reasoning about things (intellectual products, physical items, groups of people, living beings) beginning to exist; it serves as a hook to determine of a terminus post quem and ante quem”. This is a subclass of *E5 Event*, and corresponds to the Wikidata class *Q23956340 Beginning of Existence*. We also adopted the following subclasses of *E63*:
  - *E12 Production*, which “comprises activities that are designed to, and succeed in, creating one or more new items”, and corresponds to the Wikidata class *Q739302 Production*;
  - *E65 Creation*: the class “comprises events that result in the creation of conceptual items or immaterial products, such as legends, poems, texts, music, images, movies, laws, types etc”, and corresponds to the Wikidata class *Q19068184 Creation*;
  - *E66 Formation*, which “comprises events that result in the formation of a formal or informal E74 Group of people, such as a club, society, association, corporation or nation”; not found in Wikidata;
  - *E67 Birth*: this class “comprises the births of human beings. E67 Birth is a biological event focussing on the context of people coming into life”, and corresponds to the Wikidata class *Q34581 Birth*;

- *E81 Transformation*, which “comprises the events that result in the simultaneous destruction of one or more than one E77 Persistent Item and the creation of one or more than one E77 Persistent Item that preserves recognizable substance from the first one(s) but has fundamentally different nature and identity”; equivalent to Wikidata class *Q1150070 Change*;
- *E64 End of Existence*, which “may be used for temporal reasoning about things (physical items, groups of people, living beings) ceasing to exist; it serves as a hook for determination of a terminus postquem and antequem”. This is a subclass of *E5 Event*, and corresponds to the Wikidata class *Q23956356 End of Existence*. We also adopted the following subclasses of *E64*:
  - *E6 Destruction*: this class “comprises events that destroy one or more instances of E18 Physical Thing such that they lose their identity as the subjects of documentation”, and corresponds to the Wikidata class *Q19868482 Destruction*;
  - *E68 Dissolution*: this class “comprises the events that result in the formal or informal termination of an E74 Group of people”, and corresponds to the Wikidata class *Q18603731 Dissolution*;
  - *E69 Death*: this class “comprises the deaths of human beings”, and corresponds to the Wikidata class *Q4 Death*;
  - *E81 Transformation*, defined as in the previous list of subclasses of *E63*; equivalent to the Wikidata class *Q1150070 Change*;
- *E85 Joining*, this class “comprises the activities that result in an instance of E39 Actor becoming a member of an instance of *E74 Group*. This class does not imply initiative by either party. It may be the initiative of a third party. Typical scenarios include becoming a member of a social organisation, becoming employee of a company, marriage, the adoption of a child by a family and the inauguration of somebody into an official position”. *E85* is a subclass of *E7 Activity*, and it is equivalent to Wikidata class *Q23009459 Connect*.



- *E86 Leaving*, this class “comprises the activities that result in an instance of *E39 Actor* to be disassociated from an instance of *E74 Group*. This class does not imply initiative by either party. It may be the initiative of a third party. Typical scenarios include the termination of membership in a social organisation, ending the employment at a company, divorce, and the end of tenure of somebody in an official position”. *E86* is a subclass of *E7 Activity*, and it is not found in Wikidata.

The following Wikidata classes were also added to the list of event types in order to represent events from our case study:

- *Q35856 Baptism*, “Christian rite of admission and adoption, almost invariably with the use of water”;
- *Q178561 Battle*, “part of a war which is well defined in duration, area and force commitment”;
- *Q209715 Crowning*, “ceremony marking the formal investiture of a monarch and/or his or her consort with regal power”;
- *Q476300 Competition*, “contest between organisms, animals, individuals, groups, etc.”;
- *Q93190 Divorce*, “termination of a marital union”;
- *Q8434 Education*, “learning in which knowledge and skills is transferred through teaching”;
- *Q40231 Election*, “process by which a population chooses an individual to hold public office”;
- *Q188863 Exile*, “event by which a person is forced away from home”;
- *Q1725430 Lawmaking*, “process of crafting legislation”;
- *Q8445 Marriage*, “social union or legal contract between people called spouses that creates kinship”;
- *Q2761147 Meeting*, “event in which two or more people assemble”;
- *Q124734 Rebellion*, “refusal of obedience or order”;

- *Q2359691 Residence*, “presence in a specific jurisdiction”;
- *Q1763090 Sentence*, “decree of punishment in law”;
- *Q61509 Travel*, “movement of people between relatively distant geographical locations”;
- *Q198 War*, “organised and prolonged conflict between states”.

We added the event types that we found most suitable for representing a biography. An enrichment of these types has been planned as future work, possibly taking their definitions from the extensive Wikidata ontology<sup>16</sup>. The final aim is to assemble a list of event types that is sufficiently large to allow the extension of the proposed system to other subjects and other types of narratives.

### 6.3 Roles

Another important list that we had to define before the implementation of our tool is that of roles. A role is defined in the CRM as “the specific nature of an actor’s participation in an event”, and is attached to the property *P14 carried out by* through the meta-property *P14.1 with role*.

However, we found the current definition insufficient for our needs, since *P14.1* can be applied exclusively to those actors who actively carried out the event, and excludes all those who participated in the event in a passive way. For instance, in the event of a murder, we would like to identify the roles of both the killer, the victim, and any witnesses.

In order to model active and passive roles, we defined a new property *hasRole* which can be applied to any kind of event and has domain and range *role*. It should be noted that roles are sometimes dependent on the type of event, for instance an event of type *Q188863 Exile* will probably have an exiler and an exiled, a *Q8445 Marriage* will probably have an *officiant*, two participants with role *spouse*, some with role *witness*, and so on. These cases require custom definitions that are associated to each event type. The full list of roles we identified for our

<sup>16</sup>A SPARQL query for all subclasses of events with an English label that are present in Wikidata currently returns more than 56,000 results: <http://tinyurl.com/gv3ccfz>.

event types is reported in Table 6.2. The default role in each event, which is not reported in the table, is *participant*.

TABLE 6.2: Roles based on event types.

Wikidata Class	CRM Class
<i>Baptism</i>	<i>baptized, baptizer</i>
<i>Battle</i>	<i>combatant</i>
<i>Birth</i>	<i>mother, father, child</i>
<i>Creation</i>	<i>creator</i>
<i>Crowning</i>	<i>crowner, crowned</i>
<i>Dissolution</i>	<i>dissolver</i>
<i>Divorce</i>	<i>divorcee</i>
<i>Education</i>	<i>student, teacher</i>
<i>Election</i>	<i>elected, elector</i>
<i>Friendship</i>	<i>friend</i>
<i>Joining</i>	<i>member</i>
<i>Lawmaking</i>	<i>lawmaker</i>
<i>Marriage</i>	<i>spouse, officiant, witness</i>
<i>Production</i>	<i>producer</i>
<i>Rebellion</i>	<i>rebel</i>
<i>Residence</i>	<i>resident</i>
<i>Sentence</i>	<i>sentencer, sentenced</i>
<i>Transformation</i>	<i>transformer</i>
<i>Travel</i>	<i>traveller</i>

In other cases, roles can be expressed in a more general way. For instance, a meeting of the Chamber of Deputies will have one participant with role “President of the Chamber of Deputies” and many participants with role “Deputy”. This role is equivalent to the position *Q1055894 Deputy* from Wikidata, which however cannot be directly used for our purposes because it is not intended to be used as a role.

In Wikidata, *Q4164871 Position* is a subclass of *Q12737077 Occupation*, which is a subclass of *Q1914636 Activity*. An activity is an event, not a role. To work around this issue, we have decided to import the names of instances of *Q12737077 Occupation* when needed, but instead of using them directly we redefine them as roles.



## Chapter 7

# A Tool for the Creation of Narratives

After choosing the Wikidata knowledge base as a resource and developing a mapping between Wikidata and the CRM, we implemented our tool for the creation of narratives. In this Chapter we will describe the whole process of narrative creation, the user interface of the tool, and the underlying technologies that we adopted.

The tool has been designed for scholars who have already written a text about the subject of the narrative and want to create a semantic representation of it. However, a textual representation is not required, therefore the tool can also be used by anyone who wishes to represent the story of a particular subject without having previously written about it (for instance, a high-school student who is studying the life of Dante Alighieri, or another historical figure).

The tool is web-based, being written in HTML5<sup>1</sup>, CSS3<sup>2</sup>, and JavaScript (ECMAScript 5.1<sup>3</sup>), using the jQuery<sup>4</sup>, jQuery UI<sup>5</sup>, Bootstrap<sup>6</sup>, and Typeahead.js<sup>7</sup> libraries. It also makes use of an IndexedDB<sup>8</sup> database to store data locally and retrieve it on subsequent loadings.

We strived to maintain simplicity and ease of use at the core of the user interface by hiding from view all the technicalities of the ontology in favor of a clear and streamlined narrative creation process. The

---

<sup>1</sup><http://w3.org/TR/html5/>

<sup>2</sup><http://w3.org/TR/CSS/>

<sup>3</sup><http://ecma-international.org/ecma-262/5.1/>

<sup>4</sup><http://jquery.com>

<sup>5</sup><https://jqueryui.com>

<sup>6</sup><http://getbootstrap.com/>

<sup>7</sup><http://twitter.github.io/typeahead.js/>

<sup>8</sup><http://w3.org/TR/IndexedDB/>

work of the user is also made easier by the import of entities from the Wikidata knowledge base, which automates part of the process.

The main interface of the tool is composed of three views, which correspond to the three phases of the narrative creation:

1. The first view is used to select the subject of the narrative;
2. The second view is dedicated to for the main narrative creation;
3. The third view is used to define causal and mereological relations between the events of the narrative.

## 7.1 Subject Selection

When the user loads the narrative building tool for the first time, she is provided with a view that allows her to select the subject of the narrative. She can either select the subject from a few default examples, or insert the name of the person in a text field.

Figure 7.1 shows the Subject Selection view. The subject can be any entity that is present in Wikipedia. In the Figure we show a few default entities, including notable people from various historical periods, an organization (NASA), and a physical object (the Tower of Pisa). If the user selects one of these default entities, the narrative creation phase begins.

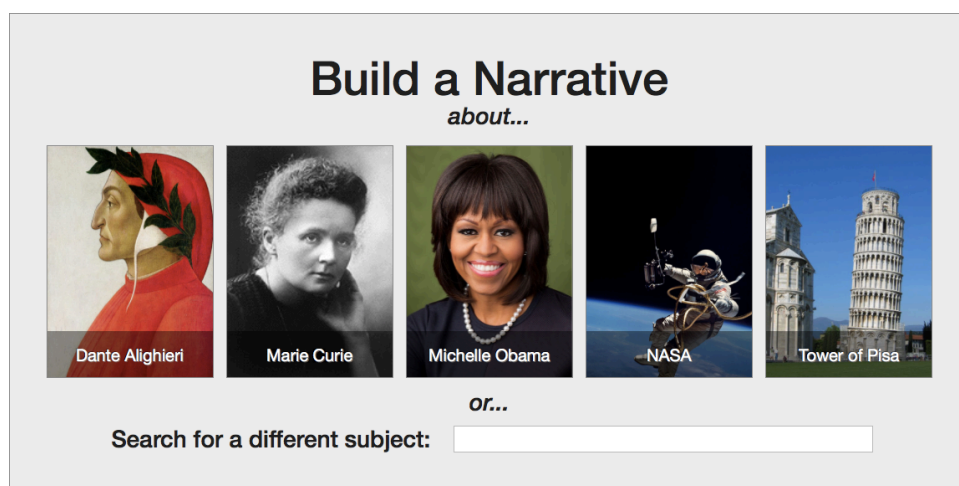


FIGURE 7.1: View of the subject selection interface.

If the user prefers to search for another subject, she can type some characters in order to be provided with a visual list of possible subjects whose names includes the inserted text, created by querying the Wikipedia API and visualized dynamically using the jQuery library. Clicking on the name of a person starts the narrative creation phase.

## 7.2 Narrative Creation

The second view of the web application is dedicated to the narrative creation phase. The screen is divided into three main sides:

- the left-hand side of the screen contains a list of all entities that can be chosen as event components, and a series of buttons to filter the displayed entities, search them by name, or add new ones;
- the right-hand side of the screen contains the main event creation form, and a series of buttons to save the event, clear the form, or switch to the relations screen;
- the bottom side of the screen contains the event timeline, which can be scrolled horizontally.

The narrative creation screen is shown in Figure 7.2.

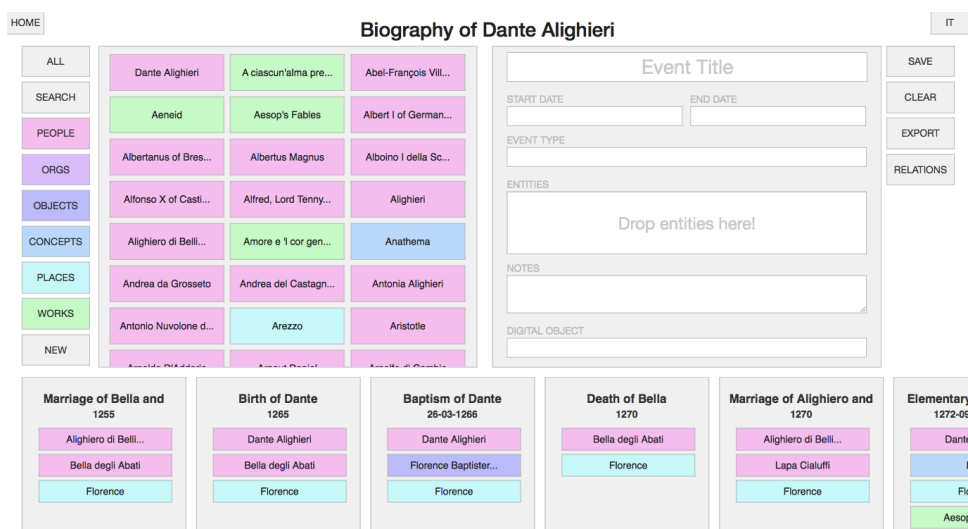


FIGURE 7.2: View of the narrative creation interface.

