A conceptualisation of narratives and its expression in the CRM

Valentina Bartalesi*, Carlo Meghini and Daniele Metilli

Istituto di Scienza e Tecnologie dell'Informazione
"Alessandro Faedo" (ISTI) – CNR,
Via G. Moruzzi 1, 56127 Pisa, Italy
Email: valentina.bartalesi@isti.cnr.it
Email: carlo.meghini@isti.cnr.it
Email: daniele.metilli@isti.cnr.it
*Corresponding author

Abstract: Current Digital Libraries (DLs) are mostly built around large collections of scarcely related objects. We aim at enriching the information space of DLs by introducing *narratives*, consisting of two main components: networks of events related to one another and to the DL resources through semantic links, and narrations of those events in texts. In order to introduce narratives in DLs, we developed a conceptualisation based on narratology and we expressed it using the CIDOC CRM and CRMinf as reference ontologies. We used this expression to validate our conceptualisation, creating a narrative of the biography of Dante Alighieri as a realistic case study. To support this experiment, we developed a semi-automated tool that collects basic knowledge about objects and events from Wikidata. The developed ontology is generally enough to be not limited to create biographies but other types of narratives as well.

Keywords: ontology; semantic web; digital libraries; narratives; CIDOC CRM; OWL.

Reference to this paper should be made as follows: Bartalesi, V., Meghini, C. and Metilli, D. (XXXX) 'A Conceptualisation of Narratives and Its Expression in the CRM', *Int. J. Metadata, Semantics and Ontologies*, Vol. X, No. Y, pp.xxx–xxx.

Biographical notes: Valentina Bartalesi received her PhD in Information Engineering at the University of Pisa. She has a master's degree in Digital Humanities, University of Pisa. She worked seven years at CELCT, the Italian Centre for the Evaluation of Language and Communication Technologies. Since 2011, she is research fellow at the Institute of Science and Technologies of CNR in Pisa and currently works in the Networked Multimedia Information Systems laboratory. Her research interests are mainly related to the Semantic Web technologies, in particular to the development of formal ontologies to represent textual content to be accessed through Web applications. She is co-author of 28 research papers at national and international conferences and journals.

Carlo Meghini is prime researcher at ISTI-CNR in the area of Conceptual Modelling, Digital Libraries and Digital Preservation. He graduated in Computer Science at the University of Pisa, Italy, in 1979 with a research thesis on distributed databases. He has participated in several EC-funded research actions in the areas of information systems, digital libraries and digital preservation. He is involved in European projects since 1988, in the areas of Digital Libraries (DELOS Network of Excellence, BRICKS, EDLNet, Europeana version 1.0, Europeana version 2.0, ASSETS) and Digital Preservation (CASPAR). From 2007 he is involved in the making of the European digital library (Europeana, www.europeana.eu), taking care of the scientific aspects of the project. He has published about 100 scientific papers in international journals, books and conferences.

Daniele Metilli is a research collaborator at the NeMIS laboratory of ISTI-CNR, working on the representation of narratives through Semantic Web technologies. He holds a master's degree in Digital Humanities from the University of Pisa and a bachelor's degree in Computer Engineering from the Polytechnic University of Milan. His current research interests revolve around the Semantic Web, Digital Libraries, and online collaborative projects. He is an administrator of the English Wikipedia and of the Wikidata knowledge base. Previously, he has been the first Wikipedian in residence at an Italian museum.

1 Introduction

Digital Libraries (DLs) are information systems that offer services over large sets of digital objects (Meghini et al., 2014). DLs were officially born about two decades ago from the Digital Library Initiative (http://www.dlib.org/dlib/july98/07griffin.html) that brought DLs in the focus of research. Much progress has been achieved since, mainly due to major breakthroughs in relevant technological fields, such as communications and multimedia. Thanks to these technological progresses, today's DLs have a wide penetration and cover a wide spectrum of digital object types, ranging from hypermedia to 3D models. However, the basic information service of a contemporary DL is essentially the same as that of a traditional library: to support users in *discovering* the digital objects that satisfy an information need, typically expressed as a query consisting of a short list of terms. Discovery works reasonably well on the web, which may be viewed as a very large DL whose objects are pages rich in textual content and interlinked between each other. On the contrary, the traditional search functionalities of DLs, such as Europeana,[1] respond to a web-like query with a ranked list of digital objects based only on the metadata descriptors that are semantically poor. Let us consider a young user wishing to know more about Dante Alighieri, the major Italian poet of the late Middle Ages. She types 'Dante Alighieri' into her favourite web search engine and most likely she gets a ranking with the Wikipedia page about Dante on top. Not willing to spend hours reading, she tries other websites, where she hopes to find something quicker and more exciting to consume than text. At some point, she lands on the web page of a DL, say Europeana, where she tries again her query. To her disappointment, the result is a long list of disparate objects, each offering a glimpse of Dante's life and works, but altogether incapable of providing an idea of who Dante was and what he did. We believe that DLs should offer something more than a ranked list of objects to information seekers. In particular, we believe that DLs should be able to provide *narratives* to their users, in the sense of networks of *events* related to one another and to the DL resources through semantic links. Indeed, the ultimate goal of our study is to promote narratives as first-class objects for DLs. Thanks to this introduction, DLs will be able to semantically connect objects to events, and through events to other objects, resulting in a much richer information structure. Such information structure will allow DLs to provide more sophisticated information services to their users, going beyond the current state. As a first step towards the introduction of narratives in digital libraries, this paper presents a conceptualisation of narratives based on notions found in narratology and in the Artificial Intelligence (AI) literature (in particular in the Event Calculus theory), its expression in a general ontology, the CIDOC CRM, and the use of this expression for validating the conceptualisation on a realistic case study: the biography of Dante Alighieri. To support this validation, we developed a semi-automated tool that collects basic knowledge about objects and events from the Wikidata knowledge base and supports the user in building narratives on top of this basic knowledge. By means of a semi-automatic tool to build narratives, we constructed a narrative about the life of Dante Alighieri (https://dlnarratives.eu/narratives.html). This narrative, visualised in form of timeline, could be imported in a DL (e.g. Europeana), and shown as the result of a search in the DL, thereby placing the digital objects related to Dante in the more general context of the biography (or a part thereof) of the poet.

The conceptualisation, its expression in the CRM and the semi-automated tool we developed are by no means limited to biographies, but they can also be applied to different narrative domains. However, in this paper we validate the developed ontology through the representation of Dante's life as case study.

In this paper, we do not detail the concrete application of the resulting ontology to DLs, which will be considered in a future work.

