Collections revisited from the perspective of historical testimonies

DOI:10.1504/IJMSO.2019.102677

(Article begins on next page)

Collections Revisited from the Perspective of Historical Testimonies

Annamaria Goy* and Diego Magro

Dipartimento di Informatica,
Università di Torino,
C. Svizzera 185, Torino,
Italy
E-mail: annamaria.goy@unito.it
E-mail: diego.magro@unito.it
*Corresponding author

Abstract:
This paper presents the results of an ontological analysis of *collective entities*, as an essential step towards the definition of a rich semantic model underlying ontology-driven applications in the historical and Cultural Heritage domains. The major contributions of our proposal are the following: (a) An explicit distinction of contingent and necessary features, that led us to formalize our ontology by means of modal logics. (b) A description of collective entities from a diachronic perspective (thus including singletons and empty collections). (c) An analysis of the inferences enabled by the characterization of collective entities and by the inclusion relationships. (d) A representation of the inclusion relationships that does not imply existential dependence. (e) A distinction between emerging and created collective entities. In the paper, we present an in-depth ontological analysis of these aspects and provide a sound formalization for it.

Keywords: collections; collective entities; sets; ontology; ontologies; semantic model; ontological analysis; ontology-driven applications; semantic metadata; inference; modal logics; historical archives; historical testimonies; Cultural Heritage.

1 Introduction

Ontology-driven applications have been designed and developed in many fields, from healthcare to e-government, and they play a major role in the Cultural Heritage domain. Data related to cultural assets (collections and catalogs) are increasingly available online, both to human users and to software applications. However, these data, although available, are often very poor – since catalog metadata usually do not describe (in details) the content of the resources (e.g., the entities and the events they refer to), and very difficult to be actually accessed, used, and exploited by a large part of their potential users. In this scenario, we think that the role of metadata, that represent the main "bridge" between users and resources, should be enhanced, in order to turn them into an actual means to discover, reach and use Cultural Heritage resources [CH14]. More precisely, we think that the enhancement of metadata can be achieved through their semantic enrichment: Standard catalog metadata should be coupled with rich formal semantic representations of the content of the resources, enabling a real content-based access to documents and objects stored in historical archives, museum collections, etc. In turn, such formal semantic representations need a formal and rich semantic model, providing the suited "vocabulary".

Therefore, the envisaged metadata semantic enrichment requires: (1) a rich semantic model (ontology); (2) a semantic knowledge base containing (formal) descriptions of the content of the resources, along with suitable ways to acquire them.

This paper focuses on the first requirement, and in particular on a specific aspect of the semantic model, i.e., the ontological characterization of *collective entities*.

The discussion about the nature and the ontological characterization of collective entities (students, citizens, Fiat-Chrysler workers, members of the USA Government, countries that are members of the European Union, laws in force in Italy, books in the Ancient Library of Alexandria, and so on) has a long tradition and has been carried on across many disciplines, ranging from Philosophy to Sociology, from Linguistics to Computer Science – see [Sea95] [Fre84] [BP01] [Obe05] among many others – as part of the wider debate about the definition of formal and computational ontologies, suited to different purposes [SS09] [GM15].

Moreover, collective entities plays a major role in the historical domain, as emerged from the study we carried out on archival documents representing historical and personal testimonies – such as biographies, letters, war reports, etc. – extracted from books narrating events occurred in Piemonte (North-West of Italy) between 1943 and 1945, and representing the Partisans fight against the Fascist regime and the Nazi occupation.

The ontological analysis and the formal characterization of collective entities presented in this paper, besides contributing to the wide multi-disciplinary debate about this topic (accounted for in Section 2.2), represents a step of paramount importance in order to define the needed semantic model (first requirement mentioned above).

The major contributions of this paper to the ontological modeling of collective entities can be summarized by the following points, that will be discussed in details in Sections 2.2 and 3:

- An explicit distinction of what is contingent from what is necessary (that is extremely relevant in history, representing our current application the domain; see Section 2.1). We believe that such an approach sheds light on several aspects of collective entities that otherwise would remain obscure. In particular, this distinction enables us to separate, as far as collective entities are concerned, temporary and enduring features, and to split enduring features into contingent and necessary features. As a consequence of this approach, we base our formalization on modal logics.

- The possibility of describing collective entities (also) from a diachronic perspective, i.e., describing their evolution over time (e.g., their creation, their changes, their end of existence). This point of view, again, is highly relevant in the historical domain. In particular, such a diachronic perspective seems to require admitting empty collections and singletons, since, for example, a just created collection can be empty, then it can grow up by first acquiring one member, then a second one, and so on. This implies that, in our model, collections with at least two members are possible, but not necessary (as, instead, they are in those approaches that do not admit empty collections and singletons).

- A detailed analysis of the main inferences enabled by the characterization of collective entities and by the inclusion relations they are involved in.

- The possibility of representing the inclusion relationships between collective entities without assuming any existential dependence between them. For instance, in a conflicting event – such as a liberation war – a collection (e.g., people fighting for liberation) can be formed by several sub-collections (e.g., groups with different political orientations) and some of those sub-collections could continue existing when the war is over and the main collection has been broken up.

- The possibility of distinguishing collective entities directly and explicitly created by an agent community (e.g., the Members of the Italian Parliament) and collective entities that emerge spontaneously or as an indirect consequence of the action of an agent community (e.g., consumers).

In Section 2 we introduce the framework of our proposal by mentioning two on-going related projects, briefly presenting the backbone structure of our historical event ontology, and sketching the role it plays in these projects (Section 2.1): This will provide the context for the contribution presented in this paper. Then, we discuss some related works, concerning event ontologies and the ontological analysis of collective entities (Section 2.2). Section 3 is the core of the paper: we start by defining collective entities as collections of individual objects, having their own individuality, and we discuss the distinction between collective entities and ontology classes, and between collective entities and sets (Section 3.1). We then propose a formal characterization of collective entities (Section 3.2): we introduce the main reused ontologies (Section 3.2.1); we define and discuss the membership relation and the issues related to singletons and empty collections (Section 3.2.2); we formally characterize collections in relation to their members (Section 3.2.3), and in relation to inclusion (Section 3.2.4), and we discuss the inferences thus enabled (Section 3.2.5); we analyze the distinction between emerging and constructed collections (Section 3.2.6).

Section 4 concludes the paper by providing an overview of the main open issues.

2 Motivations and Related Work

2.1 The HERO Ontology

The work presented in this paper is part of a larger activity, carried on in the framework of two closely related national projects:

- Harlock'900 (running 2016-2018), which involves the Department of Computer Science of the University of Torino and the Fondazione Istituto Piemontese Antonio Gramsci (www.gramscitorino.it), a non-profit institute promoting research on contemporary history. The project aims at implementing a semantic layer, based on computational ontologies of historical events, to enrich catalog metadata with information about the content of archival resources; see [CGM17].

- PRiSMHA (Providing Rich Semantic Metadata for Historical Archives) – running 2017-2020 – funded by Compagnia di San Paolo and University of Torino. It involves the Computer Science and the Historical Studies Departments of the same university, and is based on a close collaboration with the Polo del '900 (www.polodel900.it) and, in particular, with the archives and library of the Fondaz. Ist. Piemontese A. Gramsci. The project aims at designing and developing a crowdsourcing collaborative platform enabling experts and trusted users to participate in building rich semantic representations of the content of archival resources. The interaction model will be driven by the underlying ontology and will be supported with suggestions provided by automatic information extraction techniques; see [GDL^{+}17] and [Rov16, RNPG17].

In both projects, we focus our investigation on resources (documents, texts, pictures, etc.) related to the Italian history of the 20th Century. An overview of the overall approach is provided by [GMR15].

One of the most important activities, representing the first milestone of Harlock'900 and PRiSMHA, is the definition of an ontological model, HERO (Historical Event Representation Ontology), providing the "vocabulary" for representing the content of the analyzed historical resources.

Figure 1 The role of HERO in Harlock'900 and PRiSMHA projects

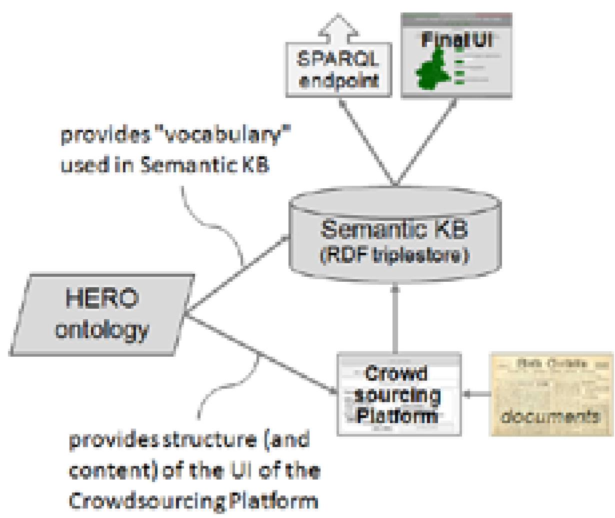

Figure 1 provides a global picture showing the role of the ontology in the two projects. In both perspectives – i.e., the Final User Interface, enabling a content-based access to archival documents, and the Crowdsourcing Platform, enabling the collaborative construction of formal semantic representations of documents content – the ontology represents the system conceptual knowledge, encompassing the classes and properties used in the Semantic Knowledge Base (KB), which contains the formal semantic representations produced through the Crowdsourcing Platform. Moreover, the HERO classes and properties provide the structure (and the content) of the form-based User Interface of the Crowdsourcing Platform.

HERO provides a rich characterization of the concepts and relations that emerged to be relevant from an accurate ontological analysis of the domain. Moreover, it relies on the well-known and cognitive-grounded foundational *DOLCE – Descriptive Ontology* *for Linguistic and Cognitive Engineering* [MBG+03, BM09] and its extension *DnS – Description and Situations* [GM03, GBCL05].

HERO is composed of five modules:

- HERO-TOP: It includes the top layer, i.e. the most general concepts, directly linked to DOLCE classes and relations.
- HERO-EVENT: It includes classes and properties related to the representation of (historical) events, their typologies, and participation modalities(see [GMR17, GMR18]).
- HERO-PLACE: It includes classes and properties modeling entities that can be georeferenced on a map (e.g., cities, rivers, countries, but also buildings and streets).
- HERO-TIME: It includes classes and properties modeling time intervals, following Allen's Interval Algebra [All83].
- HERO-ROCS: It includes classes and properties for the representation of Roles, Organizations, Collections (i.e. collective entities), and Sets.