### 7.2.1 First Load and Entity List

During the first load, the tool performs several queries to the Wikipedia API and the Wikidata Query Service (WQS). First of all, a SPARQL query is performed in order to extract all the relevant information about the subject of the narrative.

Then, the tool queries the Wikipedia API to extract all Wikipedia pages that are linked from the one about the subject of the narrative. For instance, for Dante Alighieri it would extract pages about the members of his family, his works, the places where he lived, the people he interacted with, etc.

Finally, the tool queries the WQS for the names, descriptions, and classes of each of the Wikidata items corresponding to these Wikipedia pages, using the SPARQL query shown in Figure 7.3.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX schema: <http://schema.org/>
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
SELECT DISTINCT ?uri ?type ?name ?desc ?image
WHERE {
  ?uri wdt:P31 ?class.
  ?class wdt:P279* ?type.
  OPTIONAL { ?uri wdt:P18 ?image. }
  OPTIONAL { ?uri rdfs:label ?name FILTER (lang(?name) = "en"). }
  OPTIONAL { ?uri schema:description ?desc FILTER (lang(?desc) = "en"). }
  VALUES ?uri { [list of Wikidata entity IDs] }
  VALUES ?type { [list of Wikidata class IDs] }
}
```

FIGURE 7.3: SPARQL query to extract entities from Wikidata.

Once all the Wikidata queries have finished loading, the tool classifies them using the Wikidata-CRM mapping that we discussed in Section 6.1. The six main classes are reported in Table 7.1.

Then, the tool saves the list of entities to the IndexedDB database and displays them in container on the left-hand side of the screen. Each entity is color-coded according to its class, in order to allow the user to instantly recognize its nature. The user can filter entities by clicking



TABLE 7.1: Mapping between the tool and the CRM.

Tool Class	CRM Class
<i>person</i>	<i>E21 Person</i>
<i>organization</i>	<i>E74 Group</i>
<i>object</i>	<i>E19 Physical Object</i>
<i>concept</i>	<i>E89 Propositional Object</i>
<i>place</i>	<i>E53 Place</i>
<i>work</i>	<i>E73 Information Object</i>

on the button corresponding to each class, and each of these buttons is also color-coded.

Another way to filter entities is by clicking on the Search button. When clicking on the Search button, a search field appears on top of the list, allowing the user to filter the entities by name. A detailed view of the entities container is shown in Figure 7.4.



FIGURE 7.4: View of the entities container.

When clicking on each entity from the list, a *popover*<sup>9</sup> appears, showing the user the name, description, an image of the entity loaded from the Wikimedia Commons<sup>10</sup> repository (when available), and links to the corresponding Wikipedia and Wikidata pages. An example of entity popover is shown in Figure 7.5.

<sup>9</sup>The *popover* is an interface element provided by the Bootstrap library.

<sup>10</sup><https://commons.wikimedia.org>

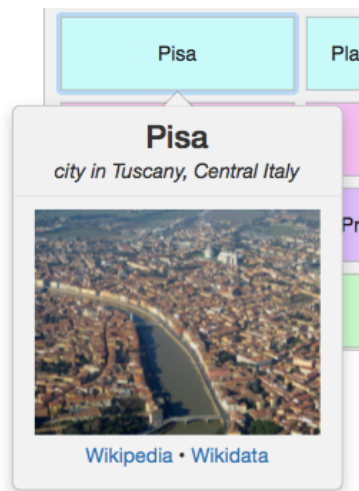


FIGURE 7.5: An example of entity popover for the city of Pisa, Italy.

The user can also create new entities by clicking on the New button found on the left side of the entity container. Upon clicking, the user is shown a popover through which she can either:

- insert a Wikipedia URL<sup>11</sup>, Wikidata URL or URI<sup>12</sup>, or Wikidata ID<sup>13</sup> to automatically load the entity from Wikidata, or,
- insert a name, description, and class to create a new entity.

The new entity popover is shown in Figure 7.6. When the user adds a new entity through the popover, the entity is saved to the IndexedDB database and added to the main list.

## 7.2.2 Event Creation Form

The right-hand side of the screen contains the main event creation form. Here the user can insert the basic data of each event, including:

- the event title (a string of text),
- the start date of the event (a string of text, automatically parsed),
- the end date of the event (a string of text, automatically parsed),

<sup>11</sup>For instance, [https://en.wikipedia.org/wiki/Dante\\_Alighieri](https://en.wikipedia.org/wiki/Dante_Alighieri).

<sup>12</sup>Wikidata URLs are of the form <https://wikidata.org/wiki/Q1076>, while Wikidata entity URIs are of the form <https://www.wikidata.org/entity/Q1076>.

<sup>13</sup>For instance, Q1076 for the entity representing Dante Alighieri.

FIGURE 7.6: The popover to add new entities.

- the event type (a string of text, selected from a list or inserted manually by the user),
- one or more Wikidata or user-defined entities that the user has identified as components of the event (added through drag-and-drop from the main entity list),
- an optional textual note (a string of text),
- an optional link to a digital object (a URL, automatically parsed).

A detailed view of the event creation form is shown in Figure 7.7.

When the user begins creating an event, she types the title of the event in the “Event Title” field of the event creation form. Then she types the start and end dates of the event, and she inserts an event type. An event type auto-complete menu is pre-populated from the list that was created in Section 6.2, but the user can manually insert new ones if needed.

The following step is the linking of the event to its components such as people, places, objects, etc. This is done through a simple drag-and-drop, which we developed using the jQuery UI library. When the user drags an entity from the left-hand table onto the “Entities” field that is found at the center of the event creation form, a popover appears, showing the user more data entry options. This popover is shown in Figure 7.8.

This popover, which we call “sourcing popover”, is used to enter information about the sources which the narrative is based on. The

FIGURE 7.7: View of the event creation form.

main sections in the popover are: Secondary source, Primary source(s), and Notes.

In the Secondary source section, the user can insert the book, chapter, and fragment of the secondary source stating that this particular entity took part in the event. For instance, if the user is a biographer of Dante Alighieri, she will insert the book, chapter, and fragment of the biography she has written.

In the Primary source section, the user can add one or more primary sources, including their author, title, the fragment of the secondary source that references it, and the fragment of the primary source itself. To facilitate the addition of primary sources and make sure that there are no duplicates, the primary source title and author fields have been constructed as auto-complete fields using the Typeahead.js library.

When the user inserts the title of a work, she is provided with a list of all works in Wikidata with that title. If she selects a particular work, the author's name is filled automatically. If instead she starts by inserting a name in the author field, she is provided with a list of all authors found in Wikidata under that name. Upon selecting a particular author from the list, the title field automatically changes to show only the works by that author.

The screenshot shows a 'Florence Baptistery' sourcing popover. At the top, three entity boxes are visible: 'Dante Alighieri' (pink), 'Florence Baptistery' (blue), and 'Florence' (teal). The popover itself is titled 'Florence Baptistery' and is divided into several sections:

- Secondary source:** Includes a 'Book' field with value '1' and a 'Chapter' field with value '3'. Below is a 'Text' field containing the text: 'A Firenze Dante fu battezzato sul fonte del Battistero intitolato a San'.
- Primary source (1):** Includes 'Author' (Dante Alighieri), 'Title' (Inferno), 'Fragment' (XIX 17), and 'Text' (il mio bel San Giovanni).
- Primary source (2):** Includes 'Author' (Dante Alighieri), 'Title' (Paradiso), 'Fragment' (XXV 8-9), and 'Text' (in sul fonte del mio battesimo).
- Notes:** A large empty text area for additional information.

At the bottom of the popover, there is a link for 'Wikipedia • Wikidata'.

FIGURE 7.8: View of the sourcing popover.

In both cases, the primary sources are stored internally with their title and corresponding Wikidata URIs. If the source or the author are not present in Wikidata, they are assigned custom URIs.

Finally, the last section allows the user to insert a textual note about the sources or the relation of the entity to the event.

If the entity is a person, the sourcing popover contains an additional section which allows the user to insert the role that the person played in the event (for instance, in an event of writing, the person who performed the writing will have role “writer”). The role section is present only for people, since other classes are not allowed to have a role in the ontology.

The role can be chosen from a list that is obtained through the process we described in Section 6.3. The list is initially filtered based on the specific entity, e.g. a journalist who also writes books would have “journalist” and “writer” as suggested roles. These role suggestions are obtained through SPARQL queries to the Wikidata knowledge base. An example is shown in Figure 7.9.

Additional suggestions are based on the event type, e.g. if the event is of type “birth” the role menu will contain “mother”, “father”, and “child”. However, the user can also bypass the suggestions and select

```

PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
SELECT DISTINCT ?occupation ?position
WHERE {
  OPTIONAL { wd:Q1067 wdt:P106 ?occupation. }
  OPTIONAL { wd:Q1067 wdt:P39 ?position. }
}

```

FIGURE 7.9: SPARQL query to extract occupations and positions held for entity *Q1067 Dante Alighieri*.

a role from the full list that we described in Section 6.3 by selecting “Other” in the menu. When the role is undefined, it defaults to “participant”.

When the user has filled the sourcing popover, she can close it and move on to another entity. After adding all related entities to the event, she can optionally add a textual note about the event and a link to a digital object. Then she can save the event or, if she made a mistake, clear the form and start again using the appropriate buttons on the right-hand side of the event form.

### 7.2.3 Default Events

In the case of biographical narratives, several of the events that the user would create can be inferred from the knowledge that is stored in Wikidata. For instance, when describing the life of Dante Alighieri, the user will probably have to insert an event titled *Death of Dante Alighieri* with a date of 14 September 1321. This information is already present in Wikidata, but since its ontology is not event-based it is not represented as an event.

Instead, Wikidata contains the statement *Q1067 (Dante Alighieri) P569 (date of death) 14 September 1321*, i.e. it uses the property *P569 date of death* to link the entity that represents Dante Alighieri to the date value “14 September 1321”.

In order to facilitate the user’s work, we developed a feature that extracts basic events from the knowledge contained in Wikidata and uses them to populate the timeline at the beginning of the narrative creation. The extraction is performed through an initial SPARQL query to the Wikidata Query Service that retrieves RDF triples having the

subject of the narrative as their subject. An example of SPARQL query is shown in 7.10.

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
SELECT DISTINCT ?birth ?death
WHERE {
  OPTIONAL { wd:Q1067 wdt:P569 ?birth. }
  OPTIONAL { wd:Q1067 wdt:P570 ?death. }
}
```

FIGURE 7.10: SPARQL query to extract birth and death dates for entity *Q1067 Dante Alighieri*.

The results are then used to populate the timeline with several “default events”. The user is then able to edit the default events, or delete them if they are not needed in her narrative. Currently, the tool is able to handle basic events such as *birth*, *death*, and *marriage* for humans, *foundation* for organizations, *creation* for physical objects. We plan to extend the list to other kinds of events in the future.

### 7.2.4 Event Timeline

The bottom side of the screen is a timeline containing a simple view of each event created by the user in chronological order. The timeline is shown in Figure 7.11. Whenever an event is saved or loaded as a default event, it is automatically added to the bottom timeline.

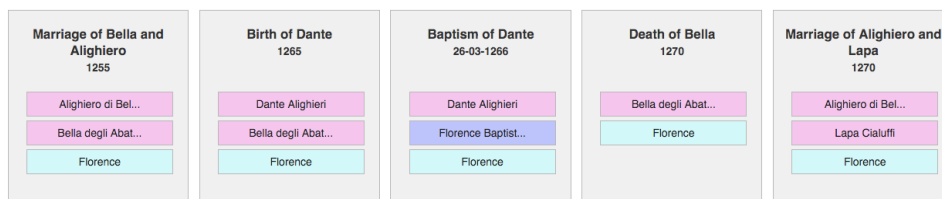


FIGURE 7.11: View of the event timeline.

Each event in the timeline shows an abridged representation of the data entered by the user, i.e. the title, the dates, and the entities that are related to the event. By clicking on an event in the timeline, the user can reload the event into the event creation form for subsequent editing and corrections. The user can also delete events from the timeline by

clicking on the button which appears in the top-right corner of each event when hovering the mouse cursor over it.

### 7.3 Causal and Mereological Relations

The user can also add relations among events by loading the third view of the interface, which is called “Causal and Mereological Relations” and can be accessed through a button on the right-hand side of the event creation form. In this view, shown in Figure 7.12, the user can link each event to the events which caused it, and also to those that contain it mereologically. These correspond to the relations that we defined in Chapter 5, except for the temporal properties which are assigned based on dates<sup>14</sup>

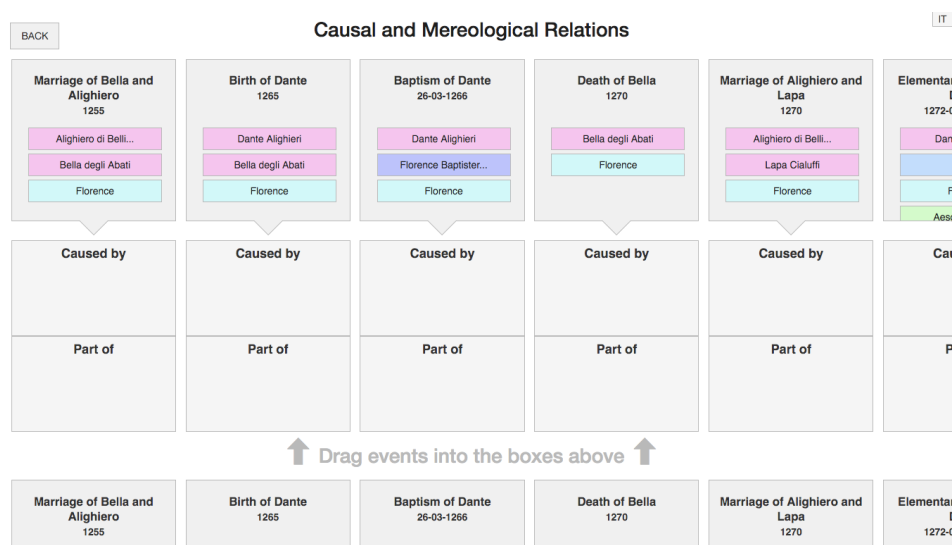


FIGURE 7.12: Causal and mereological relations.