The paper is structured as follows: Section 2 introduces the theory of narratology and reports the formal logic definitions of the components of narratives. Section 3 describes several ontologies representing events. Section 4 reports the users' requirements and gives a conceptualisation of the narrative structure and its expression using the CIDOC CRM and CRMinf as reference vocabularies. Section 5 presents the experimental validation of the conceptualisation for narratives. Section 6 introduces the tool we developed to construct and visualise narratives. Section 7 reports a first qualitative evaluation of the ontology. Finally, Section 8 contains our conclusive remarks.

2 A brief overview of narratives

In this section, we briefly review some fundamental notions on the narrative structure from the theory of narratology and on the components of narratives from the Event Calculus, along with some definitions from the area of AI.

2.1 Narratology

In literary theory, narratology is the study of narrative structure derived from literary criticism. People conventionally refer to an occurrence taking place at a certain time at a specific location as an *event*. The concept of event is a core element of narratives and generally speaking of the theory of narratology. This research field studies the logic, principles, and practices of narrative representation (Meister, 2003). An antecedent to the modern science of narratology could be found in Aristotle's Poetics (Kenny, 2013). Aristotle considers narratology as being an imitation of a real action (*praxis*) that constitutes an argument (*logos*), providing the basis of the plot (*mythos*). The plot is formed by the events selected and ordered. Despite its antecedents in classical theories of aesthetics, the theoretical principles of narratology derive from linguistic-centred approaches to literature defined by Russian formalists (in the earlier years of 20th century) and later developed through European structuralism. Russian formalism distinguishes between the series of events, that is actions or occurrences taking place at a certain time at a specific location, that compose the story (*fabula*) and the particular way that story gets narrated (*syuzhet*). *Fabula* refers to the sequence of events of the narrative in chronological order; *syuzhet* (or plot) is the way in which these events are presented in a narrative (Kenny, 2013; Shklovsky, 1965). Bal (1997) defines a third level, called *presentation*, that constitutes the concrete representation of the content that is conveyed to the audience (e.g. the text in a novel). The theory of narratology was further developed by mid-20th-century structuralism (Greimas et al., 1983; Levi-Strauss, 1963; Todorov, 1969) and since the 1980s by post-structuralist perspectives. In particular, Cognitive Narratology considers narratology as a psychological phenomenon, and proposes a study of narrative aspects from a cognitive perspective (Herman, 2000). Ontologies based on this approach have also been developed (Lombardo et al., 2015). Some studies have shown that many common-sense concepts are difficult to characterise in terms of necessary/sufficient conditions without resorting to prototypical knowledge, which cannot be adequately represented by commonly used monotonic description logics (Frixione and Lieto, 2012; Lieto and Damiano, 2014). However, for the purposes of our current research, we are not interested in representing the full scope of a cognitive approach to narratives, focusing instead on the basic definition of narrative structure as provided by Russian formalism.

2.2 Artificial intelligence

The Event Calculus (EC) (Kowalski and Sergot, 1989; Miller and Shanahan, 2002; Mueller, 2014) is a logic language for representing actions that have duration and can overlap with each other. Unlike the Situation Calculus (McCarthy, 1961; McCarthy and Hayes, 1969), which represents and reasons about situations resulting from actions performed in the world, EC is based on time rather than situations. The basic concepts of EC useful for our representation of narratives are reported below.

- *Fluents*. Fluents are functions and predicates that vary over time, used to describe the effects of actions. Fluents in EC are time-dependent and EC axioms define a fluent true at a point in time if 'the fluent was initiated by an event at some time in the past and was not terminated by an intervening event' (Russell et al., 2003, p.446).

- *Events or Actions*. In EC the terms Actions and Events are interchangeable and represent changes performed over time. On the other hand, Davidson's theory (Davidson, 2001) defines actions as a particular subclass of events, that is the events endowed with intentionality.

- *Generalised events*. In a physical universe with a spatio-temporal dimension, a generalised event is a space-time chunk which generalises concepts like actions, locations, times, fluents and physical objects such as things, animals, agents, humans.

- *Narrative*. As reported in Van Harmelen et al. (2008), a narrative is a possibly incomplete specification of a set of actual event occurrences (Miller and Shanahan, 1994; Sandewall, 1989). The Event Calculus is narrative-based, unlike the Situation Calculus in which an exact sequence of hypothetical actions is represented.

3 Related works

In this section, we describe several ontologies, which include events or play event roles in more detail. Ontologies can be divided into following categories: upper ontologies and domain ontologies (Poli et al., 2010). An upper ontology is a domain-independent ontology, from which more domain-specific ontologies may be derived. A domain ontology specifies concepts, which belong to a specific domain of interest.

Various models have been developed for representing events on the Semantic Web (Astrova et al., 2014), in the following we briefly describe some of them.

OpenCyc (http://opencyc.org) is the open source version of the Cyc ontology. It is an upper ontology (http://www.cyc.com), which is used for representing human knowledge about the objects and events of everyday life. The Cyc knowledge base contains about 500,000 terms, including about 15,000 types of relations, and about five million assertions relating these terms. OpenCyc distinguishes between static situations and events. The first ones are situations that are extended in time but do not change, whereas events are situations that are extended and change in time.

The *Suggested Upper Merged Ontology (SUMO)* (Pease et al., 2002) is a comprehensive upper ontology which is fully mapped to the WordNet lexical database (Niles and Pease, 2003). The ontology has been successfully applied to the representation of narratives, in particular to automated story generation systems (Cua et al., 2010), and it is has also been used to model the cause-effect relations found in narratives (Ang et al., 2011).

DOLCE is the first module of the WonderWeb Foundational Ontologies Library (Masolo et al., 2003). DOLCE aims at representing the ontological categories underlying natural language and human common sense. The current implementation of DOLCE is DOLCE+.

The *CIDOC CRM* (http://www.cidoc-crm.org) (CRM for short) is a high-level ontology that allows to integrate the information contained in data of the cultural heritage domain along with their correlation with knowledge stored in libraries and archives (Doerr, 2003). The CRM is one of the most widely adopted ontologies in the domain of Cultural Heritage, where both digital libraries and narratives belong. CRM is also an ISO standard since 2006 (ISO21127:2006) and renewed 2014 (ISO21127:2014). Both these factors are crucial to attain semantic interoperability, based on sharing existing ontologies.

The *Europeana Data Model* (EDM) (Doerr et al., 2010) is a model that aims at structuring and representing data delivered to Europeana (http://www.europeana.eu) by the various contributing cultural heritage institutions. In the

EDM Primer (Isaac, 2013) two approaches to provide contextual information about objects are reported: object-centric and event-centric. With the former, descriptive metadata, such as for example title or creator, are attached to the provided object. With the event-centric approach, relations between different entities are described by means of Events, and metadata are attached to such events.

The *Event Ontology* (Raimond and Abdallah, 2007) was developed in the Centre for Digital Music of the Queen Mary University of London. This ontology, which can be used in conjunction with other music-related ontologies, has not specific terms related to the music domain so it can be used in other domains as well. The top-level class in the Event Ontology is the class Event.