  In this paper, we present the ontological analysis underpinning the part of the HERO-ROCS module devoted to *collective entities*.

2.2 Related Work

The notion of *event* and its participants has been considered as a major pattern to model the historical domain in several projects aimed at providing a smart access to cultural resources [ST17, NPD17, Zar15]. Moreover, *organizations* and *collective entities* seem to be first-class types of participants in historical events.

Several ontological models have been developed, aimed at modeling the concept of *event*; see, for instance: SEM [vHS+11] (used in projects such as Agora and DIVE [vdAAC+10, dBOI+14]), LODE [STH14], EO [RA07], the ABC ontology [LH01], the event ontology defined within the CultureSampo project [HLTM12], Europeana Data Model [Eur16], CIDOC-CRM [Doe03, BDOS15], and the Event Model F [SFSS09].

However, none of the mentioned semantic models provide a detailed ontological characterization of collective entities (i.e., collections of objects or people). In order to enrich event ontologies with a sound and well-founded account of those types of entities, we claim that a deep ontological analysis of them, their status, their properties, and relations is required.

Probably, one of the the most influencing work on this topic is presented in [BCGL06], where collective entities (there called "collections") are modeled within the framework of DOLCE, DnS, and DDPO (DOLCE + DnS Plan Ontology; [GBCL05]): A *collection* is a *social object* containing only *endurants*. Moreover, collections must contain, at least, two entities (empty collections and singletons are not admitted). The membership relation – connecting a collection to its members – is characterized in terms of a "containment schema" defined on the basis of the notions of *description* and *role*. In this way, the membership relation is distinct from the constitution relation and also from parthood (since, being a *collection* a *social object* – i.e., a *non-physical endurant* – it can only have *non-physical endurants* as parts, while its members can be also *physical endurants*). Other remarkable relations in the analysis by Bottazzi and colleagues are *unifies*, that associates descriptions with collections, and *covers*, that associates a *role* (x) to a *collection* (y), specifying that x classifies all (but not only) the members of y. Finally, a *collective* is a *collection* unified by a plan (that is a specific kind of description) and having only agents as members. A more detailed comparison between the approach by Bottazzi and colleagues and our proposal is provided in Section 3.

Precious suggestions about the ontological characterization of collective entities can be found in [WG09, GW16] (where collective entities are called "collectives"); see also [Gal12, Gal14]. In [WG09], the authors define collectives as *concrete particulars*, and more specifically, as *continuants* (endurants), implying that they are not types or abstractions (which in turn means that they are not sets). With respect to the proposal discussed in [BCGL06], the notion of *collective* proposed by Wood and Galton is broader and includes also collections and social groups. A collective is the mereological sum of its members, and membership is modeled as a parthood relation, although only some parts of a collective are members (e.g., persons are members of a crowd, but arms or legs – that are parts of the crowd – are not members of it). In particular, for each collective, there is a property such that collective members are those parts having that property. Wood and Galton's approach admits collectives with a single member, but it does not admit empty ones. Moreover, the authors propose a classification system for collectives, based on five dimensions, namely: membership, location, coherence, roles, and depth; [WG09].

The same authors discuss in [GW16] the differences between *E-collectives* and *I-collectives*: *E-collectives* are *extensional collectives*, for which membership is necessarily fixed (i.e., they specifically depend on their members); thus, members are essential parts of a collective, and the collective itself coincides with the sum of its members. *I-collectives* are *intensional collectives*, for which membership is potentially variable (it can be fixed, but only contigently); thus, members of an I-collective are never all essential parts of a it – although some of them could be: e.g., Prof. Rossi in the collective represented by "Prof. Rossi's group" is an essential part of it [GW16]. I-collectives correspond to collections in the proposal presented in [BCGL06].

In [Gui11], the author proposes an ontological analysis of collective entities (there called "collectives"), including part-whole relations, subcollectives, and membership. The author shows why collectives cannot be identified with sets: collectives are integral wholes; both the member-collective and subcollective-collective relations are part-whole relations; all subcollectives of a collective are inseparable parts of it (e.g., a subcollective cannot exist without being a part of that specific collective). Furthermore, Guizzardi's collectives have a homogeneous structure, i.e., all members should play the same role with respect to the collective (e.g., a forest is a collective of trees, while an orchestra is not a collective, since its members play different roles with respect to it).

Finally, a more "pragmatically-oriented" approach is that of the Collections Ontology, presented in [PS14], in which – following most influential ontological analyses such as those in [BCGL06] and [Gui11] – the authors present the version 2.0 of the Collections Ontology (CO), implemented in OWL 2-DL and aimed at providing a semantic model to represent collections of resources: sets (unordered non-repeatable elements), bags (unordered repeatable elements), and lists (ordered repeatable elements). The "root" class is *co:Collection*, that is the superclass of *co:Set*, *co:Bag*, while *co:List* is a subclass of *co:Bag*. One of the major properties of CO is *co:element*, linking a collection to its members: This is a general property, which can be futher constrained in order to represent either set membership or member-collective relationships; see [Gui11].

3 Ontological Analysis and Formalization of Collective Entities (Collections)

3.1 Collective Entities vs Classes and Sets

Collective entities are collections of individual objects[a] (their members) that have their own individuality, full existence, and the capability to take part in (historical) events; they can be ascribed characteristics, properties or behavior that cannot be (conveniently) reduced to those of their members.

Collective entities should not be confused either with ontology classes, or with mathematical sets (even in those ontologies where sets are represented as instances).

Collective entities can not be reduced to ontology classes (and thus collective entity membership can not be modeled as class instantiation), the main reasons being:

- Collective entities are particulars, while ontology classes are properties.
- Collective entities are instances of classes, while classes are (possibly) instances of meta-classes.
- Collective entities may come into existence at a given time and cease to exist after some time, while classes are out of time.
- In first-order ontology formalizations, collective entities belong to the quantification domain, while classes do not.
- It does not seem to be true that for each collective entity, there can be in an ontology a corresponding class such that the members of the collective entity are the instances of the class. Let us consider, for instance, the herd belonging to John, which is now in this particular grassland: It is a collective entity (whose members are the sheep of John), but it might be odd to introduce in the ontology the class $SheepOfJohn$ (whose instances would be the sheep of John). Similarly, the employees of a particular company (e.g., Fiat-Chrysler) may be considered a collective entity (they may collectively participate in events, etc.), however it could make no sense to introduce in the ontology the class $EmployeeOfFiatChrysler$.
- It does not seem to be true that for each class in an ontology there is a corresponding collective entity such that the instances of the class are the members of the collective entity. This is not only due to the fact that it seems odd to introduce collective entities corresponding to some classes (e.s. the collective entity of endurants or the one of amounts of matter), but also to the semantics of classes in ontologies. Many ontologies (including DOLCE) adopts an eternalist view, which means that they place in the quantification domain all past, present and future entities. In such a framework, the corresponding collective entity of a rigid class class (i.e. a class such that, each entity belonging to the class *necessarily* belongs to it) would always contain as members all the instances of the class, and thus also the past and future ones, while the members of a collective entity at a given time must be present at that time (Axiom A4 in Section 3.2.2). The problem is even more serious with those ontologies (including DOLCE) that adopt a possibilist view and, thus, include in the quantification domain all possibilia.

[a] 'Objects' should be intended here in a broad sense, comprising both physical and non-physical things, as well as living beings.

It is also worth remembering the main differences between collective entities and mathematical sets – see [BCGL06, Gui11]:

- Given any objects, the set containing exactly those objects can be formed, while this is not true for collective entities: Only those collective entities that it makes sense to consider (i.e., that can be suitably conceived as fully existing individuals) are actually recognized as existing entities (e.g., by the historians) or are 'socially constructed' by a community of agents. For instance, a dozen people, randomly chosen among the pedestrians that are walking now on some street of Beijing form a set of people, but (usually) they are not considered a collective entity; on the opposite, *the Italian consumers* are usually considered a collective entity having its own characteristics (e.g. expressing collective preferences, influencing the market, reacting to advertisements, etc.).

- Sets cannot acquire or lose members during their life, while collective entities can. For instance, the *French citizens* collective entity continually changes its members, but it still maintains its identity over time.[b]

- Sets are uniquely determined by their members, while collective entities are not. For instance, two collective entities of the members of two different clubs may have the same members, still being two different collective entities.

- The union, intersection and difference between two sets is still a set, while this is not true, in general, for collective entities. For instance, the union of *French citizens* and *Bhutanese farmers* (both collective entities) is likely not a collective entity.

- Typically, sets are considered as abstract entities, i.e., entities that are outside space and time and that cannot influence concrete entities. On the opposite, collective entities are concrete (meaning that they have spatial and/or temporal features) and may be causally related to other concrete entities (for instance, a fleet of warships may destroy a seaport).

- To the above-stated "classical" differences, we add the following: Non-empty sets of non-existing (possibly no longer or not yet existing) entities may exist, while this is not true for collective entities (see Axiom A4 below).

3.2 Formal Characterization

The works by Bottazzi and colleagues [BCGL06], Guizzardi [Gui11], Wood and Galton [WG09, GW16] strongly influenced the ontological analysis that led to our formal characterization of collective entities. For this reason, in the following we will also emphasize the main similarities and differences between our conceptual and formal framework and those presented in the aforementioned works.

We adopt the term *collection* to refer to collective entities in general, in line with [BCGL06] (while [WG09, GW16] as well as [Gui11] generically refer to collective entities as *collectives*).

We believe that the notions of *necessity* and *possibility* are important to reach an ontologically well-founded account for collections, even though such notions were not

[b]This does not mean that collective entities necessarily lose or acquire members during their life. We can consider, for instance the Second Triumvirate, i.e. the political alliance between Augustus, Mark Antony, and Lepidus: this was a collective entity which always had the same three members during its whole lifespan [GW16].

exploited in the above-cited works. To properly represent such aspects, we provide a formalization in a modal logic language. In particular, we express our ontology in the same logic as the reference version of DOLCE, namely the modal logic S5, plus the Barcan Formula [BvBW07]. We provide a formalization consisting of axioms, definitions and theorems. Each theorem of the presented theory can be easily proved by applying a tableau method for modal logic S5. For the sake of conciseness, we will not report any proof, but, for each theorem, we will specify the minimal set of definitions, axioms and theorems that entails it.