When the user clicks on the “Relations” button, the bottom timeline slides up, revealing boxes where events can be freely dragged. These are taken from a duplicate timeline, which is displayed on the bottom of the screen. This way, the user can freely connect events to each other without losing sight of the bigger view.

At any moment, the user can go back to the Event Creation view using the “Back” button. All inserted relations are automatically saved in

<sup>14</sup>A method for representing uncertain dates in the tool has not yet been developed, therefore it is not currently possible to link events temporally without inserting dates (see also Section 10.1).



the IndexedDB database of the web application, thus a specific “Save” action is not present in the interface.

## 7.4 Subsequent Loadings and Export

On subsequent loadings of the tool, all data inserted by the user is automatically loaded from the IndexedDB database and the tool does not need to query Wikipedia nor Wikidata, thus significantly reducing the load time<sup>15</sup>. The list of entities and the timeline of events are re-created locally and can be freely edited by the user.

When the user has finished adding events and relations, she can download a JSON file containing the full list of events (including their relations) by using the “Export” button. The JSON file contains an array with all events in sequence, and each event is a JSON object endowed with the properties reported in Table 7.2.

TABLE 7.2: Properties of a JSON event.

Property	Value
<i>id</i>	<i>numeric ID of the event</i>
<i>title</i>	<i>title of the event</i>
<i>type</i>	<i>type of the event</i>
<i>start</i>	<i>start date of the event</i>
<i>end</i>	<i>end date of the event</i>
<i>objurl</i>	<i>optional URL of a digital object</i>
<i>notes</i>	<i>optional textual notes</i>
<i>props</i>	<i>key/value map of propositions</i>

Each proposition in the key/value map is identified by a key corresponding to the Wikidata or user-defined ID of the related entity that constitutes the object of the proposition. Each proposition is a JSON object endowed with the properties reported in Table 7.3.

<sup>15</sup>In our limited testing with the Dante Alighieri use case, the load time was reduced by about 50% when using the IndexedDB database. Proper performance testing will be performed in a future evaluation.

TABLE 7.3: Properties of a JSON proposition.

Property	Value
<i>name</i>	<i>name of the related entity</i>
<i>class</i>	<i>class of the related entity</i>
<i>role</i>	<i>optional role (applicable only to people)</i>
<i>book</i>	<i>optional book number of the secondary source</i>
<i>chapter</i>	<i>optional chapter number of the secondary source</i>
<i>fragment</i>	<i>optional fragment of the secondary source</i>
<i>notes</i>	<i>optional textual notes</i>
<i>sources</i>	<i>array of primary sources</i>

Each JSON proposition can contain an array of primary sources. Each primary source in the array is a JSON object endowed with the properties reported in Table 7.4.

TABLE 7.4: Properties of a JSON primary source.

Property	Value
<i>authorId</i>	<i>Wikidata or user-defined ID of the author</i>
<i>authorName</i>	<i>name of the author</i>
<i>sourceId</i>	<i>Wikidata or user-defined ID of the primary source</i>
<i>sourceTitle</i>	<i>title of the primary source</i>
<i>textFrag</i>	<i>optional textual fragment</i>
<i>refFrag</i>	<i>optional reference fragment</i>

The representation of an example event (the birth of Dante Alighieri), including three propositions and a single primary source, is shown in Figure 7.13.

```
{
  "id": "1",
  "title": "Birth of Dante Alighieri"
  "type": "birth",
  "start": "1265",
  "end": "1265",
  "book": "1",
  "chapter": "3",
  "fragment": "Certainly Dante was born in Florence",
  "notes": null,
  "objurl": null,
  "props": {
    "Q1067": {
      "name": "Dante Alighieri",
      "class": "person",
      "role": "child",
      "notes": "",
      "sources": []
    },
    "Q3637724": {
      "name": "Bella degli Abati",
      "class": "person",
      "role": "mother",
      "notes": "",
      "sources": []
    }
  }
  "Q2044": {
    "name": "Florence",
    "class": "place",
    "sources": [
      {
        "authorId": "Q1067",
        "authorName": "Dante Alighieri",
        "sourceId": "Q4509219",
        "sourceTitle": "Inferno",
        "textFrag": "io fui nato e cresciuto sopra 'l bel
          fiume d'Arno, alla gran villa",
        "refFrag": "XXIII 94-95"
      }
    ]
  }
}
```

FIGURE 7.13: JSON representation of the event *Birth of Dante Alighieri*.

## 7.5 Triplification

Once the user has completed the creation of the narrative, the events are saved to a JSON object along with the entities that compose them. We developed a Java software to automatically translate the JSON object to an OWL graph, in which the knowledge is represented as triples subject-predicate-object.

First of all, the software imports a JSON file and converts each event, each related entity, and each relation between events to a Java object. The Java objects are organized into three lists<sup>16</sup>: (i) a list of events, (ii) a list of related entities, and (iii) a list of relations between events.

The software uses the Apache Jena framework that provides a programmatic environment for OWL. Jena is an open source Semantic Web framework for Java. It provides an API to extract data from and write to RDF and OWL graphs. The graph is represented as an abstract model, which can be sourced with data from files, knowledge bases, URIs or a combination of these.

We used Jena to define an OWL model (reported in Appendix A), i.e. a grammar that describes the structure of our ontology. We defined the classes and properties that compose the model, including those that are not present in the reference ontologies and defined by us. Each class and property was defined as an OWL resource, and for each property we also indicated its domain and range.

The model is applied to the knowledge contained in the three Java array lists. The resulting graph is saved to file in RDF/XML<sup>17</sup> and Turtle<sup>18</sup> formats. Then, the OWL graph is imported into a Virtuoso triple store (Erling and Mikhailov, 2009) using the Virtuoso Jena Provider<sup>19</sup>.

Since we used the CIDOC CRM as reference ontology, we imported into Virtuoso the full OWL representation of the CRM. We adopted the Erlangen CRM/OWL<sup>20</sup>, an OWL DL 1.0 implementation of the CRM (Görz, Oischinger, and Schiemann, 2008). The Erlangen CRM/OWL

---

<sup>16</sup>Specifically, we use the ArrayList class from the Java Collections Framework.

<sup>17</sup><https://w3.org/TR/rdf-syntax-grammar/>

<sup>18</sup><https://w3.org/TR/turtle/>

<sup>19</sup><http://virtuoso.openlinksw.com/dataspace/doc/dav/wiki/Main/VirtJenaProvider>

<sup>20</sup><http://erlangen-crm.org/>

is “an interpretation of the CRM in a logical framework attempting to be as close as possible to the text of the specification“<sup>21</sup>. We also imported the full OWL representation of the FRBRoo ontology, also developed by Erlangen<sup>22</sup>.

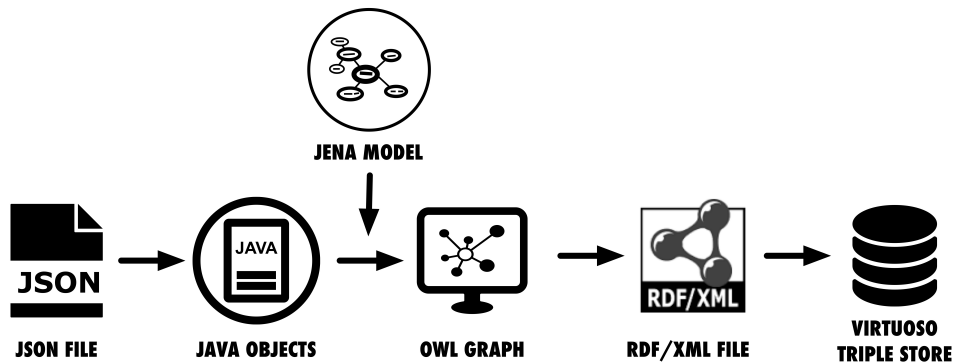


FIGURE 7.14: The process we followed to create the OWL graph.

Figure 7.14 summarizes the entire process that we currently follow to create the OWL graph and load it into the triple store:

1. we start from a JSON file, exported from the web interface;
2. we load the JSON data into appropriate Java objects;
3. we apply the Jena model to the inserted knowledge;
4. we obtain an OWL graph as a result;
5. we export the OWL graph as an RDF/XML file;
6. we load the RDF/XML file into a Virtuoso triple store.

At the end of the process, we obtain an OWL graph containing the knowledge that was inserted by the user in the tool. The triple store can be queried through SPARQL in order to extract knowledge for the following phase, the visualization.

<sup>21</sup><http://new.cidoc-crm.org/Resources/erlangen>

<sup>22</sup><http://erlangen-crm.org/efrbroo>



# Chapter 8

## Visualizations

Following the requirements outlined in Chapter 4, we developed several visualization components for visualizing the knowledge included in a narrative. For each visualization, a SPARQL query is performed on the knowledge base that has been created using the tool and the triplifier, and stored into the Virtuoso triple store.

### 8.1 Timeline Visualization

First of all, in order to give a complete overview of the events that compose the narrative, we placed them on a timeline, as shown in Figure 8.1. We adopted the TimelineJS<sup>1</sup> JavaScript library for the implementation. The timeline allows the user to visually navigate the semantic network of events.



FIGURE 8.1: Visualization of the events on a timeline.

<sup>1</sup><https://timeline.knightlab.com>

The timeline visualization shows, for each event, the following pieces of knowledge:

1. the title of the event;
2. the date of the event;
3. the fragment of the secondary source (if present);
4. the digital object representing the event (if present);
5. the list of primary sources (if present);
6. an image describing one of the related entities, if available, or otherwise a widget linking to the Wikipedia page which describes the entity.

By clicking on a specific button in the Figure 8.2 shows an example of the visualization of a digital object from the timeline. The image is shown in a Bootstrap popover, similar to those that we showed in use in the tool in Chapter 7.

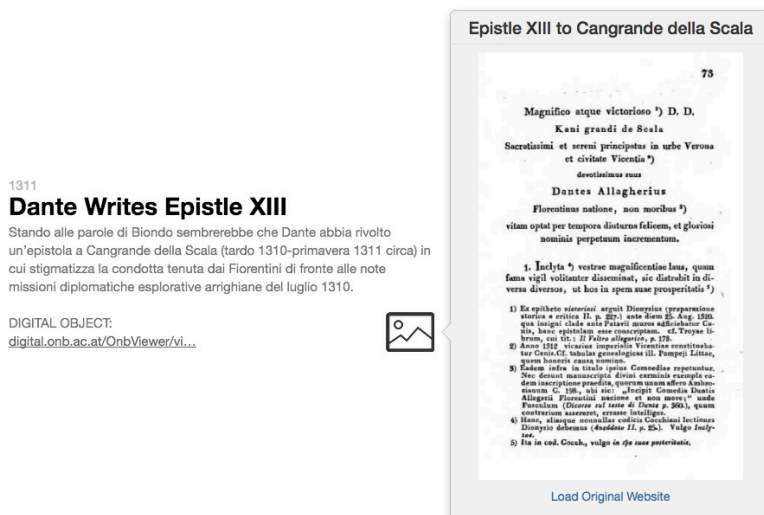


FIGURE 8.2: View of a digital object in the timeline.

The pieces of knowledge from 1 to 6 are directly extracted from the knowledge base using a SPARQL query. The query is shown in Figure 8.3.



```

PREFIX cnt: <http://www.w3.org/2011/content#>
PREFIX narr: <http://dantesources.org/narratives/>
PREFIX crm: <http://erlangen-crm.org/current/>

SELECT ?eventTitle ?startDate ?endDate ?fragmentText ?authorURI
       ?primarySourceURI ?primaryTextFrag ?primaryRefFrag ?reference
FROM <http://dantesources.org/narratives/example>

WHERE {
  ?eventURI crm:P149_is_identified_by ?eventAppellation.
  ?eventAppellation crm:P3_has_note ?eventTitle.
  ?eventURI crm:P4_has_time-span ?timeSpan.
  ?timeSpan crm:P79_beginning_is_qualified_by ?startDate.
  ?timeSpan crm:P80_end_is_qualified_by ?endDate.

  ?propositionURI rdf:subject ?eventURI.
  ?propositionURI rdf:object ?entityURI.

  ?exprFrag crm:P129_is_about ?propositionURI.
  ?exprFrag cnt:chars ?fragmentText.
  ?propositionURI narr:hasSource ?primarySourceURI.
  ?creationEvent crm:P94_has_created ?primarySourceURI.
  ?creationEvent crm:P14_carried_out_by ?authorURI.

  ?primarySourceURI crm:P149_is_identified_by ?primarySourceAppellation.
  ?primarySourceAppellation crm:P3_has_note ?primarySourceTitle.
  ?authorURI crm:P149_is_identified_by ?authorAppellation.
  ?authorAppellation crm:P3_has_note ?authorName.

  ?propositionURI narr:hasTextFragment ?primaryTextFrag.
  ?primarySourceURI crm:P106_is_composed_of ?primaryTextFrag.
  ?primaryTextFrag narr:hasReference ?reference.
}

```

FIGURE 8.3: SPARQL query to extract the knowledge shown in the timeline visualization.

In order to load the last piece of knowledge that is present in the timeline visualization, i.e. an image of a related entity or a widget linking to the Wikipedia page describing the entity, we perform a SPARQL query on the Wikidata Query Service to extract all available images of Wikidata entities that are present in the narrative.

The images are not stored in Wikidata but in the Wikimedia Commons<sup>2</sup> repository, a related project of the Wikimedia Foundation. Nevertheless, they are easily accessible from the WQS. The SPARQL query that we use to extract them is shown in Figure 8.4.

```
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
SELECT DISTINCT ?entity ?image
WHERE {
    ?entity wdt:P18 ?image.
}
VALUES {?entity [list of Wikidata entities]}
```

FIGURE 8.4: SPARQL query to extract images from Wikidata.

If no image is available, we perform a second SPARQL query on the Wikidata Query Service to extract links to Wikipedia pages describing the entities. The SPARQL query is shown in Figure 8.5.