The *Linking Open Descriptions of Events* (LODE) ontology is a 'minimal model that encapsulates the most useful properties for describing events' (Shaw et al., 2009, p.160). The aim of LODE is to permit interoperability when modelling the factual aspects of events. Those aspects are characterised in the four aspects: what happened, where did it happen, when did it happen, and who was involved.

The *ABC ontology* is a basic model and an ontology to facilitate the development of a domain, role, or community specific ontologies, in particular 'it is a basic Ontology, which provides a basic model for domain-related or community-related development' (Wenjun et al., 2005, p.303). The ABC ontology was developed for modelling physical, digital and analogue objects contained in libraries, archives, and museums and on the internet (Hunter, 2003).

The *Simple Event Model (SEM)* (Van Hage et al., 2011) allows representing events in different domains, independently from the domain-specific vocabularies that can be used. An event in SEM is defined as everything that happens, even if fictional. SEM is developed with a minimum of semantic commitment to maximise its interoperability. The core classes of SEM are Event, Actor, Place and Time. These represent the main aspects of an event: what happens, who or what participated, where and when did it happen.

The *Drammar Ontology* is a semantic model for the representation of drama features, featuring an SRWL-based rule layer to provide automatic reasoning (Lombardo et al., 2015).

4 A simple ontology for narratives

In this section, we provide a conceptualisation of the narrative structure. The conceptualisation is subsequently expressed using the CIDOC CRM (Doerr, 2003) and CRMinf (Doerr et al., 2015) as reference ontologies.

4.1 Ontology users' requirements

As specified in the ISO 9241-210:2010 standard, understanding the needs and requirements of the users is the first step to develop successful systems and products. Indeed, the result of this analysis can bring a project an increase of productivity, a better quality of the work, smaller costs for providing support and training, and improvement of users' satisfaction (Doll and Torkzadeh, 1988).

The first step in users' requirements analysis is to collect background information about the users and the processes that currently take place, through structured interviews. Using this approach, an initial set of requirements was developed, in order to create an ontology for representing the knowledge on narratives. The preferred users for this research are scholars who want to create and access narratives about the life and the works of the authors they study. Based on the results of the data collected by interviewing three scholars of the University of Pisa, expert in Dante Alighieri's biography, and on the study of the literature, principal users' requirements were identified. These are reported below:

- Representing the events that compose the narrative in a chronological order.
- Reconstructing the plot of the narrative, i.e. as the events are narrated.
- Representing the provenance, i.e. the inferential process of the narrator who composes a narrative from primary sources whose contents are propositions about the events of the narration.

4.2 Conceptualisation

This section introduces a formal computable representation of narratives in an informal way, based on the narratology background reported in Subsection 2.1 and on the basic notions of Event Calculus as introduced in Subsection 2.2. In particular, the idea is that a *narrative* consists of three main elements:

1 The *fabula*, directly representing the fabula as defined by Russian formalism, i.e. the sequence of the events that composes the narrative, each with its own features, in chronological order.

2 One or more texts that narrate the fabula, called *narrations* and that correspond to Bal's definition of *presentation*.

3 A *reference* function that connects the narrations to the fabula and allows to derive the *syuzhet* (or plot) as defined by Russian formalism.

The fabula is built on top of events (including actions), connected to each other by three kinds of relations:

1 A *mereological* relation, relating events to other events that include them as parts, e.g. the birth of Dante is part of the life of Dante.

2 A *temporal occurrence* relation, associating each event with a time interval during which the event occurs; an event occurs before (or during, or after) another event just in case the period of occurrence of the former event is before (or during, or after) the period of occurrence of the latter event.

3 A *causality* relation, relating events that in normal discourse are predicated to have a *cause-effect* relation, e.g. the eruption of the Vesuvius caused the destruction of Pompeii.

Each narration of a fabula consists of one or more narrators and a text, which is *authored by* (another relation in the conceptualisation) the narrator(s) and constitutes the narration proper. The text may be structured in parts by a *textual mereology* relation; each part, or a subpart thereof, is connected by an *event reference* relation to the event of the fabula that the part narrates. This relation can be used to obtain the *plot* of the narrative, i.e. the sequence of events not in the chronological order but in the order defined by the narrator (with flashbacks, previews of the future, etc.). Our model reflects an interpretation of the text that represents the pieces of knowledge we considered important to create a narrative on the basis of the study of the theory of narratology. However, this is not an objective representation of the complete meaning of a text and other aspects could have been taken into account, for example the psychological aspects of the actors. At the same time, from the users' point of view, our model equally accommodates the possibility that they may be thinking of representing one possible interpretation of a text or its objective contents (whatever they may be).

A mathematical specification of the conceptualisation, aimed at giving precise definitions of the notions described above, is reported in Bartalesi et al. (2016).

4.3 Validation of the conceptualisation using the CRM

4.3.1 Selection of a reference ontology

In order to express the conceptualisation, we analysed the existing ontologies, with a preference for standards, for obvious interoperability reasons. Of course, existing ontologies would have to be extended with notions that are suited to describe narratives. However, it was paramount to minimise the number of such extensions, in order to minimise the idiosyncrasies in this research. Two top ontologies were analysed in order to understand whether their vocabularies are rich enough to capture the logic definitions of the components of narratives described in Subsection 2.2:

- The CRM, a high-level ontology that allows to integrate the information contained in data of the cultural heritage domain along with their correlation with knowledge stored in libraries and archives (Doerr, 2003).
- DOLCE+, an extension of DOLCE, the first module of the WonderWeb Foundational Ontologies Library (Masolo et al., 2003).

The analysis we performed shows that both DOLCE+ and the CRM are adequate to express the conceptualisation of narratives. Indeed, both these ontologies possess the fundamental notions of the conceptualisation, and the conditions mathematically stated in the conceptualisation can be implemented by creating suitable specialisations of the corresponding notions. However, our choice falls on the CRM for the following reasons:

- The CRM is an ISO standard since 2006 and renewed in 2014. As such, it offers a stronger guarantee under many aspects: it is widely known, it is regularly revised, it is universally accessible; at present, the same can be said also about DOLCE+, but DOLCE+ not being an ISO standard, it is unpredictable how long the present status will last.
- Since in many cases a narration is expressed through a text, the ontology should be able to represent this textual component. The CRM is specifically conceived for the cultural heritage domain, and as such it is closer to the domain of narratives especially intended as textual expressions. On the other hand, DOLCE+ is built according to a software engineering approach. For example, DOLCE+ does not include classes for specifically representing textual contents. On the contrary, the CRM has been harmonised with the FRBR ontology (FRBRoo; Doerr et al., 2008), a core ontology for bibliographic information. Therefore, it provides fundamental notions for text modelling (e.g. expressions and expression fragments) that are important for our aims. Indeed, in many cases a narration is expressed through a text and the ontology should be able to represent this textual component. FRBRoo provides notions that allow representing the aspects of text that are relevant to our purposes in a very general way. In particular, the classes Expression and ExpressionFragment are particularly useful because they allow representing any kind of text. Indeed, the class Expression in FRBRoo documentation is defined as a class that 'comprises the intellectual or artistic realisations of works in the form of identifiable immaterial objects, such as texts, poems, jokes, musical or choreographic notations, movement pattern, sound pattern, images, multimedia objects, or any combination of such forms that have objectively recognisable structures […] Expressions cannot exist without a physical carrier, but do not depend on a specific physical carrier and can exist on one or more carriers simultaneously' (https://www.ifla.org/files/assets/cataloguing/FRBRoo/frbroo_v_2.4.pdf). In the same way, an instance of the class Expression Fragment 'can be due to accident, such as loss of material over time, e.g. the only remaining manuscript of an antique text being partially eaten by worms, or due to deliberate isolation, such as excerpts taken from a text by the compiler of a collection of excerpts' (https://www.ifla.org/files/assets/cataloguing/FRBRoo/frbroo_v_2.4.pdf).
- The Special Interest Group of the CRM continuously works for expanding the domain of applicability of the ontology, and a number of extensions have been already devised. One of these extensions, the CRMinf, was exploited during the experimental validation of the ontology.

- A datalog-based, efficiently implementable logical specification of the CRM has been recently developed (Meghini and Doerr, 2015), which provides a basis for a practical experimentation of the ideas presented in this paper.

4.3.2 *Expression of the conceptualisation in the CRM*

Events are expressed in the CRM as instances of class *E5 Event*,[2] which 'comprises changes of states in cultural, social or physical systems, regardless of scale, brought about by a series or group of coherent physical, cultural, technological or legal phenomena'.[3] The CRM also makes it possible to express actions as instances of class *E7 Activity*, comprising 'actions intentionally carried out by instances of *E39 Actor* that result in changes of state in the cultural, social, or physical systems documented'. The class *E7 Activity* is a subclass of *E5 Event*.

Time intervals are represented in the CRM by the 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 relations defined on the events (and actions) of the *fabula* are expressed by the following CRM properties:

- The *mereological* relation is represented using the property *P9 consists of (forms part 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'. Note that *E5 Event* is a subclass of *E4 Period*, therefore *P9* can also be used as an event mereology.

- The *event occurrence* relation is represented by the CRM property *P4 has time-span (is time-span of)*, which 'describes the temporal confinement of an instance of an *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 occurrence time. To this end, it introduces seven properties (*P114* to *P120*) mirroring the temporal relations formalised by Allen's temporal logic (Allen, 1984). For instance, the CRM property *P117 occurs during (includes)* 'allows the entire *E52 Time-Span* of an *E2 Temporal Entity* [including events] to be situated within the time-span of another temporal entity that starts before and ends after the included temporal entity', and mirrors the *during* relationship of Allen's temporal logic.

- The *causality* relation is represented by introducing a new property of *causal dependency*, named *causallyDependsOn*. The only causal property of the CRM, *P17 was motivated by*, cannot be used for narratives since it relates activities but not events. Indeed, CRMsci (Stead and Doerr, 2015), an extension of CRM for science, defines a direct causality relation, via the property *O13 triggers*, which 'associates an instance of *E5 Event* that triggers another instance of *E5 Event* with the latter […]; in that sense it is interpreted as the cause'. However, this property is inadequate to the needs of narratives, whose events may be separated by possibly long periods of time.

Turning to the expression of narrations, narrators are represented by the instances of the CRM class *E21 Person*. The authoring relation between a narrator and the text of the narration is represented through the event of the creation of the text, an instance of the CRM class *E65 Creation*. This creation event connects to the narrator by the CRM property *P14 carried out by* and to the narration by the CRM property *P94 has created*. The narration itself is an instance of the CRM class *E73 Information Object*, which identifies 'immaterial items, such as poems, jokes, data sets, images, texts, multimedia objects, procedural prescriptions, computer program code, algorithm or mathematical formulae, that have an objectively recognizable structure and are documented as single units'. The mereology of the text is represented using the CRM property *P106 is composed of*, connecting a structural whole to its parts. However, *P106* represents the structure of the text as defined by the author, and the units of this structure do not necessarily coincide with the units that narrate a single event. It may happen that a structural unit narrates more than one event, but even if it narrates a single event, the event may be narrated in a small portion, or fragment, of a unit (e.g. the portion 'Dante was baptised in Florence' of the sentence 'On Saturday 26 March 1266 Dante was baptised in Florence, as he states himself in Inferno XIX 17'). In order to factor out the chunks of text that narrate a single event, we use therefore a specific class, the FRBRoo class *F23 Expression Fragment*. FRBRoo (Doerr et al., 2008) is a bibliographic ontology resulting from the harmonisation of the FRBR ontology and the CRM. As such, it extends the CRM with bibliographic-specific classes and properties. In particular, class *F23* comprises 'parts of Expressions and these parts are not Self-Contained Expressions themselves' and is a subclass of *E73 Information Object*. Expression fragments are connected to the structural units of text where they belong via the *P106* property. Finally, in order to express the relation between a fragment of text and the event narrated by the fragment, we assume that a fragment is related to a number of propositions, which collectively formalise the content of the fragment. Such relation is expressed by the property *hasTextFragment*. Each such related 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 (not exclusively) of a subject, a predicate and an object. The relations between a proposition and its constituent parts are expressed by borrowing three properties borrowed from the RDF vocabulary: *rdf:subject*, *rdf:predicate* and *rdf:object*. Overall, then, the relation between a fragment of text and the narrated event is obtained as follows: an expression fragment is about some proposition having the event as a subject (see Figure 3). This completes the ontology for narratives.

5 Ontology experimental validation

In order to validate the ontology in a realistic setting, we have used it to represent a narrative of a biography of Dante Alighieri as case study. To this end, we established a collaboration with an authoritative Dante biographer who provided the text from which we extracted the narrative (Indizio, 2014). The biographer was also interested in expressing the primary sources supporting the plot, so we had to enrich our ontology to also document the process through which a certain narrative came into existence. To this end, we viewed 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 inference are borrowed from another 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. The usage of the CRMinf will be illustrated in due course, along with other minor aspects of the ontology, which helped in contextualising the represented knowledge. Generally speaking, the use of the CRMinf allowed us to describe the knowledge *provenance*, as the process of tracing the origins of knowledge (PREMIS Editorial Committee, 2015).

5.1 The semantic network

This section describes the experimental semantic network that we built over the ontology for representing our case study: the biography of Dante Alighieri.