For the sake of readability, in the following formulae we omit both the universal quantifier and the necessity modal operator in front of them, hence each formula $\alpha(x_1, \ldots, x_n)$ should be interpreted as $\Box(\forall x_1) \ldots (\forall x_n)\alpha(x_1, \ldots, x_n)$.

3.2.1 Reused Ontologies

As stated, we performed our ontological analysis and defined our theory of collections within the ontological and formal framework provided by DOLCE [MBG$^+$03, BM09], and its extension through DnS [GM03, GBCL05, MVB$^+$04]. Here we provide a brief informal description for the concepts and relations of DOLCE and DnS that we explicitly reuse in the present paper, in order to make it as self-contained as possible. For them, we use the prefixes $dolce$: and dns:, respectively, while we do not use any prefix for the concepts and relations strictly belonging the proposed ontology (HERO).

DOLCE distinguishes four basic categories of particulars, namely: endurants (in the following: $dolce{:}Endurant(x)$), perdurants ($dolce{:}Perdurant(x)$), qualities, and abstracts. Endurants are those "particulars that are wholly present [...] at any time they are present" [MBG$^+$03] (such as buildings, streets, organizations, countries, laws, etc.). Perdurants are those entities that "happen in time" (such as events, activities, processes, etc.). Qualities are the "basic entities that we can perceive or measure" [MBG$^+$03]. Abstracts are particulars outside time and space. Each endurant participates in at least one perdurant during some time interval – e.g., a human person participates in its life; when she speaks, a person participates in that speaking activity, etc.; conversely, each perdurant has at least one object that participates in it. Among endurants, it is worth mentioning physical endurants (such as buildings, streets, etc.) and non-physical endurants ($dolce{:}NonPhysicalEndurant(x)$, such as private desires, organizations, countries, laws, etc.). Many non-physical endurants somehow depend on a community of agents (such as organizations, countries, laws, etc.): Within DOLCE, they are called *social objects* ($dolce{:}SocialObject(x)$). Time intervals ($dolce{:}TimeInterval(t)$) are special kinds of abstracts.

DOLCE distinguishes logical existence from actual existence: the latter is specified by a the presence predicate $dolce{:}presentAt(x, t)$, stating that x is actually existing during time interval t. An entity x may depend on another specific entity y (for instance, any human being specifically depends on his/her brain): Such relation is expressed in DOLCE by the predicate
$dolce{:}specificallyConstantlyDependsOn(x, y)$, whose meaning is that "at any time t, x can't be present at t unless y is also present at t" [MBG$^+$03].

DnS is an extension of DOLCE that offers a set of notions accounting for reified concepts and relations. Such reified entities belong to the domain of discourse and thus they can be predicated on. Any concept (e.g., the concepts of person, male, king, citizen, Italian citizen, worker, etc.) may be reified and placed in the domain of discourse as an instance of the $dns{:}Concept$ class, a subclass of $dolce{:}SocialObject$. The immediate relation between the reified concepts and the entities that they make reference to (e.g. the relation between the notion of king and each specific king) is expressed by the predicate $dns{:}classifies(x, y, t)$, which means that the concept x makes reference to y during time interval t (i.e., y is a x during t – for example, Louis XIV of France was a king from 14th May 1643 to 1st September 1715). Descriptions ($dns{:}Description(x)$) represent conceptualizations (e.g., theories, plans, laws, legal system, etc.) and are special kinds of $dolce{:}SocialObject$s. A description uses ($dns{:}uses(x, y)$) – more specifically it may define – concepts (e.g., an economic theory uses the concept of *unemployed people*, the Italian legal system defines the concept of *Italian citizen*, etc.).

3.2.2 The Basics

The basic relation between a collective entity (hereinafter referred to as a $collection$) and its members is expressed by the predicate $hasMember(x, y, t)$ that we take as a primitive relation and whose intended meaning is that *during time t, x has y as a member*. With such a primitive relation, we can define collections as those non-physical endurants that may have members:

(D1) $Collection(x) =_{def}$
$$dolce{:}NonPhysicalEndurant(x) \wedge \Diamond(\exists y, t)hasMember(x, y, t)$$

Definition D1 reveals much of the ontological and formal choices that underpin our ontology of collections and allows us to provide a first comparison with the other above-mentioned approaches:

1. In our approach, a collection is always a non-physical endurant. This aspect marks a difference with respect to the conceptualizations proposed in [WG09] and [Gui11], while it places our conceptualization close to the one described in [BCGL06]. In fact, in the former works, a collection is viewed as a whole and the members are parts of it. This implies both that the physical or non-physical nature of a collection depends on the nature of its members (e.g., a collection of physical endurants is a physical endurant, while a collection of non-physical endurants is a non-physical endurant), and that a collection cannot have both physical and non-physical members.[c] We believe that conceiving a collection as a non-physical endurant, no matter the nature of its members, provides a more uniform view on collections, which better captures their common traits. Such a conceptualization is consistent with the notion of collection presented in [BCGL06], while not being strictly equivalent to it. In fact, in that work, the authors adopt a constructivist approach and define a collection as a special kind of social object, i.e. as an "immaterial [...] product of a community. In this sense a social entity depends on agents who constitute, make use of, communicate about, and 'recognize' or 'accept' it by means of some sort of agreement" [BCGL06], p. 192. Moreover, according to Bottazzi and colleagues, each collection specifically depends on at least one description. Such a notion accounts for many relevant kinds of collective entities, however in some cases its scope is too narrow, as already pointed out in [WG09]). This narrowness does

[c]This is true as far as we admit that a physical endurant can have only physical parts and non-physical endurants can have only non-physical parts, which seems a reasonable assumption, accepted also in the DOLCE's framework (see Axioms Ad11 and Ad12 in [MBG$^+$03]).

reveal itself, for instance, in the historical domain, where we find both "emerging" and "socially constructed" collections (see Section 3.2.6). Therefore, we have to adopt a more general notion of collection, still viewing collections as immaterial entities (i.e., $dolce{:}NonPhysicalEndurant$s), but encompassing both "socially constructed" and "emerging" collections.

2. Being a DOLCE non-physical endurant, any collection participates in at least one perdurant during some time interval[d].

3. An ontology of collections should account for both collections of physical endurants (e.g., a fleet of ships) and collections of non-physical endurants (e.g., a union of states).[e] Therefore, given that we consider all collections non-physical endurants and that we also assume that a non-physical endurant can have only non-physical parts, our membership relation cannot be a part-whole relation. This is consistent with the formalization provided in [BCGL06], but it differs from the both the perspectives maintained in [WG09], according to which a collection is the mereological sum of its members, and in [Gui11], which views collections as integral wholes, whose members are parts of the whole.

4. In [BCGL06], the authors define the membership relation according to a containment schema. We here adopt the simpler approach of taking $hasMember$ as a primitive, by still being consistent with the notion of membership by Bottazzi and colleagues.

5. In [BCGL06], as well as in [Gui11], the authors only admit collections with at least two members, while in [WG09] the authors do admit also singleton collections. None of them admit empty collections. These restrictions may seem reasonable and are almost a forced choice if the membership relation is viewed as a part-whole relation. However, there are also good reasons for admitting both singleton and empty collections, as borderline cases. Let us consider, for instance, the collection of the laws made by the Italian Republican Parliament. If we admitted only collections with at least two members, either we would have to state that such an entity came into the world only immediately after the second law was made or we would have to admit that such an entity changes its nature during its lifespan, becoming a collection only with the appearance of the second law and being something different until just before that event. In both cases, it would be difficult to trace the products of the Italian Republican Parliament legislative process from the very beginning and in a uniform way. In general, besides the specific example, it would be difficult to consider collections under a diachronic perspective, which is of paramount importance in the historical domain. These kinds of situations suggest that at least singleton collections should be conveniently admitted, and this is the decision that we took.

   The question arises whether empty collections should be admitted, too: We believe that they should. Let us consider, for instance, the case of Italian ordinary statute regions. The Italian Republican Constitution (which is in force since 1^{st} January 1948) specifies that Italy is administratively organized into regions, five of which have a special statute and the others have an ordinary one. However, the latter have

[d] Axiom Ad35 in [MBG$^+$03].

[e] We also admit collections of both physical and non-physical endurants (e.g., the customers of a company can be both other companies or persons, and thus such a collection has both physical and non-physical members – as far as companies and persons are considered non-physical and physical endurants, respectively).

been actually created only in 1970: We can say, therefore, that the collection of the Italian ordinary statute regions was born in 1948 (since it was constructed when the Italian Constitution came into force), it remained empty until 1970 and in that year it acquired its members. Besides temporarily empty collections, we should also admit permanently empty collections (i.e. collections that remain empty for their whole lifespan), as suggests the case of local elections in the Italian Municipality of San Luca in June 2016: No one ran for those elections, thus in that case the collection of candidates (that we can always consider in each election) was defined (and created) by the administrative process that held those elections, but it remained empty for its whole lifespan.

It is important to stress that even though there can be permanently empty collections, Definition D1 entails that collections cannot be *necessarily* empty, therefore the collection of, say, married bachelors is not a valid collection in our ontological framework.

All that being said, we acknowledge that the claim that collections *must* always have at least two members flows from a valid intuition. However, we believe that we can better capture this intuition by stating that each collection always *may* have at least two members, than by stating that it *must*. This new sense of the constraint can be easily expressed in a modal framework and it actually belongs to our theory of collections (see Axiom A5).

6. It is worth pointing out that considering the membership relation a part-whole relation poses some problems with both empty and singleton collections. Let us consider the case of empty collections first.

   If we assume that the members of a collection must exhaust the collection (i.e. each part of a collection is composed of members or of parts of members; [WG09]), then an empty collection would be a sort of mereologically null instance, whose existence is at least rather counterintuitive. Releasing such exhaustiveness constraint, and thus admitting that the members may not exhaust collections (still being parts of them), would change the problem, but it would not solve it. In fact, in this case an empty collection would no longer be necessarily a sort of mereologically null instance; however, the role of those parts that are neither composed of members nor of parts of members would not be clear.