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX schema: <http://schema.org/>
SELECT ?url
WHERE
{
    ?url schema:about wd:Q1067.
    ?url schema:inLanguage "en".
    ?url schema:isPartOf <https://en.wikipedia.org/>.
}
VALUES {?uri [list of Wikidata entities]}
```

FIGURE 8.5: SPARQL query to extract links to the English Wikipedia pages describing the entities.

For each event, the visualization component checks if an image was retrieved from the results of the first query and, if available, it visualizes it. An example is shown in Figure 8.6.

<sup>2</sup><https://commons.wikimedia.org>



Georges Jansoone, *Battistero di San Giovanni*

(from Wikimedia Commons)

March 26, 1266

### Baptism of Dante

A Firenze Dante fu battezzato sul fonte del Battistero intitolato a San Giovanni

DIGITAL OBJECT:

[www.europeana.eu/portal/record...](http://www.europeana.eu/portal/record...)

PRIMARY SOURCES:

- Dante Alighieri, *Inferno* XIX 17
- Dante Alighieri, *Paradiso* XXV 8-9

FIGURE 8.6: Visualization of the image of the Baptistery of Florence in the timeline event *Baptism of Dante*.

If no image is available, the visualization component loads the corresponding Wikipedia page from the results of the second query and automatically shows an excerpt from it, using the rich media functionality provided by the TimelineJS library<sup>3</sup>. An example of this visualization is shown in Figure 8.7.

<p><b>W</b> <b>Henry VII, Holy Roman Emperor</b>  <small>From Wikipedia, the free encyclopedia</small></p> <p><b>Henry VII</b> (German: <i>Heinrich</i>; c. 1275 – 24 August 1313) was the King of Germany (or <i>Rex Romanorum</i>) from 1308 and Holy Roman Emperor from 1312. He was the first emperor of the House of Luxembourg. During his brief career he reinvigorated the imperial cause in Italy, which was racked with the partisan struggles between the divided Guelph and Ghibelline factions, and inspired the praise of Dino Compagni and Dante Alighieri; however, his premature death undid his life's work.</p>	<p>December 1, 1310 — January 31, 1311</p> <p><b>Dante Meets Henry VII</b></p> <p>Sebbene ignoriamo ogni dettaglio, che Dante abbia raggiunto la corte di Arrigo a poca distanza dal suo ingresso in Italia è fuori d'ogni sensata discussione. Dell'incontro (milanese?) riferisce il poeta stesso, con la consueta assenza di dettagli, nell'Epistola VII (§ 9)</p> <p>PRIMARY SOURCES:</p> <ul style="list-style-type: none"> <li>• Dante Alighieri, <i>Epistle</i> XII Nam et ego qui scribo tam pro me quam pro aliis, velut decet imperatoriam maiestatem benignissimum vidi et clementissimum te audivi, cum pedes tuos manus mee tractarunt et labia mea debitum persolverunt</li> </ul>
--	--

FIGURE 8.7: Visualization of the Wikipedia page about Henry VII in the timeline event *Dante Meets Henry VII*.

## 8.2 Event-Centered Graph

Another requirement defined in agreement with the scholars is the visualization of the entities that compose each event. We developed a SPARQL query, shown in Figure 8.8, to extract this information.

<sup>3</sup><https://timeline.knightlab.com/docs/media-types.html>

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX crm: <http://erlangen-crm.org/current/>

SELECT ?eventTitle ?entityURI ?entityName
FROM <http://dantesources.org/narratives/example>

WHERE {
  ?eventURI crm:P149_is_identified_by ?eventAppellation.
  ?eventAppellation crm:P3_has_note ?eventTitle.
  ?propositionURI rdf:subject ?eventURI.
  ?propositionURI rdf:object ?entityURI.
  ?entityURI crm:P149_is_identified_by ?entityAppellation.
  ?entityAppellation crm:P3_has_note ?entityName.
}

```

FIGURE 8.8: SPARQL query to extract all entities related to all events that compose the narrative.

In Figure 8.9 we report an example of visualization of an event with its related entities. We used the `vis.js`<sup>4</sup> library to implement the visualization. The network graph visualizes the event at its center and the related entities are represented with different icons according to their ontological classes. By clicking on each entity, the user can visualize the corresponding Wikipedia page or, for user-created entities, the textual description of the entity stored in the knowledge base.

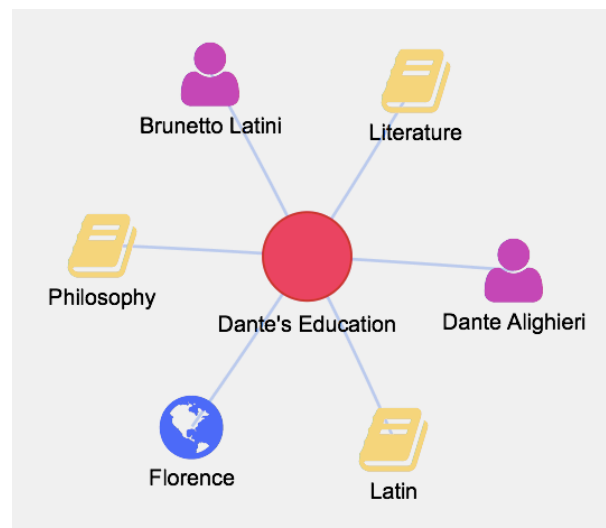


FIGURE 8.9: Visualization of the event *Dante's Education* and its related entities.

<sup>4</sup><http://visjs.org>

## 8.3 Entity-Centered Graph

In order to extract all events that are related to a specific entity, e.g. all events that occurred in the city of Florence, we developed a SPARQL query to extract this information from our knowledge base and we presented it as a vis.js network graph. The query is shown in Figure 8.10.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX crm: <http://erlangen-crm.org/current/>
PREFIX wd: <http://www.wikidata.org/entity/>

SELECT ?eventURI ?eventTitle
FROM <http://dantesources.org/narratives/example>

WHERE {
  ?eventURI crm:P149_is_identified_by ?eventAppellation.
  ?eventAppellation crm:P3_has_note ?eventTitle.
  ?propositionURI rdf:subject ?eventURI.
  ?propositionURI rdf:object wd:Q2044.
}
```

FIGURE 8.10: SPARQL query to extract all events which are linked to the specific entity *Florence*.

In Figure 8.11 we report an example of visualization of an entity with all the events which the entity is linked to.

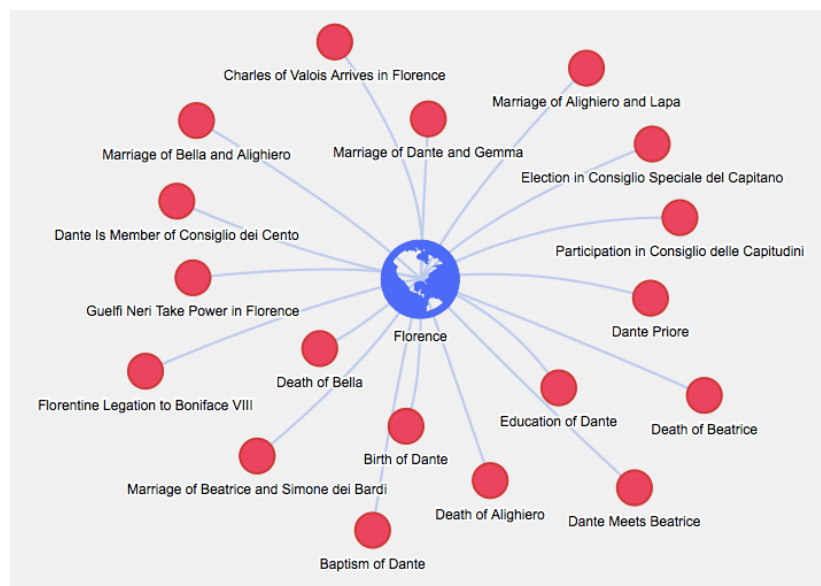


FIGURE 8.11: Visualization of the entity *Florence* and its related events.

At the center of the graph is the entity “Florence”, which belongs to the ontological class *E53 Place* from the CRM. The edges connect this entity to all the events that took place in Florence, represented with a circle along with their titles.

## 8.4 Primary Sources Table

One of the most important requirements for a scholar who studies historical events is the knowledge of their primary sources. In our case study, for each event, the user can add the following information about primary sources:

- the title of the primary source
- the author of the primary source
- a textual fragment of the primary source describing the event
- the bibliographic reference of the textual fragment

For the visualisation of this knowledge we adopted a table format, as shown in Figure 8.12.

Event	Date	Primary Source	Author	Text Fragment	Reference Fragment
Birth of Dante	1265	Inferno	Dante Alighieri	'io fui nato e cresciuto sovra 'l bel fiume d'Arno, alla gran villa'	XXIII 94-95
Baptism of Dante	1266-03-26	Inferno	Dante Alighieri	'il mio bel San Giovanni'	XIX 17
Baptism of Dante	1266-03-26	Paradiso	Dante Alighieri	'in sul fonte del mio battesimo'	XXV 8-9
Marriage of Dante and Gemma	1277-02-09	Instrumentum Dotis	Uguccione di Baldovino	'Domine Gemme vidue, uxori olim Dantis Allagherii et filie condam domini Manetti domini Donati, pro iure sue dotis librarum CC florenorum parvorum ut de instrumento dotis constat manu ser Ranaldi filii condam Oberti Baldovini de Florentia notarii, ex inbreviaturis ser Uguiccionis Baldovini notarii, facto in anno Domini M° CCLXXVI, indictione VI, die VIIIJ' mensis februarii'	Capitani di Parte Guelfa, Numeri Rossi, 42, c. 3r
Dante's First Travel to Bologna	1286-1287	Trattatello in laude di Dante	Giovanni Boccaccio	'Egli li primi iniziï, sì come di sopra è dichiarato, prese nella propria patria, e di quella, sì come a luogo più fertile di tal cibo, n'andò a Bologna'	

FIGURE 8.12: Table view of the events and their primary sources.

When an event has more than one primary source, the table shows a row for each source, e.g. the event *Baptism of Dante* has two primary sources: Dante's *Inferno* XIX 17 and *Paradiso* XXV 8-9.

It is also possible that the reference fragment is absent, such as when the primary source has no internal subsections, e.g. Giovanni Boccaccio's work *Trattatello in laude di Dante*.

The SPARQL query that we used to extract this information from the knowledge base is shown in Figure 8.13.

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX narr: <http://dantesources.org/narratives/>
PREFIX crm: <http://erlangen-crm.org/current/>
PREFIX cnt: <http://www.w3.org/2011/content#>

SELECT DISTINCT ?eventTitle ?startDate ?endDate ?primarySourceTitle ?authorName ?t
FROM <http://dantesources.org/narratives/example>

WHERE {
  ?eventURI crm:P149_is_identified_by ?eventAppellation.
  ?eventAppellation crm:P3_has_note ?eventTitle.
  ?propositionURI rdf:subject ?eventURI.
  ?propositionURI narr:hasSource ?primarySourceURI.
  OPTIONAL {
    ?propositionURI narr:hasTextFragment ?textFragment.
    ?textFragment cnt:chars ?text.
    ?primarySourceURI crm:P106_is_composed_of ?textFragment.
  }
  OPTIONAL {
    ?textFragment narr:hasReference ?refFragment.
    ?refFragment cnt:chars ?reference.
  }
  ?creationEvent crm:P94_has_created ?primarySourceURI.
  ?creationEvent crm:P14_carried_out_by ?authorURI.
  ?primarySourceURI crm:P149_is_identified_by ?primarySourceAppellation.
  ?primarySourceAppellation crm:P3_has_note ?primarySourceTitle.
  ?authorURI crm:P149_is_identified_by ?authorAppellation.
  ?authorAppellation crm:P3_has_note ?authorName.
  ?eventURI crm:P4_has_time-span ?timeSpan.
  ?timeSpan crm:P79_beginning_is_qualified_by ?startDate.
  ?timeSpan crm:P80_end_is_qualified_by ?endDate.
}

```

FIGURE 8.13: SPARQL query to extract all events with their dates and their primary sources.

## 8.5 Events in a Specific Time Range

A user has the possibility to select a specific range of time and visualize all events that occurred in that period. The user can freely insert the start and end of the period using a widget. To make sure that the query always returns a result, we show the start and end dates of the narrative as boundaries. Figure 8.14 shows the SPARQL query to extract events that happened between 1265 and 1277.

```

PREFIX crm: <http://erlangen-crm.org/current/>

SELECT ?eventTitle ?startDate ?endDate
FROM <http://dantesources.org/narratives/example>

WHERE {
  ?eventURI crm:P149_is_identified_by ?eventAppellation.
  ?eventAppellation crm:P3_has_note ?eventTitle.
  ?eventURI crm:P4_has_time-span ?timeSpan.
  ?timeSpan crm:P79_beginning_is_qualified_by ?startDate.
  ?timeSpan crm:P80_end_is_qualified_by ?endDate.

  FILTER (xsd:integer(substr(?startDate, 1, 4)) > 1265).
  FILTER (xsd:integer(substr(?endDate, 1, 4)) < 1277).
}
ORDER BY ?startDate

```

FIGURE 8.14: SPARQL query to extract all events that occurred from 1265 to 1277.

The results of the query are displayed in a table, where shows the dates for each event.

Event	Date
Baptism of Dante	1266-03-26
Death of Bella	1270
Marriage of Alighiero and Lapa	1270
Elementary Education of Dante	1272-1277
Death of Alighiero	1276
Marriage of Dante and Gemma	1277-02-09

FIGURE 8.15: Visualization of the events that occurred in a selected time range.



## 8.6 Relations Between Events

In our conceptualization, we defined three types of relations between events: mereological, temporal and causal. We developed a visualization component that shows these relations using a network graph. Figure 8.16 shows a SPARQL query which extracts this information.

```

PREFIX crm: <http://erlangen-crm.org/current/>

SELECT ?event1Title ?event2Title
FROM <http://dantesources.org/narratives/example>

WHERE {
  ?event1 ?crm:P9_consists_of ?event2
  ?event1 crm:P149_is_identified_by ?eventAppellation.
  ?eventAppellation crm:P3_has_note "Education of Dante".
  ?event2 crm:P149_is_identified_by ?eventAppellation.
  ?eventAppellation crm:P3_has_note ?event2Title.
}

```

FIGURE 8.16: SPARQL query to extract all sub-events that compose the super-event *Education of Dante*.