The semantic network that we constructed to model Dante's biography and its provenance includes the *fabula* and the *narration*, as defined in our conceptualisation; 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 the CRM names, while terms reused from other vocabularies are named by 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.

Figure 1 A portion of the *fabula* (see online version for colours)

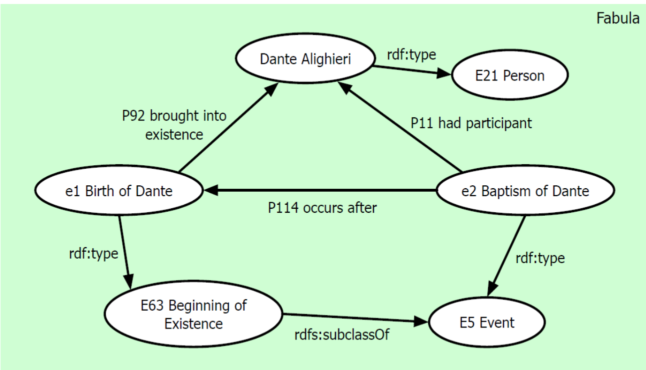

Figure 1 shows a graphical view of part of the fabula, including just two events: e1, the birth, and e2, the baptism, of Dante Alighieri. Both these events are instances of *E5 Event*, in addition e1 is an instance of *E63 Beginning of Existence*, which is a subclass of *E5 Event*. Note that we use the *rdf:type* property from the RDF vocabulary to model instance-of. Dante himself is represented by object *d*, an instance of *E21 Person*. The property *P114 occurs after* links the event e1 with the event e2.

In addition to the fabula's properties proper, an event is contextualised in terms of other properties, which we used heavily in our validation experiment. These properties can be grouped as: *Where* and *When* an event happened, and *Who* (persons) and *What* (things) were involved in it (Shaw et al., 2009; Van Hage et al., 2011). In particular:

- *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 (witnessed)* in case the location of an event is described in terms of a physical object (such as the baptism of Dante taking place in the Baptistery of Florence).
- *When* is represented by property *P4 has time-span*, which links an event with the instance of *E52 Time-Span* giving its period of occurrence.
- *Who* is represented by property *P11 had participant*, which links an event with the instance(s) of *E39 Actor*, giving the person(s) who participated in the event. 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 CRM provides the subproperty *P14.1 in the role of* in order to specify the nature of the actor's participation. Since our aim is also to give a role to the persons who do not have an active participation in an event (e.g. Pliny the Elder is an observer of the Vesuvius eruption), we related an event to an entity that represents the actor and the corresponding role. This entity is an instance of the class *ActorWithRole*, which was introduced as a new class of the ontology. The event is related to the corresponding instance of *ActorWithRole* using the new property had participant that was created for this aim. This entity has two properties: (i) *role* that links it to a string that defines the role of the actor in natural language and (ii) *subject* that links the entity to the instance of the class E39 Actor, who is the person or the person group who participates to the event.
- *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. For example, the baptism of Dante *P21 occurred in the presence of* the baptismal font, which is an instance of *E19 Physical Object*. *P12* can also be applied to concepts, which are represented as instances of *E89 Propositional Object*.

For example, ‘Rhetoric’ as mentioned in ‘Dante studied rhetoric with Brunetto Latini’.

Finally, we found it useful to relate an event with: (a) a digital object that can be used as a representation of it (e.g. the engraving ‘The Death of Beatrice’ (http://www.europeana.eu/portal/it/record/9200365/BibliographicResource_1000055750062.html)) through the property *P67 refers to*; and (b) to a textual annotation (e.g. a comment in natural language that the narrator can write in order to add some additional explanation to the event) through the property *P3 has note*. The representation of digital objects is useful in order to apply our work to the search functionality of DLs, where the events that compose a narration can be used to link together different digital objects that would remain unrelated otherwise. The textual annotation is useful to the narrator, in order to add a textual explanation of the event.

Figure 2 The *narration* (see online version for colours)

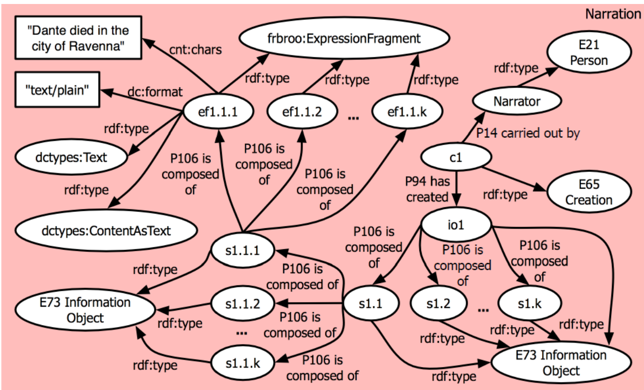

Figure 2 shows a portion of the *narration* of the life of Dante. At the bottom of the figure, the biography is represented by object io1, an instance of the class *E73 Information Object*, resulting from the event c1 of its creation carried out by the biographer. At the top of the figure, the biography is structured in chapters, paragraphs and textual fragments. We used the class *E73 Information Object* for representing chapters and paragraphs and *F23 Expression Fragment* of FRBRoo (Doerr et al., 2008) for textual fragments.

Each instance of *F23 Expression Fragment* has a textual content which is modelled according to the recommendations of the W3C’s Content in RDF (Koch et al., 2017), as shown in the top left corner of Figure 2. Specifically:

1 The *format* property of the Dublin Core (http://dublincore.org), giving the MIME media type of the instance. This allows distinguishing, e.g. embedded content in plain text from content encoded in HTML.

2 The *chars* property in the Content Namespace in RDF, giving the sequence of the characters of the content.

3 The *dctypes:Text* class, to indicate that the instance represents a resource primarily intended to be read.

4 The *cnt:ContentAsText* class, whose instances represent textual content.

Figure 3 The structure of proposition p1 (see online version for colours)

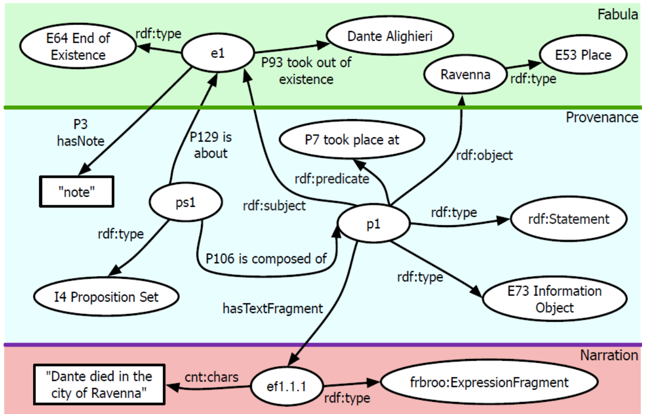

Figure 3 shows a portion of the provenance, centred around proposition p1. This proposition expresses that its subject, the death of Dante (event e1), stands according to its predicate *P7 took place at* with its object, the city of Ravenna. Therefore, p1 says that Dante died in Ravenna. Note that p1 is part of ps1, a *I4 Proposition Set*.