   Singleton collections raise similar problems: If a singleton collection is viewed as mereologically distinct from its member (still considering the latter a part of the former), it is not clear what the mereological difference between the collection and its member actually represents; in the other case, it should be clarified what actually makes the difference between the collection and its sole member.

   The latter considerations, along with the need of including in our universe of discourse both empty and singleton collections, provide further support to the intuition that membership is not parthood.

As in [BCGL06, WG09, Gui11], we consider only collections of endurants:

(A1) $hasMember(x, y, t) \rightarrow$

$$Collection(x) \wedge dolce{:}Endurant(y) \wedge dolce{:}TimeInterval(t)$$

Membership is both an irreflexive (Axiom A2) and asymmetric (Axiom A3) relation, as it is in [BCGL06, Gui11]:

(A2) $\neg hasMember(x, x, t)$

(A3) $hasMember(x, y, t) \rightarrow \neg hasMember(y, x, t)$

Obviously, membership is not transitive.[f]

If the membership relation holds between a collection and an entity at a certain time, then both the collection and the entity must be present at that time:

(A4) $hasMember(x, y, t) \rightarrow dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(y, t)$

This implies that a collection of non-existing (possibly no longer/not yet existing) entities – such as the married bachelors, the above-mentioned candidates in the June 2016 local elections in San Luca, the Greek hoplites or the students of that master that will start next year – either does not exist (possibly no longer or not yet) or it is empty.

In this regard, it is worth pointing out that the above-listed examples, though referring to collections whose members do not exist nowadays, actually illustrate different situations. As already stated, the collection of married bachelors would be *necessarily empty*, thus it is simply ruled out by our formal notion of collection (Definition D1); consequently, we cannot talk about it within our framework (from a formal point of view, such an entity does not belong to the domain of quantification, therefore it *does not exist in the logical sense*).

The case of the collection of the candidates in the June 2016 elections in San Luca is not a necessarily empty collection, but only a *contingently permanently empty* one that *actually existed in the past* (during the process that administered those elections). Therefore, such a collection is not ruled out by our formalization and we can talk about it within our framework, no matter if it had had no members for its whole lifespan and even though *nowadays it does not actually exist any more*, in exactly the same way as we can talk about Julius Caesar (from a formal point of view, such an entity is included in the domain of quantification, thus *it exists in a logical sense*, even though nowadays its existence is not actual any more, i.e. – in DOLCE's terminology – *it is no longer present* [g]).

The collection of the Greek hoplites illustrates another different situation. Such a collection had some members centuries ago, thus it is *neither a necessarily empty nor an contingently permanently empty collection*; it *actually existed in the past* but *nowadays it does not actually exist any more*: As in the previous case, we can talk about it (i.e. *it exists in a logical sense*, still *being no longer present*).

As regards the last example, we likely all agree that the collection of the students of that master that will start next year is *not necessarily empty*. Moreover, we can view it either as a *both currently actually existing and empty* collection or as a *not yet actually existing collection that will surely or possibly come into the world in the future*. In both cases we can talk about it within our framework (i.e. *such a collection does exist in a logical sense*); in the former case, the collection would also be *currently present*, while in the latter case it *would not be present, yet*; however it *would be surely or possibly present in the future*.

As already stated above, while it is not mandatory for a collection to have at least two members, it is always possible:

---

[f] For Guizzardi [Gui11], the membership relation is antitransitive (i.e., for any x,y and z, if x is member of y and y is member of z, then x is not member of z). For the sake of generality, we do not impose such a restriction.

[g] It is worth remembering that DOLCE places all past, present and future entities into the domain of quantification, according to its eternalist view. Similarly, it places in it all the possibilia too, according to its possibilist view [MBG$^+$03].

(A5) $Collection(x) \rightarrow$
$$\Diamond(\exists y_1, y_2, t)(y_1 \neq y_2 \wedge hasMember(x, y_1, t) \wedge hasMember(x, y_2, t))$$

"Collection" is a rigid property, in the sense specified in [WG01], i.e. everything that is a collection it is necessarily a collection:

(A6) $Collection(x) \rightarrow \Box Collection(x)$

Rigidity implies that if anything is a collection at a certain time, then it is a collection during its whole lifespan.

3.2.3 Characterizing Collections

Features shared by the members of a collection are important for characterizing the collection itself. Such features can be specified by means of concepts.

Concepts can be organized in a taxonomy, according to the $subConceptOf$ relation. As stated by Definition D2[h], a concept x is a subconcept of a concept y iff necessarily, each time x classifies an entity w, y classifies w, too. For instance, *woman* and *italian-citizen* are both subconcepts of *person*.

(D2) $subConceptOf(x, y) =_{def} dns{:}Concept(x) \wedge dns{:}Concept(y) \wedge$
$$\Box(\forall w, t)(dns{:}classifies(x, w, t) \rightarrow dns{:}classifies(y, w, t))$$

In the following, we introduce a set of predicates that allow us to distinguish those features that are temporarily shared by the members of a collection from those that are permanently shared by them and, among the latter, those that are contingently permanently shared from those that are necessarily shared. Such distinctions are of paramount importance when accounting for historical phenomena.

We say that a collection x has a covering concept y at a certain time t, iff each member w of x at time t possesses the property represented by y, at that time:

(D3) $hasCoveringConcept_t(x, y, t) =_{def}$
$$Collection(x) \wedge dns{:}Concept(y) \wedge dolce{:}TimeInterval(t) \wedge$$
$$dolce{:}presentAt(x, t) \wedge$$
$$(\forall w)(hasMember(x, w, t) \rightarrow dns{:}classifies(y, w, t))$$

The $hasCoveringConcept_t$ predicate allows us to express temporary features shared by the members of a collection. For instance, we can say that *eu-citizen* is a covering concept for the *Romanian citizens* collection since 1st January 2007 until now. In fact, on the 1st January 2007, Romania joined the European Union and thus, from that moment on, every Romanian citizen is (also) an EU citizen. Thus, being an EU citizen is not a permanent property for Romanian citizens, and therefore, before 1st January 2007 *eu-citizen* was not a covering concept for the *Romanian Citizens* collection.

It may happen that the members of a collection permanently (i.e., during the whole lifespan of the collection) share some features. Those cases may be captured by the non temporary version of the predicate (Definition D4):[i]

[h] Definition D2 rephrases the notion of subconcept specified in [MVB+04] within our modal framework.

[i] $hasCoveringConcept(x, y)$ corresponds to the $covers(x, y)$ predicate in [BCGL06]. However, in that work, the first argument of $covers(x, y)$ is restricted to roles (which are specific kinds of concepts), while we admit all concepts (provided that they classify endurants).

(D4) $hasCoveringConcept(x, y) =_{def}$
$Collection(x) \wedge dns{:}Concept(y) \wedge$
$(\forall w, t)(dolce{:}presentAt(x, t) \wedge hasMember(x, w, t) \rightarrow dns{:}classifies(y, w, t))$

For instance, the collection of the members of the 1^{st} Italian Fanfani Government (from 18^{th} January 1954 to 8^{th} February 1954) was covered by the *male* concept, since all of them were males.

Obviously, from the fact that a feature is shared by the members of a collection during the whole lifespan of the collection, we cannot infer that it is a permanent feature for each member. For instance, stating that the *student* concept covers a collection – e.g., a student group during the Protests of 1968 – does not mean that every member of that collection is a student during her whole life: In some sense, covering concepts characterize collections more than their members.

From Definitions D2 and D3, we can derive Theorem T1:

(T1) $hasCoveringConcept_t(x, y, t) \wedge subConceptOf(y, z) \rightarrow hasCoveringConcept_t(x, z, t)$

For instance, given that *eu-citizen* is a covering concept for the *Romanian Citizens* collection since 1st January 2007 until now, and that *eu-citizen* is a subconcept of *person*, we can derive that *person* covers the collection during that same time interval.

From Definitions D2 and D4, we have:

(T2) $hasCoveringConcept(x, y) \wedge subConceptOf(y, z) \rightarrow hasCoveringConcept(x, z)$

For instance, given that *male* is a covering concept for the collect ion of the members of the 1^{st} Italian Fanfani Government and that *male* is a subconcept of *person*, we can derive that *person* covers that collection, too.

It is worth noting that the notion of covering concept does not distinguish between *necessary* and *contingent* classification of the members of a collection. For example, the *male* concept only *contingently* classifies all the members of the 1^{st} Italian Fanfani Government, since, in principle, a female could have been a member of that government (the 1^{st} Italian Scelba Government, that immediately followed it, for instance, had a female member who was Maria Jervolino). Differently, the *person* concept *necessarily* classifies those members, since a member of an Italian Republic government necessarily is a person (and, therefore, *person* covers the collection of the members of the 1^{st} Italian Fanfani Government, too).

Moreover any (contingently) permanently empty collection is covered by any concept, since the antecedent of the implication in the right hand side of Definition D4 is always false. Therefore, the collection of the candidates in the June 2016 elections in San Luca, for instance, is covered by the concepts *candidate-in-June-2016-elections-in-san-luca*, *election-candidate*, *person*, but also by any other concept, such as *worker-at-fiat*, *fisherman*, *film-director*, etc.

To capture necessary (intrinsic) characteristics of collections, we introduce the $hasDescribingConcept(x, y)$ (Definition D5) and $hasDefiningConcept(x, y)$ (Definition D6) predicates. These two predicates do not have a temporal parameter, since they express necessary (and thus, in particular, constant over time) features.

A describing concept y for a collection x is a concept that necessarily classifies the members of the collection during the whole lifespan of the collection:

(D5) $hasDescribingConcept(x, y) =_{def}$
$Collection(x) \wedge dns{:}Concept(y) \wedge$
$\Box(\forall w, t)(dolce{:}presentAt(x, t) \wedge hasMember(x, w, t) \rightarrow dns{:}classifies(y, w, t))$

For instance, both the *romanian-citizen* and the *person* concepts describe the *Romanian citizens* collection.