In Figure 8.17 we represented the mereological relations between the event *Education of Dante* and its sub-events, again using the vis.js library for the implementation. The figure also shows the temporal relations that occur between the sub-events.

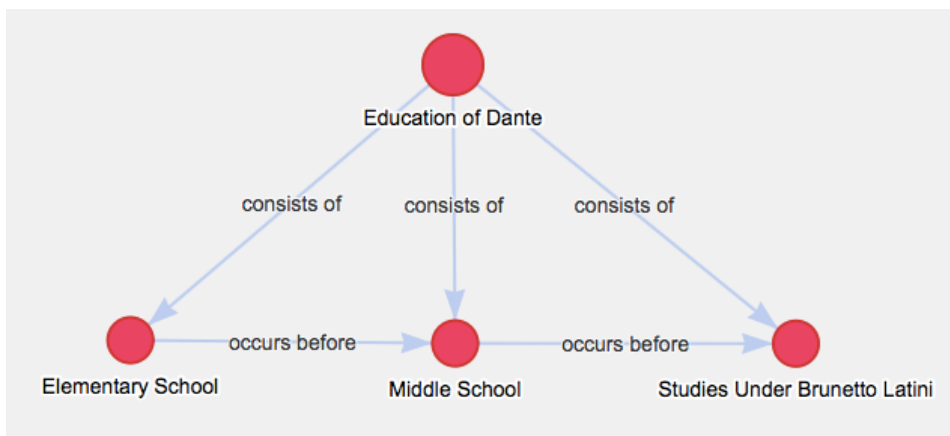


FIGURE 8.17: Visualization of the mereological relations and temporal between events.



## Chapter 9

# The Dante Alighieri Case Study

After developing our tool for the creation of narratives, we applied it to the case study we had previously selected: the biography of late Middle Ages Italian poet Dante Alighieri. We chose this particular poet because we are collaborating to the Italian National Research Project “Towards a Digital Dante Encyclopedia”<sup>1</sup>, which aims to create a digital encyclopedia about Dante Alighieri and his works. The project has produced three web applications: DanteSources<sup>2</sup>, a digital library about Dante’s primary sources; DanteSearch<sup>3</sup>, a lemmatization of the works of Dante; and DaMA<sup>4</sup>, a digital archive containing the text of Dante’s works and several of his primary sources.

### 9.1 Narrative Representation

One of the two experts provided us with a detailed text of 40 pages (83,688 characters) describing the main events of Dante’s life. From this text, we identified and extracted the events and representing them using the tool. The tool automatically imported from Wikidata the entities corresponding to the main people, places, objects, concepts, and works that were involved in Dante’s life. We used these Wikidata entities to represent the events that we found in the text provided by

---

<sup>1</sup><http://perunaenciclopediadantescadigitale.eu>

<sup>2</sup><http://dantesources.org>

<sup>3</sup><http://perunaenciclopediadantescadigitale.eu:8080/dantesearch/>

<sup>4</sup><http://perunaenciclopediadantescadigitale.eu/istidama/>

the expert. When a specific entity was not present in Wikidata, we added it using the form described in Section 7.2.1.<sup>5</sup>

In the beginning, a few default events (see Section 7.2.3) were automatically imported from Wikidata by the tool, including the birth, death, and marriage of Dante, and the creation events of his main works. Starting from his birth in 1265, we added all events that we found in the text in which Dante participated directly until his death in 1321. We also added some events that we deemed important as context, such as the marriage of his parents in 1255, the crownings and deaths of emperors or the elections and deaths of popes, and the death of his wife Gemma Donati in 1343.

For each event we inserted a title, an event type, its start and end dates and, when available, a related digital object. Then we linked the people or organizations that participated in the event, the place where the event happened, and the related objects, concepts and works. For each of these links, representing a proposition in our ontology (see Section 5.3), we added a fragment of the secondary source that we were using (i.e. the text provided to us by the expert) and also one or more primary sources cited by the expert.

The final result is a timeline of 80 events and 206 propositions, whose representation through the tool took about eight hours of work for two people. The events were subsequently linked to each other using mereological and causal relations (see Section 5.3). The resulting JSON file was loaded into the triplifier which produced an RDF/XML file, which was finally loaded into the Virtuoso triple store and visualized using the visualization components presented in Chapter 8.

## 9.2 Limited Evaluation

A proper full-scale evaluation of the tool has not yet been started. Nevertheless, we can hereby report the first results of the initial testing phase of the tool, which involved the representation of the biography of Dante Alighieri and enlisted the help of the two experts on Dante's life and works who collaborated with us during the development of

---

<sup>5</sup>An alternate solution to this problem could be the direct addition of these entities to Wikidata, given its open and collaborative nature. The only requirement is that they be within the scope of the knowledge base (see also Section 10.1).

the project. Specifically, we can report whether, in the opinion of the two experts, the tool is able to fulfill all of the requirements that were identified in Section 4.2.

Through semi-structured interviews, we specifically asked the two experts if, in their opinion, the tool and the visualization components are able to satisfy the requirements outlined in Chapter 4.2. Their opinion was very positive. Specifically, they liked the ease-of-use and visual nature of the tool, the possibility to cite secondary and primary sources, and also the fact that the visualization components provide a bird's eye view on the complex web of events that is Dante's life.

The experts also confirmed that, in their opinion, the tool fulfills all of the requirements that we had identified with their help. Specifically:

- the tool allows a semi-automated creation of the events of the narrative, thanks to the reliance on the Wikidata knowledge base to import resources and reduce the data insertion by the user;
- through the event creation form shown in Section 7.2.2, the tool allows: (i) the definition of the factual aspects of each event (such as time, persons, locations, physical objects, and conceptual objects), (ii) the definition of event types, choosing from an expandable list of pre-defined options, (iii) the linkins of a digital object (e.g. an object found in a digital library) with a specific event;
- through the Sourcing Popover shown in Section 7.2.2, the tool allows: (i) the identification of the role that a specific person performed in an event, (ii) the storage of a textual fragment representing an event in the narrative, (iii) the connection of a textual fragment to its author, (iv) the storage of the primary sources from which the events are defined, i.e. the provenance of the inserted knowledge;
- through the Relations view described in Section 7.3, the tool permits the linking of events to each other using the appropriate semantic relations (causal and mereological in an explicit way, temporal through the insertion of dates);
- the tool allows the export of the narrative in JSON format, as explained in Section 7.4.

The experts also verified that, in their opinion, the visualization components that we have developed also fulfill the functionalities that they had deemed essential during the identification of requirements:

- extracting the events along with their primary sources, as shown in Sections 8.1 and 8.2;
- extracting the events with a specific primary source, as shown in Section 8.4;
- extracting the events featuring a particular related entity (location, person etc.), as shown in Section 8.3;
- extracting the events that happened in a defined range of time, as shown in Section 8.5;
- extracting the events that are linked by a causal or mereological relation, as shown in Section 8.6.

We are currently in the process of further refining the biography of Dante with the help of the experts. Eventually, the result of this work will be published on the DanteSources website. Later on, we intend to perform a full-scale evaluation of the interface of the tool and also of the visualization components.

## Chapter 10

# Conclusions

In this thesis we presented a tool for the creation of narratives. Starting from a narratology and Semantic Web background, we described the full development process of the tool. We described the selection of a case study, the biography of Dante Alighieri, the identification of the requirements with the aid of two experts, and the ontology for narratives which the tool is based on.

We motivated the choice of Wikidata as a knowledge base from which to import entities in order to facilitate the work of the users, and presented the challenges that we encountered in mapping the Wikidata ontology to the CIDOC CRM. We presented in detail the interface of the tool, the triplification component that transforms the inserted data in OWL, and a few visualization components that we are developing. Finally, we also reported the first results of a limited evaluation of the tool by the experts that collaborated with us.

The end result is a tool that can be used to represent historical narratives (with a special focus on biographies) and, with minor adjustments or a little more effort on the part of the user, can be used for fictional narratives as well. As we said in the Introduction, we think the tool is able to fulfill the creation, construction, and curation aspects of narrative representation.

It is not limited to a specific narrative or case study, instead being open to any kind of user input. We envision it being used in the Digital Humanities to aid the work of historians, teachers, or journalists, and also to improve the search functionalities of digital libraries by linking digital objects through narratives.

However, the tool is by no means limited to these applications. We hope that when it is released publicly it will be used in novel ways

that we have yet to imagine. We further hope that the narratives that are created with the tool will be shared and compared, enriching the critical discussion of historical events and the creative potential of fictional stories alike.

## 10.1 Future Developments

In Chapter 9 we verified that, according to the expert, the tool is able to fulfill all the requirements that we identified in Chapter 4.2. However, this limited set of requirements is by no means exhaustive. We have envisioned a series of possible future developments that could further enhance and refine the tool, in order to make it even better for users.

**Better Integration.** The next step in the development of the tool is a better integration between its components. To this end, we will develop a backend which will tie together the web interface, the triplification component and the visualization components. This will allow the users to move from the knowledge insertion phase to the final result in a much more seamless way.

**Full-Scale Evaluation.** A crucial next step in the development of the tool is a full-scale evaluation involving a higher number of users. The results of this user test will allow us to identify criticalities that we might have overlooked in the current version of the tool and solve them in future updates, in order to provide a better experience to users.

**Bypassing Wikipedia.** Currently, the initial load of the tool makes a query to the Wikipedia API to obtain all entities linked from the page dedicated to the subject of the narrative. We initially deemed this step necessary since it would allow us to select all entities that have a relation with the subject, even though this relationship has not yet been formalized in Wikidata. However, since the beginning of our project Wikidata has matured and we now feel confident enough to test the removal of this initial step. A side-effect of this removal would be the ability to start a narrative for any entity contained in Wikidata, not just those that have a corresponding Wikipedia page in the specific language of the web interface.



**Giving Back to Wikidata.** Another interesting development that could be explored is the addition of a functionality to add knowledge to Wikidata, instead of just loading knowledge from it. This is very easy to do using the tools that Wikidata provides, and it is not only permitted by the rules of Wikidata, but encouraged<sup>1</sup>. For instance, a user could be using the tool to represent the biography of Dante Alighieri and suddenly realize that his stepmother, Lapa Cialuffi, is not present in Wikidata. With a simple click of a button, the entity that the user has created in the tool could be immediately transferred to Wikidata and automatically linked with the Wikidata entities representing Dante and his father Alighiero.

**Better Representation of Time.** One of the challenges that we have encountered during the development of the tool is the representation of time. Currently, the tool accepts as input well-formed dates in the ISO 8061<sup>2</sup> format using literals with the `xsd:date` datatype<sup>3</sup>. We are considering improving the representation by adopting the OWL Time Ontology<sup>4</sup>, however this does not appear sufficient to represent uncertain or vague dates such as “sometimes between Spring and Summer of 1301”, which are often encountered in historical research. We will continue to explore alternative solutions to this issue.

**Advanced Visualizations.** Currently, the visualization components that we have developed are very simple in nature and separated from one another. We concentrate on a specific aspect of the narrative and show the user a graphical representation of it. However, we envision more advanced visualizations where the user could inspect the narrative in finer detail. For instance, the timeline visualization could be provided with a new functionality that, for every event or connected entity, allows the user to visualize the graph centered on that event or entity. Other possible visualizations that we are considering include maps, social graphs, and genealogies.

---

<sup>1</sup>[https://wikidata.org/wiki/Wikidata:Data\\_donation](https://wikidata.org/wiki/Wikidata:Data_donation)

<sup>2</sup><http://iso.org/iso/home/standards/iso8601.htm>

<sup>3</sup><https://w3.org/TR/xmlschema11-2/#date>

<sup>4</sup><https://w3.org/TR/owl-time>

**Named Entity Recognition.** Another functionality that is missing from the tool is the possibility to directly import a text and automatically extract the entities found in it. This feature, called Named Entity Recognition (Nadeau and Sekine, 2007), would significantly enhance the narrative creation process, therefore we are seriously considering it as an addition to the tool. The difficulty would be making sure that the recognition is accurate, especially with regard to the linking of the recognized entities with the corresponding ones from Wikidata.

## 10.2 Final Remarks

We started this thesis by talking about stories, what they are, what they are made of and what they are used for. This thesis is not the place to answer this question. A simple interface for inserting events and visualizing the results may seem too little to help further this study, and maybe even somewhat out-of-scope. However, it is only through the continuous analysis of existing narratives and the process of their creation that we can truly understand their inner nature.

A standard Ontology of Narratives would certainly help further this goal. This is why we have contributed to its development, and the first application built on it is indeed the tool that we have presented in this thesis. Moreover, we think that the coming together of a rigorous Ontology of Narratives with a collaborative, community-driven knowledge base such as Wikidata can help further the scope of narratological research. The present work is just a first step in this direction.

It is only through rigorous evaluation that we will be able to tell if this is the right direction. Be it historians or teachers or users of digital libraries, whoever uses the tool or interacts with the resulting narratives will hopefully find her workflow enriched by our contribution, and it is by this empowering of users that we will measure the success of the tool. And maybe, with the help of our small contribution, narratives will finally find their rightful place in our digital world.

# Appendix A

## OWL Schema

This appendix contains the provisional OWL schema of the Ontology of Narratives in RDF/XML format. We listed all classes and properties that we directly re-used from the CIDOC CRM, FRBRoo, and CRMinf ontologies, along with all classes and properties that are newly defined. For the CRM, we used Erlangen CRM<sup>1</sup> as reference encoding. For FRBRoo, we used Erlangen FRBRoo<sup>2</sup>. The schema is compatible with the OWL DL<sup>3</sup> profile. The current version of the schema is provided here for reference, however please note that it is provisional and may be subject to change in the future.