Figure 4 The *provenance* (see online version for colours)

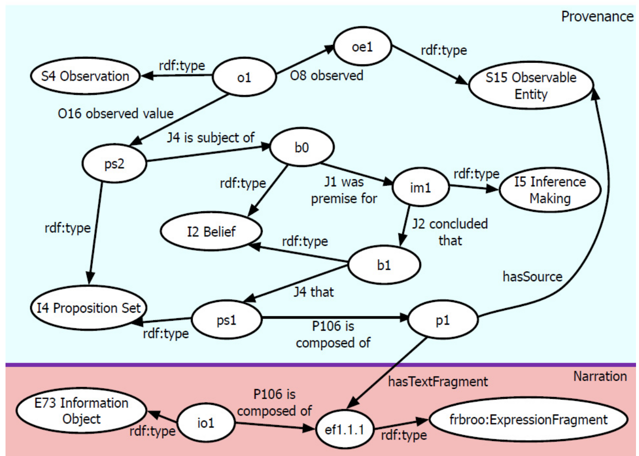

Figure 4 shows another portion of the *provenance*, representing the inferential process of the biographer. Everything starts from the reading of a source, modelled as an event of observation, o1, that is an instance of the class *S4 Observation*. o1 *08 observed* the observable entity oe1, while o1’s *O16 Observed value* is a proposition set ps2. Because they result from direct observation, the propositions in ps2 are believed to be true by the biographer, so ps2 *J4 is subject of* a belief b1 which *J1 was the premise for* an inference making event im1, which *J2 concluded that* belief b2 is the case. b2 is the belief of a set of propositions ps1 which, as it has already been mentioned, includes the proposition p1 asserting that Dante was born in Florence. So, the connection between the reading of the Divine Comedy and the proposition that Dante was born in Florence is established through an inference making event, between the propositions resulting from the reading and the propositions including the conclusion. Inference making, in turn, requires that inference maker believes both the premises and the conclusions, so beliefs are involved in the process.

6 Narrative construction

In order to construct the semantic network required for experimentation, we developed a tool supporting the user in creating a narrative through a web-based interface developed in HTML5 and JavaScript. The tool automatically populates the top classes of our ontology by looking up the Wikidata knowledge base (https://www.wikidata.org). Wikidata is a free collaborative knowledge base developed and hosted by the Wikimedia Foundation, containing structured knowledge extracted from Wikipedia and several other collaborative projects (https://www.wikidata.org/wiki/Wikidata:Sister_projects). Once the classes are populated, the tool allows the user to connect the so collected entities in a narrative, according to the ontology illustrated thus far. The automatic class population happens in two steps. In the first step, the tool queries the Wikipedia SPARQL endpoint (https://query.wikidata.org) for a list of all encyclopedia pages that are linked to the one about the given subject of the narrative (e.g. Dante Alighieri). As a result, a list of resources related to the subject is obtained. In the second step, the tool queries the Wikidata API to download JSON representations of the retrieved resources that correspond to Wikipedia pages. This allows to obtain the Wikidata URIs of these resources, which the tool then automatically assigns as instances to the different classes of our ontology. The assignment takes place through a mapping that links the definitions of the Wikidata classes to those of our ontology. The mapping is given in Table 1.

Table 1 Mapping between the classes of Wikidata and those of our ontology

Wikidata	*Our ontology*
Q5 human	E21 Person
Q16334295 group of humans	E74 Group
Q7184903 abstract object	E89 Propositional Object
Q223557 physical object	E19 Physical Object
Q17334923 location	E53 Place
Q234460 text	E73 Information Object
Q478798 image	
Q340169 media	
Q186081 time interval	E52 Time-Span
Q1190554 event	E5 Event

In compliance with the Linked Data recommendations, for each instance of the classes, the tool extracts not only a URI of the resource, but also its name and its description in several natural languages. The user can also create objects that are not present in Wikidata, if needed. User-created instances are assigned URIs defined automatically by the tool, and their class is selected manually by the user. After Wikidata resources are identified, the user can start creating events (e.g. the birth of Dante) in order to construct a narrative (e.g. the biography of Dante). Each event requires at least a title and a time span. It can also contain a textual note and one or more connected digital objects. Events must also be connected to one or more Wikidata or user-created resources that take part in the event (e.g. the mother of Dante in the event of the birth of Dante). Each connection between an event and a resource is expressed by a proposition in our model, such as p1 above stating that the birth of Dante took place in Florence. For any proposition, the user can provide: (i) one or more primary sources from which the proposition has been inferred; (ii) for the biography of Dante we took as case study, the information about the biographer's work (e.g. the text and structure) from which the proposition was extracted. Once the user has completed the creation of the narrative, it is saved to a JSON object. Then, the tool allows to automatically translate the data included in the JSON object into an OWL (Web Ontology Language) graph using an SW we developed in Java. Finally, the OWL graph is stored in a Blazegraph (https://www.blazegraph.com) triple store. The visualisation of the stored knowledge is available in the form of: (i) events organised on a timeline, (ii) graphical visualisation of the entities that compose each event and of the entities with their related events, and (iii) tables reporting information about the primary sources of the events and the events that occurred in a specified period of time.

The representation of the Dante Alighieri case study using the semi-automated tool took about 5 hours of work for two people, starting from a Word document of 40 pages (83,688 characters) provided by the biographer. The resulting semantic network contains 83 events of the life of Dante Alighieri and 206 propositions, for a total of 7379 triples.

Unfortunately, we cannot freely disseminate the narrative based on the text provided by the biographer because it is under copyright. However, another example of a narrative of Dante Alighieri's life (https://en.wikipedia.org/wiki/Dante_Alighieri), based on the Wikipedia page of the poet and developed using our tool, is available on our website (https://dlnarratives.eu/narratives.html).

The tool that we developed is semi-automated. In order to make it fully automated, it could be possible to rely on NLP techniques; however, there is a trade-off between level of automation and accuracy of the information, in the sense that automatic techniques are prone to introducing errors in narratives. We followed a semi-automated approach because in our research accuracy is paramount.

7 Representational adequacy of the ontology

The experiment to build a formal narration of Dante's life allowed performing a first evaluation of the representational adequacy of the ontology. During the creation of the narrative, the problem of finding the appropriate classes and properties of the ontology was faced, in order to formally represent the content of the biography. This formal representation was performed without changing the ontology. Since the ontology directly reflects the structure of a narrative as defined in narratology, this can be considered as an indication that narratology has reached a certain maturity, and that the structure of narratives as defined by narratology is solid, in spite of the fact that certain notions remain difficult to formalise.