However, it should be noted that *male*, which – as stated above – is a covering concept for the collection of the members of the 1^{st} Italian Fanfani Government, it is not a describing concept for it, since it only contingently, but not necessarily, classifies its members. Moreover, for the same reason, even though a contingently permanently empty collection is covered by any concept, it is not true that it is also described by any concept. For instance, the *election-candidate* concept describes the collection of candidates in the June 2016 elections in San Luca, but the concepts *worker-at-fiat*, *fisherman* and *film-director* do not describe that collection (although they cover it).

From definitions D4 and D5, we derive that a describing concept is also a covering concept:

(T3) $hasDescribingConcept(x, y) \rightarrow hasCoveringConcept(x, y)$

For instance, given that both the *romanian-citizen* and the *person* concepts describe the *Romanian citizens* collection, we can derive that they also cover it.

From Definitions D2 and D5, we obtain:

(T4) $hasDescribingConcept(x, y) \wedge subconceptOf(y, z) \rightarrow hasDescribingConcept(x, z)$

For example, if *romanian-citizen* is a describing concept for the *Romanian citizens* collection and *romanian-citizen* is a subconcept of the *citizen* concept, then *citizen* is a describing concept for that collection.

A defining concept y for a collection x is a concept that necessarily classifies all and only the members of the collection, during the whole lifespan of the collection:

(D6) $hasDefiningConcept(x, y) =_{def}$
$Collection(x) \wedge dns{:}Concept(y) \wedge$
$\Box(\forall w, t)(dolce{:}presentAt(x, t) \wedge hasMember(x, w, t) \leftrightarrow$
$dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(w, t) \wedge dns{:}classifies(y, w, t))$

For instance, *romanian-citizen* is obviously a defining concept for the *Romanian citizens* collection, while neither *citizen* nor *person* are defining concepts for that collection.

Each collection has a corresponding definining concept:

(A7) $Collection(x) \rightarrow (\exists y)hasDefiningConcept(x, y)$

From Definitions D5 and D6, we can derive that a defining concept is also a describing concept:

(T5) $hasDefiningConcept(x, y) \rightarrow hasDescribingConcept(x, y)$

From Definitions D4 and D6, we derive that a defining concept is also a covering concept:

(T6) $hasDefiningConcept(x, y) \rightarrow hasCoveringConcept(x, y)$

For instance, given that *romanian-citizen* defines the *Romanian citizens* collection, we can derive that it also describes and covers it.

From theorems T4 and T5, we infer that a super-concept of a defining concept is a describing concept (but, in general, not a defining concept):

(T7) $hasDefiningConcept(x, y) \wedge subConceptOf(y, z) \rightarrow hasDescribingConcept(x, z)$

For instance, given that *romanian-citizen* defines the *Romanian citizens* collection, and it is a subconcept of the *citizen* concept, we have that *citizen* describes that collection.

3.2.4 Inclusion Relations

In many cases, it is important to specify that a collection is contained in another (i.e., that the members of a collection also belong to another collection). For example, *Romanian citizens* are now all *EU citizens*; the members of the 1^{st} Italian Fanfani Government were all males (i.e., its members where all members of the collection of males, too); in the Italian Republic, all *senators* (as defined by the Italian Constitution currently in force) are *parliamentarians*. In these cases, we should be able to distinguish temporary from permanent and contingent from necessary inclusion relationships between collections.

We represent possibly contingent inclusion by the predicates $includedIn_t(x, y, t)$ – that contains a temporal parameter (Definition D7) – and $includedIn(x, y)$ – if the inclusion is permanent (Definition D8). We represent necessary inclusion by the predicate $subCollectionOf(x, y)$.

A collection x is included in a collection y, during a time interval t, if and only if during that time both collections are present and the members of x are also members of y:

(D7) $includedIn_t(x, y, t) =_{def}$

$Collection(x) \wedge Collection(y) \wedge dolce{:}TimeInterval(t) \wedge dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(y, t) \wedge$

$(\forall w)(hasMember(x, w, t) \rightarrow hasMember(y, w, t))$

An example of a temporal (contingent) inclusion is the case of the *Romanian citizens* collection, that is included in the *EU citizens* collection since the 1st January 2007 (but before that date it was not).

From Definition D7 we can derive both the reflexivity (Theorem T8) and the transitivity (Theorem T9) of the $includedIn_t(x, y, t)$ relation:

(T8) $Collection(x) \wedge dolce{:}TimeInterval(t) \wedge dolce{:}presentAt(x, t) \rightarrow includedIn_t(x, x, t)$

(T9) $includedIn_t(x, y, t) \wedge includedIn_t(y, z, t) \rightarrow includedIn_t(x, z, t)$

We consider a collection x permanently included in a collection y iff they coexist for a time interval and the members of x are also members of y, when the collections are both present:

(D8) $includedIn(x, y) =_{def}$

$Collection(x) \wedge Collection(y) \wedge (\exists t)(dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(y, t)) \wedge$

$(\forall w, t)(dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(y, t) \wedge hasMember(x, w, t) \rightarrow hasMember(y, w, t))$

According to this definition, the collection of the members of the 1^{st} Italian Fanfani Government is included in the collection of males. Differently, we cannot state that the *Romanian citizens* collection is included in the *EU citizens* collections, since there was a period of time (before 1st January 2007) during which the two collection did coexist, but the members of the former were not members of the latter.

It should be noted that Definition D8 only requires that the two collections coexist for a time interval, while no dependence constraint is specified between them. Therefore, in particular, the included collection may exist also when the including one does not (i.e., before the including collection comes into existence or after it has disappeared). For example, Italy was a founding member of EU, therefore, the *Italian citizens* collection is included in the *EU citizens* one from the moment when the latter came into existence until now (temporary inclusion). If EU (but not Italy) ceased to exist now, we could say that the *Italian citizens* collection is included in the *EU citizens* one (without the need of specifying any temporal restriction). In fact, in that case, the members of the former would be also members of the latter for the whole timespan during which the two collections coexist. However, in that case, the *Italian citizens* collection (i.e., the included one) not only existed before the birth of the *EU citizens* one (i.e., the including collection), but it would still exist after the latter disappeared.

This shows that our notion of inclusion is different from that discussed in [Gui11], who considers any included collection an inseparable part of the including one (i.e., any included collection cannot possibly exist without being part of the including one).

From Definition D8, it immediately follows the reflexivity of the $includedIn$ relation:

(T10) $Collection(x) \wedge (\exists t)(dolce{:}presentAt(x, t)) \rightarrow includedIn(x, x)$

The $includedIn$ relation is not transitive, even among collections that coexist for a time interval.[j]

To have the transitivity, we must add the condition stating that the collection y is always present whenever collections x and z are both present (from Definition D8):

(T11) $(\forall t)(dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(z, t) \rightarrow dolce{:}presentAt(y, t)) \wedge$

$(\exists t)(dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(z, t)) \wedge$

$includedIn(x, y) \wedge includedIn(y, z) \rightarrow$

$includedIn(x, z)$

If two collections coexist for a timespan and the first one is always empty when they are both present, then the first is included in the second one (from Definition D8):

[j]I.e., the following is NOT a theorem:
$(\exists t)(dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(y, t) \wedge dolce{:}presentAt(z, t)) \wedge includedIn(x, y) \wedge includedIn(y, z) \rightarrow includedIn(x, z)$.

(T12) $Collection(x) \wedge Collection(y) \wedge$

$(\exists t)(dolce{:}presentAt(x,t) \wedge dolce{:}presentAt(y,t)) \wedge$

$(\forall w,t)(dolce{:}presentAt(x,t) \wedge dolce{:}presentAt(y,t) \rightarrow$

$\neg hasMember(x,w,t)) \rightarrow$

$includedIn(x,y)$

As a corollary of the latter theorem, we have that each contingently permanently empty collection is included in any collection coexisting with it for some timespan:

(T13) $Collection(x) \wedge Collection(y) \wedge (\exists t)(dolce{:}presentAt(x,t) \wedge$
$dolce{:}presentAt(y,t)) \wedge$

$(\forall w,t)(\neg hasMember(x,w,t)) \rightarrow$

$includedIn(x,y)$

For instance, given Theorem T13, we have that the collection of the candidates in the June 2016 elections in San Luca is included in the collection of the Italian Republic election candidates, but also in any other collection, such as that of fishermen, that of film directors, etc. However, inclusion relations like the latter two, although true, are usually of no interest.

In general, it should be clear that not all the inclusion relations between collections have the same nature and some of them can be of little interest. For instance, it happened that the collection of the members of the 1^{st} Italian Fanfani Government was included in the collection of males. However, as already stated, this was completely contingent, since, in principle, a female could have been member of that government. Differently, in the Italian Republic, the collection of the *senators* is necessarily included in that of the *parliamentarians*.[k] Similarly, the (contingently permanently empty) collection of the candidates in the June 2016 elections in San Luca is necessarily included in the collection of the Italian Republic election candidates, while the former is only contingently included in the collection of fishermen and in that of film directors.

The following definition formalizes the necessary inclusion relation (here called $subCollection$ relation) between collections: A subcollection relation between two collections holds iff they coexist for a time interval and *necessarily* the members of one of them are also members of the other one, when the collections are both present:

(D9) $subCollectionOf(x,y) =_{def}$

$Collection(x) \wedge Collection(y) \wedge (\exists t)(dolce{:}presentAt(x,t) \wedge$
$dolce{:}presentAt(y,t)) \wedge$

$\Box(\forall w,t)(dolce{:}presentAt(x,t) \wedge dolce{:}presentAt(y,t) \wedge hasMember(x,w,t) \rightarrow$
$hasMember(y,w,t))$

[k] Here we refer to these two collections as they are defined in the current Italian Constitution. Thus we are not stating that necessarily *senators* are also *parliamentarians*, but only that, considering those two specific (socially created) collections, one of them is necessarily included in the other in force of their definitions, contained in the current Italian Constitution. Should the Constitution be changed and some senators outside the Parliament be admitted, then two new and different collections of senators and parliamentarians would be defined (and created) for which the inclusion relation would not necessarily hold.

We thus have that the collection of the Italian Republic *senators* is a subcollection of the *parliamentarians*. Nevertheless, the collection of the members of the 1^{st} Italian Fanfani Government is not a subcollection of the collection of males. Moreover, the collection of the candidates in the June 2016 elections in San Luca is a subcollection of the collection of the Italian Republic election candidates, but it is neither a subcollection of fishermen or of film directors.

