Ontologies for the Digital Humanities: Learning from the Life Sciences?

Ludger JANSEN[a,1]
[a] *University of Rostock* and *Ruhr University Bochum*

Abstract. Ontologies are more and more used in the Digital Humanities, although still in an unsystematic and uncoordinated fashion. In contrast, ontology development in the life sciences has become a highly coordinated community effort, with emerging tools, standards, guidelines and good practice rules. The paper argues that DH ontologies can benefit from these or similar standards. For this purpose, the paper first discusses two case studies, namely the treatment of social entities in the National Cancer Institute Thesaurus (NCIT) and a toy ontology from a DH textbook. It then presents a number of objections to the claim, which are discussed and dismissed. Learning from the life sciences, ontology development in DH should thus be envisioned as a coordinated, principle-guided community effort.

Keywords. Digital Humanities, Applied Ontology, Social Ontology, Culture, Guidelines, NCIT

1. Introduction

Human culture and history are complex affairs. Digital methods allow for processing massive amounts of data, traditionally called 'source material'. For this purpose, digital methods should be able to integrate data from diverse sources, and cope with the highly complex domain. These challenges are parallel to the challenges posed to bioinformatics. Biological life is similarly complex, and digital methods in the life sciences also have to cope with a complex domain and with data from diverse sources.

There are various kinds of artefacts that have been developed to organize data. One strategy is to annotate data with terms from controlled vocabularies, collections of subject headings, or thesauri. The most expressive method on offer is the use of formal ontologies, where terms are not only standardized, but also rigorously characterized by means of logical formulae. The idea behind this is that such a standardization and formal characterization answers several desiderata at once. First, by fixing not only the term but also its meaning through formal characterization, information retrieval no longer has to rely on the search for a string – which might be ambiguous or only one of several synonyms used for a certain thing, thus impairing recall and precision of searches. Having fixed the meaning, however, allows semantically disambiguated searching – or, to put it in a slogan, searching for things instead of strings. In addition, a formal characterization in a computer-processable logical dialect (like, e.g., description logic),

[1] Corresponding Author, Ludger Jansen, Institute of Philosophy, University of Rostock, 18051 Rostock, Germany; ludger.jansen@uni-rostock.de.

allows the use of automatic reasoning programs. This requires logically consistent data – for any inconsistency will allow inferring any statement, including any false statement. For this reason, rigorous semantics and their enforcement are very much called for, as well as formally characterized relations between terms. If done well, ontologies allow for information integration between different databases or across different versions of the same database.

In this paper, I want to point to various aspects where ontology development in the Digital Humanities can learn from the achievements reached in the domain of the life sciences. For this purpose, I start with the state of the art of ontology development in the life sciences (Section 2). Then I analyze the use of ontologies for the social domain, first by in the Digital Humanities in general (Section 3), then by means of two case studies. First, I analyze the treatment of social entities in the National Cancer Institute Thesaurus (Section 4); second, I discuss a toy ontology from a Digital Humanities textbook (Section 5). Finally, I discuss and dismiss several possible objections against the transfer of ontology development guidelines from the life sciences to the social domain (Section 6). I conclude by pointing out various aspects where ontology development for the social domain can learn from the life sciences (Section 7).

2. Ontologies in the Life Sciences

In the last two decades, much of the incentive to further develop applied ontology came from the biomedical domain. The biomedical domain does not only feature a huge diversity of life forms, but also deals with highly complex systems and subsystems, like organisms and cells, and complex biological processes as well as functional aspects involved in, e.g., genetics and metabolism. This has turned important parts of biology into data science [24], very much relying on computer-based data representation and processing. Combined with the desire to use all available knowledge to develop cures against continuing medical threats, it led to the development of tools supporting the representation and processing of biomedical data, beginning with rather conventional tools like the Medical Subject Headings (MeSH) and statistical classifications like the International Classification of Diseases (ICD) to computer-based terminologies like the Systematized Nomenclature of Medicine (SNOMED).