### A.1 Schema

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:cnt="http://www.w3.org/2011/content#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcmi="http://purl.org/dc/dcmitype/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:ecrm="http://erlangen-crm.org/current/"
```

---

<sup>1</sup><http://erlangen-crm.org>

<sup>2</sup><http://erlangen-crm.org/efrbroo>

<sup>3</sup><https://w3.org/TR/owl-ref/#OWLDL>

```

xmlns:efrbroo="http://erlangen-crm.org/efrbroo/"
xmlns:crminf="http://www.cidoc-crm.org/cidoc-crm/CRMinf/"
xmlns:crmsci="http://www.ics.forth.gr/isl/CRMsci/"
xmlns:narr="http://dantesources.org/narratives/"
>

<!-- CIDOC-CRM Classes -->

<owl:Class rdf:about="http://erlangen-crm.org/current/E5_Event">
  <rdfs:label xml:lang="en">E5 Event</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E4_Period"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:someValuesFrom>
        <owl:Class
          rdf:about="http://erlangen-crm.org/current/E77_Persistent_Item"/>
        </owl:someValuesFrom>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P12_occurred_in_the_presence_of"/>
        </owl:onProperty>
      </owl:Restriction>
    </rdfs:subClassOf>
  </owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E6_Destruction">
  <rdfs:label xml:lang="en">E6 Destruction</rdfs:label>
  <rdfs:subClassOf
    rdf:resource="http://erlangen-crm.org/current/E64_End_of_Existence"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:someValuesFrom
        rdf:resource="http://erlangen-crm.org/current/E18_Physical_Thing"/>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P13_destroyed"/>
        </owl:onProperty>
      </owl:Restriction>
    </rdfs:subClassOf>
  </owl:Class>

```

```
<owl:Class rdf:about="http://erlangen-crm.org/current/E7_Activity">
  <rdfs:label xml:lang="en">E7 Activity</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E5_Event"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:someValuesFrom
        rdf:resource="http://erlangen-crm.org/current/E39_Actor"/>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P14_carried_out_by"/>
        </owl:onProperty>
      </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E12_Production">
  <rdfs:label xml:lang="en">E12 Production</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E11_Modification"/>
  <rdfs:subClassOf
    rdf:resource="http://erlangen-crm.org/current/E63_Beginning_of_Existence"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P108_has_produced"/>
        </owl:onProperty>
      <owl:someValuesFrom>
        <owl:Class
          rdf:about="http://erlangen-crm.org/current/E24_Physical_Man-Made_Thing"/>
        </owl:someValuesFrom>
      </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E19_Physical_Object">
  <rdfs:label xml:lang="en">E19 Physical Object</rdfs:label>
  <rdfs:subClassOf
    rdf:resource="http://erlangen-crm.org/current/E18_Physical_Thing"/>
  <rdfs:subClassOf>
    <owl:Restriction>
```

```

    <owl:maxCardinality
      rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 1
    </owl:maxCardinality>
    <owl:onProperty>
      <owl:ObjectProperty
        rdf:about="http://erlangen-crm.org/current/P54_has_current_permanent_location"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:maxCardinality
      rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 1
    </owl:maxCardinality>
    <owl:onProperty>
      <owl:ObjectProperty
        rdf:about="http://erlangen-crm.org/current/P55_has_current_location"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E21_Person">
  <rdfs:label xml:lang="en">E21 Person</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E39_Actor"/>
  <rdfs:subClassOf
    rdf:resource="http://erlangen-crm.org/current/E20_Biological_Object"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P152_has_parent"/>
        </owl:onProperty>
      <owl:minCardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#int"> 2
    </owl:minCardinality>
    </owl:Restriction>
  </rdfs:subClassOf>
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:onProperty>

```

```
        <owl:ObjectProperty
            rdf:about="http://erlangen-crm.org/current/P98i_was_born"/>
        </owl:onProperty>
    <owl:cardinality
        rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 1
    </owl:cardinality>
</owl:Restriction>
</rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E28_Conceptual_Object">
    <rdfs:subClassOf
        rdf:resource="http://erlangen-crm.org/current/E71_Man-Made_Thing"/>
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:onProperty>
                <owl:ObjectProperty
                    rdf:about="http://erlangen-crm.org/current/P94i_was_created_by"/>
                </owl:onProperty>
                <owl:someValuesFrom
                    rdf:resource="http://erlangen-crm.org/current/E65_Creation"/>
            </owl:Restriction>
        </rdfs:subClassOf>
    </owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E36_Visual_Item">
    <rdfs:label xml:lang="en">E36 Visual Item</rdfs:label>
    <rdfs:subClassOf
        rdf:resource="http://erlangen-crm.org/current/E73_Information_Object"/>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E39_Actor">
    <rdfs:label xml:lang="en">E39 Actor</rdfs:label>
    <rdfs:subClassOf
        rdf:resource="http://erlangen-crm.org/current/E77_Persistent_Item"/>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E41_Appellation">
    <rdfs:label xml:lang="en">E41 Appellation</rdfs:label>
    <rdfs:subClassOf
        rdf:resource="http://erlangen-crm.org/current/E90_Symbolic_Object"/>
```

```

</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E52_Time-Span">
  <rdfs:label xml:lang="en">E52 Time-Span</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E1_CRM_Entity"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:someValuesFrom>
        <owl:Class
          rdf:about="http://erlangen-crm.org/current/E2_Temporal_Entity"/>
        </owl:someValuesFrom>
        <owl:onProperty>
          <owl:ObjectProperty
            rdf:about="http://erlangen-crm.org/current/P4i_is_time-span_of"/>
          </owl:onProperty>
        </owl:Restriction>
      </rdfs:subClassOf>
      <rdfs:subClassOf>
        <owl:Restriction>
          <owl:onProperty>
            <owl:ObjectProperty
              rdf:about="http://erlangen-crm.org/current/P84_had_at_most_duration"/>
            </owl:onProperty>
            <owl:maxCardinality
              rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 1
            </owl:maxCardinality>
          </owl:Restriction>
        </rdfs:subClassOf>
        <rdfs:subClassOf>
          <owl:Restriction>
            <owl:maxCardinality
              rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 1
            </owl:maxCardinality>
            <owl:onProperty>
              <owl:ObjectProperty
                rdf:about="http://erlangen-crm.org/current/P83_had_at_least_duration"/>
              </owl:onProperty>
            </owl:Restriction>
          </rdfs:subClassOf>
        </owl:Class>

```



```

<owl:Class rdf:about="http://erlangen-crm.org/current/E53_Place">
  <rdfs:label xml:lang="en">E53 Place</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E1_CRM_Entity"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:maxCardinality
        rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 1
      </owl:maxCardinality>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P59i_is_located_on_or_within"/>
        </owl:onProperty>
      </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E63_Beginning_of_Existence">
  <rdfs:label xml:lang="en">E63 Beginning of Existence</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E5_Event"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:someValuesFrom>
        <owl:Class
          rdf:about="http://erlangen-crm.org/current/E77_Persistent_Item"/>
        </owl:someValuesFrom>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P92_brought_into_existence"/>
        </owl:onProperty>
      </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E64_End_of_Existence">
  <rdfs:label xml:lang="en">E64 End of Existence</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E5_Event"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:someValuesFrom>
        <owl:Class
          rdf:about="http://erlangen-crm.org/current/E77_Persistent_Item"/>

```

```

        </owl:someValuesFrom>
        <owl:onProperty>
            <owl:ObjectProperty
                rdf:about="http://erlangen-crm.org/current/P93_took_out_of_existence"/>
            </owl:onProperty>
        </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E65_Creation">
    <rdfs:label xml:lang="en">E65 Creation</rdfs:label>
    <rdfs:subClassOf
        rdf:resource="http://erlangen-crm.org/current/E63_Beginning_of_Existence"/>
    <rdfs:subClassOf>
        <owl:Class rdf:about="http://erlangen-crm.org/current/E7_Activity"/>
    </rdfs:subClassOf>
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:someValuesFrom>
                <owl:Class
                    rdf:about="http://erlangen-crm.org/current/E28_Conceptual_Object"/>
                </owl:someValuesFrom>
            <owl:onProperty>
                <owl:ObjectProperty
                    rdf:about="http://erlangen-crm.org/current/P94_has_created"/>
                </owl:onProperty>
            </owl:Restriction>
        </rdfs:subClassOf>
    </owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E66_Formation">
    <rdfs:label xml:lang="en">E66 Formation</rdfs:label>
    <rdfs:subClassOf
        rdf:resource="http://erlangen-crm.org/current/E63_Beginning_of_Existence"/>
    <rdfs:subClassOf>
        <owl:Class rdf:about="http://erlangen-crm.org/current/E7_Activity"/>
    </rdfs:subClassOf>
    <rdfs:subClassOf>
        <owl:Class
            rdf:about="http://erlangen-crm.org/current/E63_Beginning_of_Existence"/>
        </rdfs:subClassOf>
    </owl:Class>

```

```

<rdfs:subClassOf>
  <owl:Restriction>
    <owl:someValuesFrom>
      <owl:Class rdf:about="http://erlangen-crm.org/current/E74_Group"/>
    </owl:someValuesFrom>
    <owl:onProperty>
      <owl:ObjectProperty
        rdf:about="http://erlangen-crm.org/current/P95_has_formed"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E67_Birth">
  <rdfs:label xml:lang="en">E67 Birth</rdfs:label>
  <rdfs:subClassOf
    rdf:resource="http://erlangen-crm.org/current/E63_Beginning_of_Existence"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:someValuesFrom
        rdf:resource="http://erlangen-crm.org/current/E21_Person"/>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P96_by_mother"/>
        </owl:onProperty>
      </owl:Restriction>
    </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P97_from_father"/>
        </owl:onProperty>
      <owl:someValuesFrom
        rdf:resource="http://erlangen-crm.org/current/E21_Person"/>
      </owl:Restriction>
    </rdfs:subClassOf>
  </owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E68_Dissolution">
  <rdfs:label xml:lang="en">E68 Dissolution</rdfs:label>

```

```

<rdfs:subClassOf
  rdf:resource="http://erlangen-crm.org/current/E64_End_of_Existence"/>
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:onProperty>
      <owl:ObjectProperty
        rdf:about="http://erlangen-crm.org/current/P99_dissolved"/>
      </owl:onProperty>
      <owl:someValuesFrom
        rdf:resource="http://erlangen-crm.org/current/E74_Group"/>
      </owl:Restriction>
    </rdfs:subClassOf>
  </owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E69_Death">
  <rdfs:label xml:lang="en">E69 Death</rdfs:label>
  <rdfs:subClassOf
    rdf:resource="http://erlangen-crm.org/current/E64_End_of_Existence"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P100_was_death_of"/>
        </owl:onProperty>
        <owl:someValuesFrom
          rdf:resource="http://erlangen-crm.org/current/E21_Person"/>
        </owl:Restriction>
      </rdfs:subClassOf>
    </owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E73_Information_Object">
  <rdfs:label xml:lang="en">E73 Information Object</rdfs:label>
  <rdfs:subClassOf
    rdf:resource="http://erlangen-crm.org/current/E89_Propositional_Object"/>
  <rdfs:subClassOf
    rdf:resource="http://erlangen-crm.org/current/E90_Symbolic_Object"/>
  </owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E74_Group">
  <rdfs:label xml:lang="en">E74 Group</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E39_Actor"/>

```

```
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:onProperty>
      <owl:ObjectProperty
        rdf:about="http://erlangen-crm.org/current/P107_has_current_or_former_member"/>
      </owl:onProperty>
      <owl:minCardinality
        rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 2
      </owl:minCardinality>
    </owl:Restriction>
  </rdfs:subClassOf>
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:minCardinality
      rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 0
    </owl:minCardinality>
    <owl:onProperty>
      <owl:ObjectProperty
        rdf:about="http://erlangen-crm.org/current/P146i_lost_member_by"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:maxCardinality
      rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 1
    </owl:maxCardinality>
    <owl:onProperty>
      <owl:ObjectProperty
        rdf:about="http://erlangen-crm.org/current/P95i_was_formed_by"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:onProperty>
      <owl:ObjectProperty
        rdf:about="http://erlangen-crm.org/current/P144i_gained_member_by"/>
      </owl:onProperty>
      <owl:minCardinality
        rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 2
      </owl:minCardinality>
    </owl:Restriction>
  </rdfs:subClassOf>
```

```

        </owl:minCardinality>
    </owl:Restriction>
</rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E77_Persistent_Item">
    <rdfs:label xml:lang="en">E77 Persistent Item</rdfs:label>
    <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E1_CRM_Entity"/>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E81_Transformation">
    <rdfs:label xml:lang="en">E81 Transformation</rdfs:label>
    <rdfs:subClassOf
        rdf:resource="http://erlangen-crm.org/current/E63_Beginning_of_Existence"/>
    <rdfs:subClassOf
        rdf:resource="http://erlangen-crm.org/current/E64_End_of_Existence"/>
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:someValuesFrom>
                <owl:Class
                    rdf:about="http://erlangen-crm.org/current/E77_Persistent_Item"/>
            </owl:someValuesFrom>
            <owl:onProperty>
                <owl:ObjectProperty
                    rdf:about="http://erlangen-crm.org/current/P123_resulted_in"/>
            </owl:onProperty>
        </owl:Restriction>
    </rdfs:subClassOf>
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:someValuesFrom>
                <owl:Class
                    rdf:about="http://erlangen-crm.org/current/E77_Persistent_Item"/>
            </owl:someValuesFrom>
            <owl:onProperty>
                <owl:ObjectProperty
                    rdf:about="http://erlangen-crm.org/current/P124_transformed"/>
            </owl:onProperty>
        </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>

```

```
<owl:Class rdf:about="http://erlangen-crm.org/current/E85_Joining">
  <rdfs:label xml:lang="en">E85 Joining</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E7_Activity"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P144_joined_with"/>
        </owl:onProperty>
        <owl:minCardinality
          rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 1
        </owl:minCardinality>
      </owl:Restriction>
    </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P143_joined"/>
        </owl:onProperty>
        <owl:cardinality
          rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 1
        </owl:cardinality>
      </owl:Restriction>
    </rdfs:subClassOf>
  </owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E86_Leaving">
  <rdfs:label xml:lang="en">E86 Leaving</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E7_Activity"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P146_separated_from"/>
        </owl:onProperty>
        <owl:minCardinality
          rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 1
        </owl:minCardinality>
      </owl:Restriction>
    </rdfs:subClassOf>
  </owl:Class>
```