From the definition above, it immediately follows the reflexivity of the $subCollectionOf$ relation:

(T14) $Collection(x) \wedge (\exists t)(dolce{:}presentAt(x,t)) \rightarrow subCollectionOf(x,x)$

The $subCollectionOf$ relation does not express a necessary relation between collections.[l]

As expected, from Definitions D8 and D9, it follows that $subCollectionOf(x,y)$ expresses a stronger relation between collections than $includedIn(x,y)$:

(T15) $subCollectionOf(x,y) \rightarrow includedIn(x,y)$

As for $includedIn$, the transitivity of $subCollectionOf$ holds only in specific circumstances, i.e. when x and z coexist for a timespan and necessarily, the collection y is always present whenever x and z are both present (from Definition D9):

(T16) $(\Box(\forall t)(dolce{:}presentAt(x,t) \wedge dolce{:}presentAt(z,t) \rightarrow dolce{:}presentAt(y,t))) \wedge$
$(\exists t)(dolce{:}presentAt(x,t) \wedge dolce{:}presentAt(z,t)) \wedge$
$subCollectionOf(x,y) \wedge subCollectionOf(y,z) \rightarrow$
$subCollectionOf(x,z)$

Differently from what holds for the $includedIn$ relation, and consistently with our intuition, it is not true that each contingently permanently empty collection is a subcollection of any collection coexisting with it for some timespan.[m]

Thus, for example, it is not true that the collection of the candidates in the June 2016 elections in San Luca is a subcollection of the collection of fishermen or of that of film directors.

3.2.5 Deriving Characterizations from Inclusions (and vice versa)

In this section, we discuss how the relations for characterizing collections – $hasCoveringConcept_t(x,y,t)$, $hasCoveringConcept(x,y)$, $hasDescribingConcept(x,y)$ and $hasDefiningConcept(x,y)$ –

stand w.r.t. the inclusion relations – $includedIn_t(x,y,t)$, $includedIn(x,y)$ and $subCollectionOf$. In particular, we formulate several theorems that allow us to derive a characterization for collections from inclusion relations and vice versa. For the sake of brevity, some theorems (and formulas) are grouped into theorem (and formula) schemata.

[l] I.e., the following is NOT a theorem:
$subCollectionOf(x,y) \rightarrow \Box subCollectionOf(x,y)$.

[m] I.e., the following is NOT a theorem:
$Collection(x) \wedge Collection(y) \wedge (\exists t)(dolce{:}presentAt(x,t) \wedge dolce{:}presentAt(y,t)) \wedge$
$(\forall w,t)(\neg hasMember(x,w,t)) \rightarrow subCollectionOf(x,y)$.

First of all, from Definitions D3 and D7, we can derive that any concept y that covers a collection x during a time interval t, also covers any collection z included in x during that same time interval:

(T17) $hasCoveringConcept_t(x, y, t) \wedge includedIn_t(z, x, t) \rightarrow hasCoveringConcept_t(z, y, t)$

For instance, given that today the *Austrian citizens* collection is covered by the *eu-citizen* concept and that the *Carinthian citizens* collection is included in that of *Austrian citizens*, we can derive that today the *Carinthian citizens* collection is covered by the *eu-citizen* concept, too.

A similar theorem cannot be derived if we substitute either $includedIn(z, x)$ or $subCollectionOf(z, x)$ for $includedIn_t(z, x, t)$.[n] Moreover, no theorem similar to Theorem T17 holds for any combination of the (no temporary) characterizing and inclusion relations.[o]

We can however prove a set of (weaker) properties, that we group in Theorem Schema TS1, which states that if y is a covering, describing or defining concept for a collection x, and z is either included in or a subcollection of x, then y classifies all the members of z for the whole timespan in which the two collections coexist:

(TS1) $\alpha(x, y) \wedge \beta(z, x) \wedge dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(z, t) \wedge hasMember(z, w, t) \rightarrow$

$dns{:}classifies(y, w, t)$,

where $\alpha \in \{hasCoveringConcept, hasDescribingConcept, hasDefiningConcept\}$ and $\beta \in \{includedIn, subCollectionOf\}$.

Given Definitions D4 and D8, it is easy to prove the instance of Theorem Schema TS1 where
$\alpha \equiv hasCoveringConcept$ and $\beta \equiv includedIn$. Then, given Theorems T3, T5, T6 and T15, it follows that every instance of Theorem Schema TS1 is a theorem.

For example, given that the *italian-citizen* concept describes the *Italian Republic parliamentarians* collection and that the *Italian Republic senators* collection is a subcollection of the *Italian Republic parliamentarians* collection, then the *italian-citizen* concept classifies all the Italian Republic senators for the whole timespan in which the two collections coexist.

If it happens that a collection z is always present when a collection x is present, then we can derive a covering relation between z and a concept y, given any combination of the no temporary characterizing and inclusion relations, as follows:

(TS2) $\alpha(x, y) \wedge \beta(z, x) \wedge (\forall t)(dolce{:}presentAt(z, t) \rightarrow dolce{:}presentAt(x, t)) \rightarrow$

$hasCoveringConcept(z, y)$,

[n] I.e., no instance of the following schema is a theorem: $hasCoveringConcept_t(x, y, t) \wedge \beta(z, x) \rightarrow hasCoveringConcept_t(z, y, t)$, where $\beta \in \{includedIn, subCollectionOf\}$. It is easy to verify that the instance where $\beta \equiv subCollectionOf$ is not a theorem, thus, for Theorem T15, the other instance is not a theorem either.

[o] I.e., no instance of the following schema is a theorem:
$\alpha(x, y) \wedge \beta(z, x) \rightarrow \gamma(z, y)$,
where $\alpha, \gamma \in \{hasCoveringConcept, hasDescribingConcept, hasDefiningConcept\}$ and $\beta \in \{includedIn, subCollectionOf\}$. It is easy to verify that the instance where $\alpha \equiv hasDefiningConcept$, $\beta \equiv subCollectionOf$ and $\gamma \equiv hasCoveringConcept$ is not a theorem. Then, given Theorems T3, T5, T6 and T15, we can prove that no instance is a theorem.

where $\alpha \in \{hasCoveringConcept, hasDescribingConcept, hasDefiningConcept\}$ and $\beta \in \{includedIn, subCollectionOf\}$.

From Definitions D4 and D8, it is easy to prove the instance of Theorem Schema TS2 where
$\alpha \equiv hasCoveringConcept$ and $\beta \equiv includedIn$. Then, given Theorems T3, T5, T6 and T15, it follows that every instance of TS2 is a theorem.

However, from that same premises we cannot derive anything stronger than coverage.[p]

To be able to derive that a concept not only covers a collection, but it also actually describes it, we need to strengthen the premises, as follows:

(TS3) $\alpha(x, y) \land subcollectionOf(z, x) \land \Box(\forall t)(dolce{:}presentAt(z, t) \rightarrow dolce{:}presentAt(x, t)) \rightarrow$

$hasDescribingConcept(z, y)$,

where $\alpha \in \{hasDescribingConcept, hasDefiningConcept\}$.

Given Definitions D5 and D9, it is easy to prove the instance of Theorem Schema TS3 where
$\alpha \equiv hasDescribingConcept$. Then, given Theorem T5, it follows that the other instance of TS3 (where $\alpha \equiv hasDefiningConcept$) is a theorem, too. It is also worth noting that the antecedent of Theorem Schema TS3 specifies, in particular, an existential dependence for the collection z on the collection x.

For example, given that the *Italian Republic senators* collection is a subcollection of that of *Italian Republic parliamentarians* – which is both defined by the *italian-republic-parliamentarian* and described by the *italian-citizen* concepts – and that the former collection does existentially depend on the latter, we can derive that both concepts describe the *Italian Republic senators* collection.

Given that in Modal Logic S5 $\Box\psi \rightarrow \psi$ for any well-formed formula ψ, from Theorem Schema TS2, it immediately follows Theorem Schema TS4, stating that if, necessarily, a collection z is always present when a collection x is present, then we can derive a covering relation between z and a concept y, given any combination of the no temporary characterizing (α) and inclusion (β) relations:

(TS4) $\alpha(x, y) \land \beta(z, x) \land \Box(\forall t)(dolce{:}presentAt(z, t) \rightarrow dolce{:}presentAt(x, t)) \rightarrow$

$hasCoveringConcept(z, y)$,

where $\alpha \in \{hasCoveringConcept, hasDescribingConcept, hasDefiningConcept\}$ and $\beta \in \{includedIn, subCollectionOf\}$.

No combination of the characterizing and inclusion relations in schemata like TS2, TS3 and TS4 above allows us to derive that any concept defines a collection.[q]

[p] I.e., no instance of the following schema is a theorem:
$\alpha(x, y) \land \beta(z, x) \land (\forall t)(dolce{:}presentAt(z, t) \rightarrow dolce{:}presentAt(x, t)) \rightarrow \gamma(z, y)$,
where $\alpha \in \{hasCoveringConcept, hasDescribingConcept, hasDefiningConcept\}$, $\beta \in \{includedIn, subCollectionOf\}$ and $\gamma \in \{hasDescribingConcept, hasDefiningConcept\}$. It is easy to verify that the instance where $\alpha \equiv hasDefiningConcept$, $\beta \equiv subCollectionOf$ and $\gamma \equiv hasDescribingConcept$ is not a theorem. Then, given Theorems T3, T5, T6 and T15, we can prove that no instance is a theorem.

[q] I.e., no instance of the following schema is a theorem:

Furthermore, except for the instances of the Theorem Schema TS3, no other combinations of the considered predicates allows us to derive $hasDescribingConcept(z, y)$.[r]

In many cases, given a taxonomy of concepts, from the characterization of collections, we can derive some inclusion relationships between them. First of all, given two collections x and z that coexist for a time interval, if x is either described or defined by a concept y_1, which is a subconcept of a concept y_2 defining z, then x is a subcollection of z:

(TS5) $\alpha(x, y_1) \wedge hasDefiningConcept(z, y_2) \wedge subConceptOf(y_1, y_2) \wedge$
$(\exists t)(dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(z, t)) \rightarrow$
$subCollectionOf(x, z)$,
where $\alpha \in \{hasDescribingConcept, hasDefiningConcept\}$.