A second and more thorough evaluation of the semantic network was performed by the expert who provided the biography of Dante used to build the network. The exam was organised in a dialogic fashion: a brief list of questions was given to the expert, each addressing a specific aspect of this representation.

- As first question, the expert was asked to evaluate the representational adequacy of the ontology, which is the ability of the ontology to capture, in a formal way, the salient aspects of the provided narrative as well as its provenance. To this end, the expert explored the narrative on the timeline and looked at the contextualisation of the individual events of the fabula. The evaluation was *positive* (score 4, in a scale from 1 to 5) and the scholar confirmed that the ontology was able to express all the formalisable knowledge related to the events and their relations, as described in the analysed text.
- The second question was about the usefulness of using external resources to enrich the narrative. The scholar appreciated the possibility to extend the knowledge on each event using related resources, such as Wikipedia pages, digital objects descriptions relating to the event and included in external digital libraries such as Europeana, and related images extracted from Wikimedia Commons (https://commons.wikimedia.org). In his opinion, the semantic network shape of the knowledge could be exploited to make the content of the network easy to consume also for the non-experts.
- As third question, the expert was asked to evaluate the representation of the provenance of the knowledge. In his opinion, the provenance gives crucial technical information to a scholar, e.g. the primary sources supporting an event. Having this information available allows supporting the narrative with evidence and allows readers to verify the trustworthiness of the biographical reconstruction.
- Finally, as last question the expert was asked to evaluate the possibility to have different narratives of the same topic, created by different scholars using the tool. He answered that this allows comparing results thereby identifying common points and differences in the created narratives. In particular, the investigation of the different primary sources used to identify the events of the fabula could be very interesting in a historical reconstruction point of view.

8 Conclusions

In this paper, we have given a conceptualisation of narratives based on notions derived from narratology and AI. Specifically, our conceptualisation consists of a fabula and one or more narrations of it, linked to the fabula by an event association relation. In order to validate such conceptualisation, we have expressed it in an existing ontology, the CIDOC CRM, and endowed it with provenance knowledge, also expressed in an extension of the CRM named CRMinf. This expression has been used to model a narrative of the life of Dante Alighieri, provided by a biographer with whom we have collaborated. The model of the life of Dante has been created in a semi-automatic way with the support of a tool for narrative construction we implemented. The tool retrieves and assigns URIs to the instances of the classes of the ontology using the Wikidata knowledge base as resource and also facilitates the construction and contextualisation of events, and their linking to form the fabulae of narratives. A first qualitative evaluation of the ontology was performed. A brief list of questions was given to the expert, each addressing a specific aspect of this representation. The evaluation is based on the semantic network representing the narrative of Dante's life and concerns the representational adequacy of the ontology.

The conceptualisation that we developed, its expression in the CRM and the narrative-building tool are not limited to represent and build biographies. The structure of the text can be represented using classes and properties from the FRBRoo ontology, resulting from the harmonisation of the CRM with FRBR. Since we represent the textual fragment as a basic element, any text which describes events can be represented using our approach.

In order to validate the ontology on different domains, representing different types of narratives, the narrative-building tool and the underlying ontology are currently being used by a Computer Science professor at the University of Pisa to represent the history of Informatics, focussing on the Turing Award, and by a researcher in Computational Biology at the Italian National Research Council (CNR) to narrate the discoveries related to the giant squid (https://dlnarratives.eu/timeline/squid.htm). The development of the ontology and the tool constitutes the first steps towards the introduction of narratives in DLs. Indeed, narratives can improve the discovery functionality of DLs by connecting the events that compose them to the digital objects contained in the DLs. As first step to reach this aim, we developed a narrative about the life of the Austrian artist Gustav Klimt (https://dlnarratives.eu/timeline/klimt.html), in which the events are enriched with the digital objects extracted from the Europeana digital library but we aim at investigating this problem in a future study. As future work, we have also planned to extend the functionalities of our tool to automatically retrieve entities not only from Wikidata but also from XML encoded texts and traditional databases.

References

Allen, J.F. (1984) 'Towards a general theory of action and time', *Artificial Intelligence*, Vol. 23, No. 2, pp.123–154.

Ang, K., Yu, S. and Ong, E. (2011) 'Theme-based cause-effect planning for multiple-scene story generation', *Proceedings of the 2nd International Conference on Computational Creativity*, Mexico City, Mexico, pp.48–53.

Astrova, I., Koschel, A., Lukanowski, J., Martinez, J.L.M., Procenko, V. and Schaaf, M. (2014) 'Ontologies for complex event processing', *International Journal of Computer, Information Science and Engineering*, Vol. 8, No. 5, pp.556–566.

Bal, M. (1997) *Narratology: Introduction to the Theory of Narrative*, University of Toronto Press, Toronto.

Bartalesi, V., Meghini, C. and Metilli, D. (2016) 'Steps towards a formal ontology of narratives based on narratology', *Proceedings of the International Workshop on Computational Models of Narrative 2016*, Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik, Wadern, Germany, pp.1–10.

Cua, J., Ong, E., Manurung, R. and Pease, A. (2010) 'Representing story plans in SUMO', *Proceedings of the NAACL HLT 2010 Second Workshop on Computational Approaches to Linguistic Creativity*, Association for Computational Linguistics, Los Angeles, CA, USA, pp.40–48.

Davidson, D. (2001) *Essays on Actions and Events: Philosophical Essays*, Vol. 1, Oxford University Press, Oxford.

Doerr, M. (2003) 'The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata', *AI Magazine*, Vol. 24, No. 3, pp.75–92.

Doerr, M., Bekiari, C. and LeBoeuf, P. (2008) 'FRBRoo, a conceptual model for performing arts', *2008 Annual Conference of CIDOC*, Athens, Greece, pp.6–18.

Doerr, M., Gradmann, S., Hennicke, S., Isaac, A., Meghini, C. and van de Sompel, H. (2010) 'The Europeana data model (EDM)', *World Library and Information Congress: 76th IFLA General Conference and Assembly*, 10–15 August, Gothenburg, Sweden, pp.10–12.

Doerr, M., Kritsotaki, A., Rousakis, Y., Hiebel, G. and Theodoridou, M. (2015) *CRMinf: The Argumentation Model*. Available online at: http://www.ics.forth.gr/isl/CRMext/CRMinf/docs/CRMinf-0.7.pdf (accessed on 27 February 2017).

Doll, W.J. and Torkzadeh, G. (1988) 'The measurement of end-user computing satisfaction', *MIS Quarterly*, Vol. 12, No. 2, pp.259–274.

Frixione, M. and Lieto, A. (2012) 'Representing concepts in formal ontologies: compositionality vs. typicality effects', *Logic and Logical Philosophy*, Vol. 21, No. 4, pp.391–414.