```

</rdfs:subClassOf>
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:cardinality
      rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger"> 1
    </owl:cardinality>
    <owl:onProperty>
      <owl:ObjectProperty
        rdf:about="http://erlangen-crm.org/current/P145_separated"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E89 Propositional_Object">
  <rdfs:label xml:lang="en">E89 Propositional Object</rdfs:label>
  <rdfs:subClassOf
    rdf:resource="http://erlangen-crm.org/current/E28_Conceptual_Object"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:someValuesFrom>
        <owl:Class rdf:about="http://erlangen-crm.org/current/E1_CRM_Entity"/>
      </owl:someValuesFrom>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P67i_is_referred_to_by"/>
        </owl:onProperty>
      </owl:Restriction>
    </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:ObjectProperty
          rdf:about="http://erlangen-crm.org/current/P129i_is_subject_of"/>
        </owl:onProperty>
      <owl:someValuesFrom>
        <owl:Class rdf:about="http://erlangen-crm.org/current/E1_CRM_Entity"/>
      </owl:someValuesFrom>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>

```



```

    <owl:Restriction>
      <owl:onProperty>
        <owl:TransitiveProperty
          rdf:about="http://erlangen-crm.org/current/P148_has_component"/>
        </owl:onProperty>
      <owl:someValuesFrom
        rdf:resource="http://erlangen-crm.org/current/E89_Propositional_Object"/>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>

<owl:Class rdf:about="http://erlangen-crm.org/current/E90_Symbolic_Object">
  <rdfs:label xml:lang="en">E90 Symbolic Object</rdfs:label>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:TransitiveProperty
          rdf:about="http://erlangen-crm.org/current/P106_is_composed_of"/>
        </owl:onProperty>
      <owl:someValuesFrom
        rdf:resource="http://erlangen-crm.org/current/E90_Symbolic_Object"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Class rdf:about="http://erlangen-crm.org/current/E72_Legal_Object"/>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Class
      rdf:about="http://erlangen-crm.org/current/E28_Conceptual_Object"/>
  </rdfs:subClassOf>
</owl:Class>

<!-- FRBRoo Classes -->

<owl:Class rdf:about="http://erlangen-crm.org/efrbroo/F23_Expression_Fragment">
  <rdfs:label xml:lang="en">F2 Expression</rdfs:label>
  <rdfs:subClassOf
    rdf:resource="http://erlangen-crm.org/current/E73_Information_Object"/>
  <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/efrbroo/F2_Expression"/>
  <rdfs:subClassOf>
    <owl:Restriction>

```

```

        <owl:someValuesFrom
            rdf:resource="http://erlangen-crm.org/efrbroo/F28_Expression_Creation"/>
        <owl:onProperty>
            <owl:ObjectProperty
                rdf:about="http://erlangen-crm.org/efrbroo/R17i_was_created_by"/>
            </owl:onProperty>
        </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>

<!-- CRMinf Classes -->

<owl:Class rdf:about="http://www.cidoc-crm.org/cidoc-crm/CRMinf/I2_Belief">
    <rdfs:label xml:lang="en">I2 Belief</rdfs:label>
    <rdfs:subClassOf
        rdf:resource="http://erlangen-crm.org/current/E2_Temporal_Entity"/>
</owl:Class>

<owl:Class rdf:about="http://www.cidoc-crm.org/cidoc-crm/CRMinf/I4_Proposition_Set">
    <rdfs:label xml:lang="en">I4 Proposition Set</rdfs:label>
    <rdfs:subClassOf
        rdf:resource="http://erlangen-crm.org/current/E73_Information_Object"/>
</owl:Class>

<owl:Class rdf:about="http://www.cidoc-crm.org/cidoc-crm/CRMinf/I5_Inference_Making">
    <rdfs:label xml:lang="en">I5 Inference Making</rdfs:label>
    <rdfs:subClassOf
        rdf:resource="http://www.cidoc-crm.org/cidoc-crm/CRMinf/I1_Argumentation"/>
</owl:Class>

<!-- CRMsci Classes -->

<owl:Class rdf:about="http://www.ics.forth.gr/isl/CRMsci/S4_Observation">
    <rdfs:label xml:lang="en">S4 Observation</rdfs:label>
    <rdfs:subClassOf
        rdf:resource="http://erlangen-crm.org/current/E13_Attribute_Assignment"/>
</owl:Class>

<owl:Class rdf:about="http://www.ics.forth.gr/isl/CRMsci/S15_Observable_Entity">
    <rdfs:label xml:lang="en">S15 Observable Entity</rdfs:label>
    <rdfs:subClassOf rdf:resource="http://erlangen-crm.org/current/E1_CRM_Entity"/>

```

```
</owl:Class>

<!-- Content in RDF Classes -->

<owl:Class rdf:about="http://www.w3.org/2011/content#ContentAsText">
  <rdfs:label xml:lang="en">Text content</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://www.w3.org/2011/content#Content" />
</owl:Class>

<!-- Dublin Core Classes -->

<owl:Class rdf:about="http://purl.org/dc/dcmitype/Text">
  <rdfs:label xml:lang="en">Text</rdfs:label>
</owl:Class>

<!-- New Classes -->

<owl:Class rdf:about="http://dantesources.org/narratives/Actor_With_Role">
  <rdfs:label xml:lang="en">Actor with Role</rdfs:label>
</owl:Class>

<!-- CIDOC-CRM Properties -->

<owl:ObjectProperty rdf:about="http://erlangen-crm.org/current/P1_is_identified_by">
  <rdfs:label xml:lang="en">P1 is identified by</rdfs:label>
  <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E1_CRM_Entity"/>
  <rdfs:range rdf:resource="http://erlangen-crm.org/current/E41_Appellation"/>
</owl:ObjectProperty>

<owl:DatatypeProperty rdf:about="http://erlangen-crm.org/current/P3_has_note">
  <rdfs:label xml:lang="en">P3 has note</rdfs:label>
  <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E1_CRM_Entity"/>
</owl:DatatypeProperty>

<owl:ObjectProperty rdf:about="http://erlangen-crm.org/current/P4_has_time-span">
  <rdfs:label xml:lang="en">P4 has time-span</rdfs:label>
  <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E2_Temporal_Entity"/>
  <rdfs:range rdf:resource="http://erlangen-crm.org/current/E52_Time-Span"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="http://erlangen-crm.org/current/P7_took_place_at">
```

```

    <rdfs:label xml:lang="en">P7 took place at</rdfs:label>
    <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E4_Period"/>
    <rdfs:range rdf:resource="http://erlangen-crm.org/current/E53_Place"/>
</owl:ObjectProperty>

<owl:ObjectProperty
    rdf:about="http://erlangen-crm.org/current/P8_took_place_on_or_within">
    <rdfs:label xml:lang="en">P8 took place on or within</rdfs:label>
    <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E4_Period"/>
    <rdfs:range rdf:resource="http://erlangen-crm.org/current/E18_Physical_Thing"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="http://erlangen-crm.org/current/P9_consists_of">
    <rdfs:label xml:lang="en">P9 consists of</rdfs:label>
    <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E4_Period"/>
    <rdfs:range rdf:resource="http://erlangen-crm.org/current/E4_Period"/>
    <rdfs:type rdf:resource="http://www.w3.org/2002/07/owl#TransitiveProperty"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="http://erlangen-crm.org/current/P11_had_participant">
    <rdfs:label xml:lang="en">P11 had participant</rdfs:label>
    <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E5_Event"/>
    <rdfs:range rdf:resource="http://erlangen-crm.org/current/E39_Actor"/>
    <rdfs:subPropertyOf
        rdf:resource="http://erlangen-crm.org/current/P12_occurred_in_the_presence_of"/>
</owl:ObjectProperty>

<owl:ObjectProperty
    rdf:about="http://erlangen-crm.org/current/P12_occurred_in_the_presence_of">
    <rdfs:label xml:lang="en">P12 occurred in the presence of</rdfs:label>
    <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E5_Event"/>
    <rdfs:range rdf:resource="http://erlangen-crm.org/current/E77_Persistent_Item"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="http://erlangen-crm.org/current/P14_carried_out_by">
    <rdfs:label xml:lang="en">P14 carried out by</rdfs:label>
    <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E7_Activity"/>
    <rdfs:range rdf:resource="http://erlangen-crm.org/current/E39_Actor"/>
    <rdfs:subPropertyOf
        rdf:resource="http://erlangen-crm.org/current/P11_had_participant"/>
</owl:ObjectProperty>

```

```
<owl:DatatypeProperty
  rdf:about="http://erlangen-crm.org/current/P79_beginning_is_qualified_by">
  <rdfs:label xml:lang="en">P79 beginning is qualified by</rdfs:label>
  <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E52_Time-Span"/>
  <rdfs:subPropertyOf rdf:resource="http://erlangen-crm.org/current/P3_has_note"/>
</owl:DatatypeProperty>

<owl:DatatypeProperty
  rdf:about="http://erlangen-crm.org/current/P80_end_is_qualified_by">
  <rdfs:label xml:lang="en">P80 end is qualified by</rdfs:label>
  <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E52_Time-Span"/>
  <rdfs:subPropertyOf resource="http://erlangen-crm.org/current/P3_has_note"/>
</owl:DatatypeProperty>

<owl:TransitiveProperty
  rdf:about="http://erlangen-crm.org/current/P116_is_composed_of">
  <rdfs:label xml:lang="en">P116 is composed of</rdfs:label>
  <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E90_Symbolic_Object"/>
  <rdfs:range rdf:resource="http://erlangen-crm.org/current/E90_Symbolic_Object"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
</owl:TransitiveProperty>

<owl:TransitiveProperty
  rdf:about="http://erlangen-crm.org/current/P117_occurs_during">
  <rdfs:label xml:lang="en">P117 occurs during</rdfs:label>
  <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E2_Temporal_Entity"/>
  <rdfs:range rdf:resource="http://erlangen-crm.org/current/E2_Temporal_Entity"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
</owl:TransitiveProperty>

<owl:ObjectProperty
  rdf:about="http://erlangen-crm.org/current/P118_overlaps_in_time_with">
  <rdfs:label xml:lang="en">P118 overlaps in time with</rdfs:label>
  <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E2_Temporal_Entity"/>
  <rdfs:range rdf:resource="http://erlangen-crm.org/current/E2_Temporal_Entity"/>
</owl:ObjectProperty>

<owl:ObjectProperty
  rdf:about="http://erlangen-crm.org/current/P119_meets_in_time_with">
  <rdfs:label xml:lang="en">P119 meets in time with</rdfs:label>
```

```

    <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E2_Temporal_Entity"/>
    <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E2_Temporal_Entity"/>
</owl:ObjectProperty>

<owl:TransitiveProperty
    rdf:about="http://erlangen-crm.org/current/P120_occurs_before">
    <rdfs:label xml:lang="en">P120 occurs before</rdfs:label>
    <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E2_Temporal_Entity"/>
    <rdfs:range rdf:resource="http://erlangen-crm.org/current/E2_Temporal_Entity"/>
    <rdfs:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
</owl:TransitiveProperty>

<owl:ObjectProperty rdf:about="http://erlangen-crm.org/current/P129_is_about">
    <rdfs:label xml:lang="en">P129 is about</rdfs:label>
    <rdfs:domain
        rdf:resource="http://erlangen-crm.org/current/E89_Propositional_Object"/>
    <rdfs:range rdf:resource="http://erlangen-crm.org/current/E1_CRM_Entity"/>
    <rdfs:subPropertyOf rdf:resource="http://erlangen-crm.org/current/P67_refers_to"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="http://erlangen-crm.org/current/P138_represents">
    <rdfs:label xml:lang="en">P138 represents</rdfs:label>
    <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E36_Visual_Item"/>
    <rdfs:range rdf:resource="http://erlangen-crm.org/current/E1_CRM_Entity"/>
    <rdfs:subPropertyOf rdf:resource="http://erlangen-crm.org/current/P67_refers_to"/>
</owl:ObjectProperty>

<!-- CRMinf Properties -->

<owl:ObjectProperty
    rdf:about="http://www.cidoc-crm.org/cidoc-crm/CRMinf/J1_used_as_premise">
    <rdfs:label xml:lang="en">J1 used as premise</rdfs:label>
    <rdfs:domain
        rdf:resource="http://www.cidoc-crm.org/cidoc-crm/CRMinf/I5_Inference_Making"/>
    <rdfs:range rdf:resource="http://www.cidoc-crm.org/cidoc-crm/CRMinf/I2_Belief"/>
    <rdfs:subPropertyOf
        rdf:resource="http://erlangen-crm.org/current/P17_was_motivated_by"/>
</owl:ObjectProperty>

<owl:ObjectProperty
    rdf:about="http://www.cidoc-crm.org/cidoc-crm/CRMinf/J2_concluded_that">

```

```
<rdfs:label xml:lang="en">J2 concluded that</rdfs:label>
<rdfs:domain
  rdf:resource="http://www.cidoc-crm.org/cidoc-crm/CRMinf/I1_Argumentation"/>
<rdfs:range rdf:resource="http://www.cidoc-crm.org/cidoc-crm/CRMinf/I2_Belief"/>
<rdfs:subPropertyOf rdf:resource="http://erlangen-crm.org/current/P116_starts"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="http://www.cidoc-crm.org/cidoc-crm/CRMinf/J4_that">
  <rdfs:label xml:lang="en">J4 that</rdfs:label>
  <rdfs:domain rdf:resource="http://www.cidoc-crm.org/cidoc-crm/CRMinf/I2_Belief"/>
  <rdfs:range
    rdf:resource="http://www.cidoc-crm.org/cidoc-crm/CRMinf/I4_Proposition_Set"/>
  <rdfs:subPropertyOf rdf:resource="http://erlangen-crm.org/current/P116_starts"/>
</owl:ObjectProperty>