From Definitions D2, D5, D6 and D9, we can prove the instance of Theorem Schema TS5 where $\alpha \equiv hasDescribingConcept$. Then, given Theorem T5, it follows that the other instance of TS5 (where $\alpha \equiv hasDefiningConcept$) is a theorem, too.

For example, let us consider the two collections of the *Italian Garibaldian partisans* and of the *Italian partisans* (which, of course, coexisted for a time period): If we know that the former is described by the *italian-communist-partisan* concept, which is a subconcept of *italian-partisan*, a concept that defines the latter, then we can derive that *Italian Garibaldian partisans* is a subcollection of *Italian partisans*.

It is worth noting that, in Theorem Schema TS5, to derive the subcollection relationship between x and z, the concept y_1 should at least describe the collection x (thus, a covering relation between x and y_1 is not enough).[s] However, if y_1 only covers x, we can derive the weaker inclusion relationship $includedIn(x, z)$. Such a situation is represented by an instance of the following theorem schema:

(TS6) $\alpha(x, y_1) \wedge hasDefiningConcept(z, y_2) \wedge subconceptOf(y_1, y_2) \wedge$
$(\exists t)(dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(z, t)) \rightarrow$
$includedIn(x, z)$,

$\alpha(x, y) \wedge \beta(z, x) \wedge \Box(\forall t)(dolce{:}presentAt(z, t) \rightarrow dolce{:}presentAt(x, t)) \rightarrow hasDefiningConcept(z, y)$,

where $\alpha \in \{hasCoveringConcept, hasDescribingConcept, hasDefiningConcept\}$ and $\beta \in \{includedIn, subCollectionOf\}$. It is easy to verify that the instance where $\alpha \equiv hasDefiningConcept$ and $\beta \equiv subCollectionOf$ is not a theorem. Then, given Theorems T3, T5, T6 and T15, we can prove that no instance is a theorem.

[r] I.e., no instance of the following schema is a theorem:

$\alpha(x, y) \wedge includedIn(z, x) \wedge \Box(\forall t)(dolce{:}presentAt(z, t) \rightarrow dolce{:}presentAt(x, t)) \rightarrow hasDescribingConcept(z, y)$,

where $\alpha \in \{hasCoveringConcept, hasDescribingConcept, hasDefiningConcept\}$. It is easy to verify that the instance where $\alpha \equiv hasDefiningConcept$ is not a theorem. Then, given Theorems T5 and T6, we can prove that no instance is a theorem.

Moreover, it is easy to verify that the following is not a theorem, either:

$hasCoveringConcept(x, y) \wedge subCollectionOf(z, x) \wedge \Box(\forall t)(dolce{:}presentAt(z, t) \rightarrow dolce{:}presentAt(x, t)) \rightarrow hasDescribingConcept(z, y)$.

[s] I.e., the following is not a theorem:

$hasCoveringConcept(x, y_1) \wedge hasDefiningConcept(z, y_2) \wedge subConceptOf(y_1, y_2) \wedge (\exists t)(dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(z, t)) \rightarrow subCollectionOf(x, z)$.

where $\alpha \in \{hasCoveringConcept, hasDescribingConcept,$
$hasDefiningConcept\}$.

From Definitions D2, D4, D6 and D8, we can prove the instance of Theorem Schema TS6 where $\alpha \equiv hasCoveringConcept$. Then, from Theorem Schema TS5 and Theorems T5 and T15, it follows that every instance of TS6 is a theorem.

It should be noted that there is no other way to derive any inclusion relationship, given a taxonomy of concepts and a characterization of collections by means of concepts.[t]

3.2.6 Emerging vs Constructed Collections

Among collections, we distinguish between "emerging" and "constructed" collections (Definitions D11 and D12 below, respectively), where the former can be conceptualized, but are independent from any conceptualization, while the latter are always conceptualized and depend on their conceptualization.

We represent conceptualizations by means of descriptions (intended as $dns{:}Descriptions$). A description that describes a collection uses at least one describing concept:

(A8) $describesCollection(x, y) \to dns{:}Description(x) \wedge Collection(y) \wedge$
$(\exists z)(Concept(z) \wedge dns{:}uses(x, z) \wedge hasDescribingConcept(y, z))$

Similarly, a description that defines a collection uses at least one defining concept:

(A9) $definesCollection(x, y) \to dns{:}Description(x) \wedge Collection(y) \wedge$
$(\exists z)(Concept(z) \wedge dns{:}uses(x, z) \wedge hasDefiningConcept(y, z))$

From Theorem T5 and the two latter axioms (A8 and A9), we obtain:

(T18) $definesCollection(x, y) \to describesCollection(x, y)$

A description provides a conceptualization for a collection if it either describes or defines it:

(D10) $conceptualizesCollection(x, y) =_{def}$
$describesCollection(x, y) \vee definesCollection(x, y)$

For instance, the Constitution of the Italian Republic specifies a description providing a conceptualization for the collection of the *Italian Republic parliamentarians* and official documents of ISTAT (the Italian National Institute of Statistics) contain a conceptualization for the collection of the *Italian unemployed people*.

Formally, an emerging collection is a collection that has at least one member at some point and that it does not specifically depends on any of its conceptualizations:

[t]I.e., no instance of the following schema is a theorem:
$\alpha(x, y_1) \wedge \beta(z, y_2) \wedge subConceptOf(y_1, y_2) \wedge (\exists t)(dolce{:}presentAt(x, t) \wedge dolce{:}presentAt(z, t)) \to \gamma(x, z)$,
where $\alpha \in \{hasCoveringConcept, hasDescribingConcept, hasDefiningConcept\}$, $\beta \in \{hasCoveringConcept, hasDescribingConcept\}$, $\gamma \in \{includedIn, subCollectionOf\}$. It is easy to verify that the instance where $\alpha \equiv hasDefiningConcept$, $\beta \equiv hasDescribingConcept$ and $\gamma \equiv includedIn$ is not a theorem. Then, given Theorems T3, T5, T6 andT15, we can prove that no instance is a theorem.

(D11) $EmergingCollection(x) =_{def}$
$Collection(x) \wedge (\exists y, t)(hasMember(x, y, t)) \wedge$
$(\forall z)(conceptualizesCollection(z, x) \rightarrow \neg dolce{:}specificallyConstantlyDependsOn(x, z))$

For instance, *consumers*, *working class*, *Italian unemployed people* are all emerging collections.

In many cases, an emerging collection is conceptualized a posteriori, i.e. first the collection emerges, then its conceptualization is specified. For example, this is true for many (emerging) collections that are of interest for sociologists and historians: e.g., first *consumers*, *working class*, *Italian unemployed people*, etc. emerge (and participate in events), then they are conceptualized. Emerging collections may also be conceptualized for the first time when they no longer exist, for instance the collection of *dinosaurs*, that lived on Earth several millions years ago. In principle, an emerging collection can also be conceptualized before it comes into existence: for example, each time a war breaks out, we can easily envisage the respective collection of *maimed war veterans*.[u]

Differently, a constructed collection is a collection that specifically depends on a description that conceptualizes it:

(D12) $ConstructedCollection(x) =_{def}$
$Collection(x) \wedge (\exists z)(conceptualizesCollection(z, x) \wedge$
$dolce{:}specificallyConstantlyDependsOn(x, z))$

The main distinguishing feature of a constructed collection is the fact that it specifically depends on its conceptualization. This implies that its first conceptualization must be conceived before (or at the same time as) the collection comes into existence.[v] Moreover, since a description generically depends on some intentional agents [GM03, MVB$^+$04, GBCL05], we can derive that a constructed collection (indirectly) depends on some intentional agents, too.

Formally, constructed collections are not only non-physical endurants (see Definition D1), but, more specifically, they are social objects, as are all collections in the framework presented in [BCGL06]):

(A10) $ConstructedCollection(x) \rightarrow dolce{:}SocialObject(x)$

We should all agree that *Italian citizens*, *Italian citizens with right to vote*, *Italian ordinary statute regions*, *Laws made by the Italian Republican Parliament*, *Italian Republic parliamentarians*, *Members of the* 1^{st} *Italian Fanfani Government* and *Workers at FIAT* are all examples of constructed collections, as far as we agree that they depend on descriptions that conceptualize them (possibly contained in legal systems, norms, agreements, etc.).

[u] It should be clear that in the latter case too, the conceptualization process only produce in advance a description for a possibly future collection, but it is not that process that actually generates the collection.

[v] The specific constant dependence of a constructed collection on its conceptualization does not mean that the existence of the conceptualization implies the existence of the corresponding collection. For instance, the conceptualization of the collection of the *German Democratic Republic citizens* is still present (we still have copies of the German Democratic Republic Constitution and we talk about it), but nowadays the collection of *German Democratic Republic citizens* is no longer present. For a constructed collection to be present, not only the conceptualization it specifically depends on must be present too, but the latter must also be somehow "in force". An ontological analysis of the status of "conceptualization in force", although interesting, is beyond the scope of the present paper.

It is worth noting that the distinction between emerging and constructed collections is not intended to mirror any sort of dichotomy “natural” vs “artificial” or “non human-created” vs “human-created” collections. In particular, if it is true that we likely classifies as $EmergingCollection$ any collection that we consider either “natural” or “non human-created” (such as the *dinosaurs* or a *swarm of locusts*), it is nevertheless true that the category of $EmergingCollection$s may encompass also some collections that we usually classify as “artificial” or “human-created”, as far as we consider them as not depending on any describing conceptualization. For instance, (we probably all agree that) almost all the above listed examples of emerging collections (*consumers, working class*, etc.) are the results of human choices and activities and thus they may be considered as “artificial” or, at least, as “human-created”; nevertheless (we probably also all agree that) these collections do not depend on any purposely created conceptualization, and this justifies their attribution to the category of $EmergingCollection$s.