Greimas, A.J., McDowell, D. and Velie, A.R. (1983) *Structural Semantics: An attempt at a Method*, University of Nebraska Press, Lincoln, NE.

Herman, D. (2000) 'Narratology as a cognitive science', *Image and Narrative*. Available online at: http://www.imageandnarrative.be/inarchive/narratology/davidherman.htm

Hunter, J. (2003) 'Enhancing the semantic interoperability of multimedia through a core ontology', *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, No. 1, pp.49–58.

Indizio, G. (2014) *Problemi di biografia dantesca*, Longo, Ravenna. [In Italian]

Isaac, A. (2013) *Europeana Data Model Primer*. Available online at: http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf

Kenny, A. (2013) *Poetics*, Oxford World's Classics, Oxford University Press, Oxford.

Koch, J., Velasco, C.A. and Ackermann, P. (2017) *Representing Content in RDF 1.0*. Available online at: https://www.w3.org/TR/Content-in-RDF10/ (accessed on 27 February 2017).

Kowalski, R. and Sergot, M. (1989) 'A logic-based calculus of events', *Foundations of Knowledge Base Management*, Springer, Berlin, pp.23–55.

Levi-Strauss, C. (1963) 'Structural analysis in linguistics and in anthropology', *Semiotics: An Introductory Anthology*, Vol. 1, pp.110–128.

Lieto, A. and Damiano, R. (2014) 'A hybrid representational proposal for narrative concepts: a case study on character roles', *OASIcs – OpenAccess Series in Informatics*, Vol. 41, Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik, Wadern, Germany.

Lombardo, V., Battaglino, C., Pizzo, A., Damiano, R. and Lieto, A. (2015) 'Coupling conceptual modeling and rules for the annotation of dramatic media', *Semantic Web*, Vol. 6, No. 5, pp.503–534.

Manola, F. and Miller, E. (2004) *RDF Primer*, W3C Recommendation, WWW Consortium. Available online at: http://www.w3.org/TR/rdf-primer/

Masolo, C., Borgo, S., Gangemi, A., Guarino, N. and Oltramari, A. (2003) *WonderWeb Deliverable D18: Ontology Library (Final)*, ICT Project, Laboratory for Applied Ontology, Trento.

McCarthy, J. (1961) 'A basis for a mathematical theory of computation, preliminary report', Papers presented at the *May 9–11, 1961, Western Joint IRE-AIEE-ACM Computer Conference*, ACM, Los Angeles, CA, USA, pp.225–238.

McCarthy, J. and Hayes, P.J. (1969) 'Some philosophical problems from the standpoint of artificial intelligence', *Readings in Artificial Intelligence*, Vol. 12, No. 2, pp.431–450.

Meghini, C. and Doerr, M. (2015) *A First-Order Logic Expression of the CIDOC Conceptual Reference Model*. Available online at: http://new.cidoc-crm.org/sites/default/files/20150805-document.pdf

Meghini, C., Spyratos, N., Sugibuchi, T. and Yang, J. (2014) 'A model for digital libraries and its translation to RDF', *Journal on Data Semantics*, Vol. 3, No. 2, pp.107–139.

Meister, J.C. (2003) *Computing Action: A Narratological Approach*, Vol. 2, Walter de Gruyter, Berlin.

Miller, R. and Shanahan, M. (1994) 'Narratives in the situation calculus', *Journal of Logic and Computation*, Vol. 4, No. 5, pp.513–530.

Miller, R. and Shanahan, M. (2002) 'Some alternative formulations of the event calculus', *Computational Logic: Logic Programming and Beyond*, Springer, Berlin, pp.452–490.

Mueller, E.T. (2014) *Commonsense Reasoning: An Event Calculus Based Approach*, Morgan Kaufmann, San Francisco, CA.

Niles, I. and Pease, A. (2003) 'Mapping WordNet to the SUMO ontology', *Proceedings of the IEEE International Knowledge Engineering Conference*, IEEE, Palo Alto, CA, USA, pp.23–26.

Pease, A., Niles, I. and Li, J. (2002) 'The suggested upper merged ontology: a large ontology for the semantic web and its applications', *Working Notes of the AAAI-2002 Workshop on Ontologies and the Semantic Web*, Vol. 28, AAAI, Edmonton, Canada.

Poli, R., Healy, M. and Kameas, A. (2010) *Theory and Applications of Ontology: Computer Applications*, Springer, Berlin.

PREMIS Editorial Committee (2015) *Data Dictionary for Preservation Metadata, Version 3.0*. Available online at: http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf

Raimond, Y. and Abdallah, S. (2007) *The Event Ontology*. Available online at: http://motools.sourceforge.net/event

Russell, S.J., Norvig, P., Canny, J.F., Malik, J.M. and Edwards, D.D. (2003) *Artificial Intelligence: A Modern Approach*, Vol. 2, Prentice Hall, Upper Saddle River, NJ.

Sandewall, E. (1989) *Filter Preferential Entailment for the Logic of Action in Almost Continuous Worlds*, Universitetet i Linköping/Tekniska Högskolan i Linköping, Linköping.

Shaw, R., Troncy, R. and Hardman, L. (2009) 'LODE: linking open descriptions of events', *The Semantic Web*, Springer, Berlin, pp.153–167.

Shklovsky, V. (1965) *Russian Formalist Criticism: Four Essays*, University of Nebraska Press, Lincoln, NE.

Stead, S. and Doerr, M. (2015) *Definition of the CRMsci*. Available online at: http://www.ics.forth.gr/isl/CRMext/CRMsci/docs/CRMsci1.2.3.pdf (accessed on 27 February 2017).

Todorov, T. (1969) *Grammaire du Décaméron*, Mouton, The Hague. [In French]

Van Hage, W.R., Malaisé, V., Segers, R., Hollink, L. and Schreiber, G. (2011) 'Design and use of the simple event model (SEM)', *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 9, No. 2, pp.128–136.

Van Harmelen, F., Lifschitz, V. and Porter, B. (2008) *Handbook of Knowledge Representation*, Vol. 1, Elsevier, Amsterdam.

Wenjun, W., Yingwei, L., Xinpeng, L., Xiaolin, W. and Zhuoqun, X. (2005) 'Ontological model of event for integration of inter-organization applications', *International Conference on Computational Science and Its Applications*, Springer, Berlin, pp.301–310.

Notes

1 Europeana is the largest DLs in the cultural heritage sector, offering a unique access point to about forty million objectsfrom more than two thousand European institutions. See http://www.europeana.eu

2 As a notational convention, elements of the CRM vocabulary are written in italics.

3 All quotations in this Section are from the CRM specification version 6.2.2, available from http://www.cidoc-crm.org/Version/version-6.2