<!-- Content in RDF Properties -->

<owl:ObjectProperty rdf:about="http://www.w3.org/2011/content#chars">
  <rdfs:label xml:lang="en">Character sequence</rdfs:label>
  <rdfs:domain rdf:resource="http://www.w3.org/2011/content#ContentAsText" />
  <rdfs:range rdf:resource="http://www.w3.org/2000/01/rdf-schema#Literal" />
</owl:ObjectProperty>

<!-- Dublin Core Properties -->

<owl:DatatypeProperty rdf:about="http://purl.org/dc/elements/1.1/format">
  <rdfs:label xml:lang="en">Format</rdfs:label>
  <rdfs:range rdf:resource="http://purl.org/dc/terms/MediaTypeOrExtent"/>
</owl:DatatypeProperty>

<!-- New Properties -->

<owl:ObjectProperty rdf:about="http://dantesources.org/narratives/had_participant">
  <rdfs:label xml:lang="en">J4 that</rdfs:label>
  <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E5_Event"/>
  <rdfs:range rdf:resource="http://dantesources.org/narratives/Actor_With_Role"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="http://dantesources.org/narratives/subject">
  <rdfs:label xml:lang="en">subject</rdfs:label>
  <rdfs:domain rdf:resource="http://dantesources.org/narratives/Actor_With_Role"/>
```

```

    <rdfs:range rdf:resource="http://erlangen-crm.org/current/E39_Actor"/>
</owl:ObjectProperty>

<owl:DatatypeProperty rdf:about="http://dantesources.org/narratives/role">
  <rdfs:label xml:lang="en">role</rdfs:label>
  <rdfs:domain rdf:resource="http://dantesources.org/narratives/Actor_With_Role"/>
</owl:DatatypeProperty>

<owl:ObjectProperty rdf:about="http://dantesources.org/narratives/has_source">
  <rdfs:label xml:lang="en">has source</rdfs:label>
  <rdfs:domain
    rdf:resource="http://erlangen-crm.org/current/E73_Information_Object"/>
  <rdfs:range
    rdf:resource="http://erlangen-crm.org/current/E73_Information_Object"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="http://dantesources.org/narratives/has_text_fragment">
  <rdfs:label xml:lang="en">has text fragment</rdfs:label>
  <rdfs:domain
    rdf:resource="http://erlangen-crm.org/current/E73_Information_Object"/>
  <rdfs:range
    rdf:resource="http://erlangen-crm.org/efrbroo/F23_Expression_Fragment"/>
</owl:ObjectProperty>

<owl:DatatypeProperty rdf:about="http://dantesources.org/narratives/has_reference">
  <rdfs:label xml:lang="en">has reference</rdfs:label>
  <rdfs:domain
    rdf:resource="http://erlangen-crm.org/current/E73_Information_Object"/>
</owl:DatatypeProperty>

<owl:TransitiveProperty
  rdf:about="http://dantesources.org/narratives/causally_depends_on">
  <rdfs:label xml:lang="en">causally depends on</rdfs:label>
  <rdfs:domain rdf:resource="http://erlangen-crm.org/current/E5_Event"/>
  <rdfs:range rdf:resource="http://erlangen-crm.org/efrbroo/E5_Event"/>
  <rdfs:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
</owl:TransitiveProperty>

</rdf:RDF>

```



# Bibliography

- Agosti, Maristella et al. (2013). “Enhancing End User Access to Cultural Heritage Systems: Tailored Narratives and Human-Centered Computing.” In: *New Trends in Image Analysis and Processing—ICIAP 2013*. Springer, pp. 278–287.
- Alani, Harith et al. (2003). “Automatic ontology-based knowledge extraction from web documents.” In: *Intelligent Systems, IEEE* 18.1, pp. 14–21.
- Allen, James F (1984). “Towards a general theory of action and time.” In: *Artificial intelligence* 23.2, pp. 123–154.
- Antoniou, Grigoris and Martin Doerr (2007). “Web Ontology Languages.” In: *Semantic Web Services: Theory, Tools and Applications: Theory, Tools and Applications*. IGI Global.
- Auer, Sören et al. (2007). “Dbpedia: A nucleus for a web of open data.” In: *The semantic web*. Springer, pp. 722–735.
- Bal, Mieke (1997). *Narratology: Introduction to the theory of narrative*. University of Toronto Press.
- Bartalesi, Valentina, Carlo Meghini, and Daniele Metilli (2016). “Steps Towards a Formal Ontology of Narratives Based on Narratology.” In: *Proceedings of the International Workshop on Computational Models of Narrative 2016*.

- Barthes, Roland (1966). “Introduction à l’analyse structurale des récits.” In: *Communications* 8.1, pp. 1–27.
- Berners-Lee, Tim, James Hendler, Ora Lassila, et al. (2001). “The semantic web.” In: *Scientific american* 284.5, pp. 28–37.
- Cavazza, Marc and David Pizzi (2006). “Narratology for interactive storytelling: A critical introduction.” In: *Technologies for Interactive Digital Storytelling and Entertainment*. Springer, pp. 72–83.
- Committee, PREMIS Editorial et al. (2008). “PREMIS data dictionary for preservation metadata, version 2.0.” In: *Retrieved May 22*, p. 2010.
- Crawford, Chris (2012). *Chris Crawford on interactive storytelling*. New Riders.
- De Saussure, Ferdinand (1989). *Cours de linguistique générale: Édition critique*. Vol. 1. Otto Harrassowitz Verlag.
- Doerr, Martin (2003). “The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata.” In: *AI Mag.* 24.3, pp. 75–92. ISSN: 0738-4602.
- Doerr, Martin et al. (2008). “FRBRoo, a conceptual model for performing arts.” In: *2008 Annual Conference of CIDOC*, pp. 06–18.
- Elson, David K and Kathleen R McKeown (2009). “A tool for deep semantic encoding of narrative texts.” In: *Proceedings of the ACL-IJCNLP 2009 Software Demonstrations*. Association for Computational Linguistics, pp. 9–12.
- Erling, Orri and Ivan Mikhailov (2009). “RDF Support in the Virtuoso DBMS.” In: *Networked Knowledge-Networked Media*. Springer, pp. 7–24.

- Erxleben, Fredo et al. (2014). “Introducing Wikidata to the linked data web.” In: *International Semantic Web Conference*. Springer, pp. 50–65.
- Färber, Michael et al. (2015). “A Comparative Survey of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO.” In: *Semantic Web Journal*, July.
- Fernández-Aceves, Hervin (2013). “A Relational View of the Norman Kingdom of Sicily and its Royal Court: The Social Space Constructed by ‘Hugo Falcandus’” MA thesis. Budapest, Hungary: Central European University.
- Finlayson, Mark A (2011). “The Story Workbench: An Extensible Semi-Automatic Text Annotation Tool.” In: *Intelligent Narrative Technologies*.
- Franzosi, Roberto (1998). “Narrative as data: Linguistic and statistical tools for the quantitative study of historical events.” In: *International review of social history* 43.S6, pp. 81–104.
- Frixione, Marcello and Antonio Lieto (2012). “Representing concepts in formal ontologies. Compositionality vs. typicality effects.” In: *Logic and Logical Philosophy* 21.4, pp. 391–414.
- Genette, Gerard E and Jane E Lewin (1983). *Narrative discourse: An essay in method*. Cornell University Press.
- Georgakopoulou, Alexandra (2007). *Small stories, interaction and identities*. Vol. 8. John Benjamins Publishing.
- Görz, Günther, Martin Oischinger, and Bernhard Schiemann (2008). “An implementation of the CIDOC conceptual reference model (4.2.4) in OWL-DL.” In: *Proceedings of the 2008 Annual Conference of CIDOC-The Digital Curation of Cultural Heritage*.

- Greimas, Algirdas Julien, Daniele McDowell, and Alan R Velie (1983). *Structural semantics: An attempt at a method*. University of Nebraska Press Lincoln.
- Gruber, Tom (1993). “What is an Ontology.” In: *WWW Site <http://www-ksl.stanford.edu/kst/whatis-an-ontology.html>* (accessed on 07-09-2004).
- Guarino, Nicola (1998). “Formal ontology and information systems.” In: *Proceedings of FOIS*. Vol. 98. 1998, pp. 81–97.
- Halliwell, Stephen (1987). *The Poetics of Aristotle: translation and commentary*. UNC Press Books.
- Harrell Jr, Douglas Alan (2007). “Theory and technology for computational narrative: An approach to generative and interactive narrative with bases in algebraic semiotics and cognitive linguistics.” PhD thesis. University of California, San Diego.
- Herman, David (2000). “Narratology as a cognitive science.” In: *image and narrative* 1.1.
- (2011). *Basic elements of narrative*. John Wiley & Sons.
- Hernández, Daniel, Aidan Hogan, and Markus Krötzsch (2015). “Reifying RDF: What Works Well With Wikidata.” In: *Proceedings of the 11th International Workshop on Scalable Semantic Web Knowledge Base Systems*. Vol. 1457, pp. 32–47.
- Hirsch, Christian, John Hosking, and John Grundy (2009). “Interactive visualization tools for exploring the semantic graph of large knowledge spaces.” In: *Workshop on Visual Interfaces to the Social and the Semantic Web (VISSW2009)*. Vol. 443.
- Hyvärinen, Matti (2006). “Towards a conceptual history of narrative.” In:

- Kenny, A. (2013). *Poetics*. Oxford World's Classics. OUP Oxford. ISBN: 9780191635809.
- Kreiwirth, Martin (2000). "Merely telling stories? Narrative and knowledge in the human sciences." In: *Poetics Today* 21.2, pp. 293–318.
- Labov, William and Joshua Waletzky (1997). "Narrative analysis: Oral versions of personal experience." In:
- Lehmann, Jens et al. (2015). "DBpedia—a large-scale, multilingual knowledge base extracted from Wikipedia." In: *Semantic Web* 6.2, pp. 167–195.
- Levi-Strauss, Claude (1963). "Structural analysis in linguistics and in anthropology." In: *Semiotics-An Introductory Anthology*, pp. 110–128.
- Lieto, Antonio and Rossana Damiano (2014). "A Hybrid Representational Proposal for Narrative Concepts: A Case Study on Character Roles." In: *OASICS-OpenAccess Series in Informatics*. Vol. 41. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.
- Lombardo, Vincenzo and Rossana Damiano (2010). "Narrative annotation and editing of video." In: *Interactive Storytelling*. Springer, pp. 62–73.
- Mani, Inderjeet (2012). "Computational Modeling of Narrative." In: *Synthesis Lectures on Human Language Technologies* 5.3, pp. 1–142.
- (2014). "Computational narratology." In: *Handbook of narratology*, pp. 84–92.
- Manola, Frank and Eric Miller (2004). *RDF Primer*. W3C Recommendation. <http://www.w3.org/TR/rdf-primer/>. WWW Consortium.

- Meister, Jan Christoph (2012). *Narratology*. Synthesis. American Museum of Natural History. Available at <http://ncep.amnh.org>.
- Mietchen, Daniel et al. (2015). “Enabling Open Science: Wikidata for Research (Wiki4R).” In: *Research Ideas and Outcomes* 1, e7573.
- Nadeau, David and Satoshi Sekine (2007). “A survey of named entity recognition and classification.” In: *Linguisticae Investigationes* 30.1, pp. 3–26.
- Niehaus, James et al. (2014). “A Flexible Framework for the Creation of Narrative-Centered Tools.” In: *OASIS-OpenAccess Series in Informatics*. Vol. 41. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.
- Ochs, Elinor and Lisa Capps (2009). *Living narrative: Creating lives in everyday storytelling*. Harvard University Press.
- Pellissier Tanon, Thomas et al. (2016). “From Freebase to Wikidata: The Great Migration.” In: *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, pp. 1419–1428.
- Propp, Vladimir (1973). *Morphology of the Folktale*. Vol. 9. University of Texas Press.
- Ryan, Marie-Laure et al. (2007). “Toward a definition of narrative.” In: *The Cambridge companion to narrative*, pp. 22–35.
- Shaw, Ryan, Raphael Troncy, and Lynda Hardman (2009). “Lode: Linking open descriptions of events.” In: *The Semantic Web*. Springer, pp. 153–167.
- Shklovsky, Victor (1965). “Art as technique.” In: *Russian formalist criticism: Four essays* 3.
- Sikos, L. (2015). *Mastering Structured Data on the Semantic Web: From HTML5 Microdata to Linked Open Data*. Apress. ISBN: 9781484210499.

- Song, Dezhao, Christopher G Chute, and Cui Tao (2012). “Semantator: annotating clinical narratives with semantic web ontologies.” In: *AMIA Summits on Translational Science Proceedings 2012*, p. 20.
- Spitz, Andreas et al. (2016). “State of the Union: A Data Consumer’s Perspective on Wikidata and Its Properties for the Classification and Resolution of Entities.” In: *Tenth International AAAI Conference on Web and Social Media*.
- Todorov, Tzvetan (1969). *Grammaire du Décaméron*. Mouton The Hague.
- Uhlir, Jan and Martin Falc (2004). “Annotating narratives using ontologies and conceptual graphs.” In: *Database and Expert Systems Applications, 2004. Proceedings. 15th International Workshop on*. IEEE, pp. 84–88.
- Van Hage, Willem Robert et al. (2011). “Design and use of the Simple Event Model (SEM).” In: *Web Semantics: Science, Services and Agents on the World Wide Web 9.2*, pp. 128–136.
- Vrandečić, Denny and Markus Krötzsch (2014). “Wikidata: a free collaborative knowledgebase.” In: *Communications of the ACM 57.10*, pp. 78–85.
- Wolff, Annika, Paul Mulholland, and Trevor Collins (2012). “Storyspace: a story-driven approach for creating museum narratives.” In: *Proceedings of the 23rd ACM conference on Hypertext and social media*. ACM, pp. 89–98.
- Zarri, Gian Piero (2009). *Representation and management of narrative information: Theoretical principles and implementation*. Springer Science & Business Media.