Each collection is either an emerging or a constructed collection (Axiom A11), but it cannot be both (Theorem T19, from Definitions D11 and D12):

(A11) $Collection(x) \rightarrow EmergingCollection(x) \vee ConstructedCollection(x)$

(T19) $ConstructedCollection(x) \rightarrow \neg EmergingCollection(x)$

From Definition D11, Axiom A11, and Theorem T19, we can derive that any contingently permanently empty collection can never emerge and thus that such a collection necessarily is a constructed collection:

(T20) $Collection(x) \wedge \neg(\exists y, t)(hasMember(x, y, t)) \rightarrow$
$\neg EmergingCollection(x) \wedge ConstructedCollection(x)$

For instance, necessarily, the contingently permanently empty collection of *Candidates in the June 2016 elections in San Luca* is a constructed collection (and, thus, not an emerging one).

4 Conclusions

In this paper, we presented the results of an ontological analysis of collective entities (*collections*), together with a formalization, in modal logics, of their features and the relations they are involved in. Such an ontological analysis, besides contributing to clarify the notion of collection, also represents an essential step towards the definition of a formal conceptualization supporting ontology-driven applications.

As far as the ontological modeling is concerned, some issues, closely related to the semantic characterization of collections, deserve further study. First of all, the relation between collections and organizations [BF09], should be analyzed. Moreover, the relation between the participation modalities of endurants in events, that can be represented by thematic roles [GMR18, GMR17], and the collections requires an accurate analysis.

From the application point of view, the results discussed in the present paper were used as a reference ontology that guided the specification of an OWL 2 characterization of collections used within two projects in the historical domain. Obviously, the OWL 2 ontology of collections, though consistent with the one discussed here, is a lighter version of it, given that modalities, as well as some first-order logic features, are not expressible in OWL 2. Another fundamental step for the effective usage of the OWL 2 version of the ontology is the mapping between its elements and the most influencing semantic models used in the Cultural Heritage domain (see Section 2.2) in order to guarantee interoperability, especially in a Linked Data perspective [HB11].

References

[All83] J.F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983.

[BCGL06] E. Bottazzi, C. Catenacci, A. Gangemi, and J. Lehmann. From collective intentionality to intentional collectives: an ontological perspective. *Cognitive Systems Research - Special Issue on Cognition Joint Action and Collective Intentionality*, 7(2-3):192–208, 2006.

[BDOS15] P. Le Boeuf, M. Doerr, C. E. Ore, and S. Stead. *Definition of the CIDOC Conceptual Reference Model, v. 6.2.1.* ICOM/CIDOC CRM SIG, 2015.

[BF09] E. Bottazzi and R. Ferrario. Preliminaries to a dolce ontology of organizations. *International Journal of Business Process Integration and Management*, 4(4):225–238, 2009.

[BM09] S. Borgo and C. Masolo. Foundational choices in dolce. In S. Staab and R. Studer, editors, *Handbook on Ontologies, Second Edition*, pages 361–381. Springer, Berlin, Heidelberg, 2009.

[BP01] V. Borschev and B. H. Partee. Genitive modifiers, sorts, and metonymy. *Nordic Journal of Linguistics*, 24(2):140–160, 2001.

[BvBW07] P. Blackburn, J. van Benthem, and F. Wolter, editors. *Handbook of Modal Logic*. Elsevier, 2007.

[CGM17] M. Caserio, A. Goy, and D. Magro. Smart access to historical archives based on rich semantic metadata. In *Proc. KMIS 2017*, pages 93–100. SciTePress, 2017.

[CH14] E. Bailey C. M. Steiner E. C. Hillemann O. Conlan C. Hampson, S. Lawless. Metadata-enhanced exploration of digital cultural collections. *Int. J. Metadata, Semantics and Ontologies*, 10(1):155 – 167, 2014.

[dBOI+14] V. de Boer, J. Oomen, O. Inel, L. Aroyo, E. van Staveren, W. Helmich, and D. de Beurs. Dive into the event-based browsing of linked historical media. In *Proc. 13th Int. Semantic Web Conference (ISWC 2014)*, 2014.

[Doe03] M. Doerr. The cidoc conceptual reference model: An ontological approach to semantic interoperability of metadata. *AI Magazine*, 24(3):75–92, 2003.

[Eur16] Europeana. *EDM Definition of the Europeana Data Model v.5.2.7.* 2016.

[Fre84] P. A. French. *Collective and corporate responsibility*. Columbia University Press, 1984.

[Gal12] A. Galton. How is a collection related to its members? In *Proc. InterOntology 2010*. Society for the Study of Artificial Intelligence and Simulation of Behaviour, 2012.

[Gal14] A. Galton. On generically dependent entities. *Applied Ontology*, 9(2):129–153, 2014.

[GBCL05] A. Gangemi, S. Borgo, C. Catenacci, and J. Lehmann. *Task Taxonomies for Knowledge Content*. 2005. Metokis Deliverable D07.

[GDL+17] A. Goy, R. Damiano, F. Loreto, D. Magro, S. Musso, D. Radicioni, C. Accornero, D. Colla, A. Lieto, E. Mensa, M. Rovera, D. Astrologo, B. Boniolo, and M. D'ambrosio. Prismha (providing rich semantic metadata for historical archives). In *Proc. Contextual Representation of Objects and Events in Language (CREOL 2017)*, 2017.

[GM03] A. Gangemi and P. Mika. Understanding the semantic web through descriptions and situations. *LNCS*, 2888:689–706, 2003.

[GM15] A. Goy and D. Magro. What are ontologies useful for? In M. Khosrow-Pour, editor, *Encyclopedia of Information Science and Technology, Third Edition*, pages 7456–7464. IGI Global, 2015.

[GMR15] A. Goy, D. Magro, and M. Rovera. Ontologies and historical archives: A way to tell new stories. *Applied Ontology*, 10(3-4):331–338, 2015.

[GMR17] A. Goy, D. Magro, and M. Rovera. An ontological perspective on thematic roles. *LNAI*, 10180:123–126, 2017.

[GMR18] A. Goy, D. Magro, and M. Rovera. On the role of thematic roles in a historical event ontology. *Applied Ontology*, 13:19–39, 2018.

[Gui11] G. Guizzardi. Ontological foundations for conceptual part-whole relations: The case of collectives and their parts. In *Proc. Twenty-third International Conference on Advanced Information System Engineering (CAiSE 2011)*, 2011.

[GW16] A. Galton and Z. Wood. Extensional and intensional collectives and the *de re/de dicto* distinction. *Applied Ontology*, 11(3):205–226, 2016.

[HB11] T. Heath and C. Bizer. *Linked Data: Evolving the Web into a Global Data Space*. Morgan & Claypool, 2011.

[HLTM12] E. Hyvönen, T. Lindquist, J. Törnroos, and E. MÄkelä. History on the semantic web as linked data - an event gazetteer and timeline for the world war i. In *Proc. of CIDOC 2012 Enriching Cultural Heritage*, 2012.

[LH01] C. Lagoze and J. Hunter. The abc ontology and model. In *Proc. Dublin Core Metadata Initiative (DCMI 2001)*, pages 160–176, 2001.

[MBG+03] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, and A. Oltramari. *WonderWeb deliverable D18*. ISTC-CNR, 2003. Technical Report.

[MVB+04] C. Masolo, L. Vieu, E. Bottazzi, C. Catenacci, R. Ferrario, A. Gangemi, and N. Guarino. Social roles and their descriptions. In *Proc. KR 2004*, pages 267–277, 2004.

[NPD17] F. Nanni, S. P. Ponzetto, and L. Dietz. Building entity-centric event collections. In *Proc. ACM/IEEE Joint Conference on Digital Libraries (JCDL 2017)*, 2017.

[Obe05] D. Oberle. *Semantic management of middleware*. Springer, 2005.

[PS14] Ciccarese P. and Peroni S. The collections ontology: creating and handling collections in owl 2 dl frameworks. *Semantic Web*, 5(6):515–529, 2014.

[RA07] Y. Raymond and S. Abdallah. *The Event Ontology*. 2007.

[RNPG17] M. Rovera, F. Nanni, S. P. Ponzetto, and A. Goy. Domain-specific named entity disambiguation in historical memoirs. In *Proc. Fourth Italian Conference on Computational Linguistics (CLIC-IT 2017)*, 2017.

[Rov16] M. Rovera. A knowledge-based framework for events representation and reuse from historical archives. *LNCS*, 9678:845–852, 2016.

[Sea95] J. R. Searle. *The Construction of Social Reality*. Free Press, 1995.

[SFSS09] A. Scherp, T. Franz, C. Saathoff, and F. Staab. F - a model of events based on the foundational ontology dolce+dnsultralite. In *Proc. 5th International Conference on Knowledge Capture (K-CAP 2009)*, pages 137–144. ACM, 2009.

[SS09] S. Staab and R. Studer, editors. *Handbook on Ontologies, Second Edition*. Springer, 2009.

[ST17] R. Sprugnoli and S. Tonelli. One, no one and one hundred thousand events: Defining and processing events in an inter-disciplinary perspective. *Natural Language Engineering*, 23(4):485–506, 2017.

[STH14] R. Shaw, R. Troncy, and L. Hardman. Lode: Linking open descriptions of events. In *Proc. 4th Asian Conf. on The Semantic Web (ASWC 2009)*, pages 153–167, 2014.

[vdAAC+10] C. van den Akker, L. Aroyo, A. Cybulska, M. van Erp, P. Gorgels, L. Hollink, C. Jager, L. van der Meij, J. Oomen, J. van Ossenbruggen, G. Schreiber, R. Segers, P. Vossen, and B. Wielinga. Historical event-based access to museum collections. In *Proc. 1st Int. Workshop on Recognising and Tracking Events on the Web and in Real Life (EVENTS 2010)*, 2010.

[vHS+11] W. R. van Hage, V. Malaisé , R. Segers, L. Hollink, and G. Schreiber. Desing and use of the simple event model (sem). *Journal of Web Semantics*, 9(2):128–136, 2011.

[WG01] C. Welty and N. Guarino. Supporting ontological analysis of taxonomic relationships. *Data and Knowledge Engineering*, 31(1):51–74, 2001.

[WG09] Z. Wood and A. Galton. A taxonomy of collective phenomena. *Applied Ontology*, 4(3-4):267–292, 2009.

[Zar15] G. P. Zarri. A structured and in-depth representation of the semantic content of elementary and complex events. *Int. J. Metadata, Semantics and Ontologies*, 10(1):12–27, 2015.