The life sciences developed rich tools for supporting the use of ontologies in their domain. An important initiative in this field is the Open Biological and Biomedical Ontologies (OBO) Foundry. The OBO Foundry operates a repository for ontologies for the biological or biomedical domain (obofoundry.org) [40]. The work of the OBO Foundry is based on a set of good practice rules that are used for the development of a number of consistent and orthogonal community-based domain ontologies for various biomedical domains that are interoperable with each other. These ontologies are represented in data artefacts that are freely available in the public domain, and that can be used singly or in combination under the umbrella of the Basic Formal Ontology (BFO) as their common top-level ontology [1].

In addition, the biomedical ontology community provides community tools for developers, peer review, and good-practice rules for ontology development. Upper-domain ontologies like BioTop (not itself part of the OBO Foundry world; Schulz, Boeker & Martinez-Costa 2018), or ontology development guidelines like the Good Ontology Design guidelines (GoodOD) [34], support ontology developers.

One of the leading ideas that shaped ontology development in the life sciences is to move away from *ad hoc* application ontologies to principled reference ontologies, which are then used as a basis for small-scale application ontologies. Many biomedical ontologies are collected in the repository of the OBO Foundry. Reference ontologies are committed to a rigorous semantics not only of their terms, but also of the relations they use. In order to guarantee the semantic interoperability of the modular reference ontologies, a common set of formal relations has been developed, the Relation Ontology (RO), which contains formal relations like **part of** or **has participant** [39].

Reference ontologies are organized in a modular way. They should be orthogonal to each other, i.e., they should be integrated, but have no overlap. Each reference ontology should have its peculiar domain, and no class should be dealt with in more than one ontology. Instead, relevant terms from other domains should be re-used and integrated. For this purpose, the MIREOT standard has been developed, which specifies the 'Minimum Information to Reference an External Ontology Term' [6]. In order to integrate terms, the tool OntoFox can be used [46]. In order to browse these ontologies, or to find appropriate classes for re-use, there are tools like OntoBee (ontobee.org) or the Ontology Lookup Service (OLS; maintained by EBI at www.ebi.ac.uk/ols). By these means, reference ontologies are, on the one hand, designed in a modularized way, but, on the other hand, interconnected through the import of terms from other ontologies. One crucial way of connecting ontologies is their hierarchical organization, from the top-level ontology with the most general classes, down to more and more specific domains. The shared top-level ontology for the OBO Foundry is the Basic Formal Ontology (BFO) [1], which contains general classes like *Material object*, *Quality*, *Role*, and *Process*. BFO builds on the three basic distinctions between particulars vs universals, continuants (existing at a time, like material objects or qualities) vs occurrents (existing through time, like processes), and independent entities (like material objects) vs dependent entities (like qualities) [15].

3. Ontologies in the Humanities

The gain of applied ontologies is, however, not restricted to the biomedical domain. It is relevant for any data-intensive and complex field of research. The complexity of socio-cultural phenomena outruns the complexity of the biomedical domain due to the creative potential of human languages. E.g., the combinatorial possibilities of a single tweet on Twitter exceeds the estimated number of particles in the entire universe. Given this observation, applied ontology would be desperately needed for structuring data. As of now, however, it is not much used. Hence, it would be desirable to develop ontological analysis into a universal tool in the representation of socio-cultural knowledge and data.

Information scientists have developed several ways to encode large databases of computer-readable ontological knowledge. These range from ontology description languages like the Web Ontology Language (OWL) [45] via ontology editors to automated reasoning programs, which can infer new knowledge: they can make explicit what has been only implicit in the assertions made by human ontology curators. By now, information science draws heavily supported on philosophical work in ontology (Smith 2003). Though still debated [26], a realism-based paradigm emerged as a guide to good ontology design [38, 41], now described on textbook level [20, 27]. It is supported by various standards describing, e.g., the use of formal relations [39].

There is, however, still a lack of application of these new tools to the field of digital humanities. Presently, some textbooks for this field do not even mention ontologies (like [22]), or mention them in passing only (like [7], p. 154). At least one textbook does discuss ontologies to a fair extent, but in a way that fails short of the state of the art reached in the life sciences [14, pp. 162–176]. I will discuss this treatment of ontologies in Section 5. Some standards do exist, though. To mention some: Museums use the CIDOC Conceptual Reference Model (CIDOC CRM; ISO 21127) for describing material cultural heritage, libraries use standards like "Resource Description and Access" (RDA) [23] or the "Functional Requirements for Bibliographic Records" (FRBR) [13] for bibliographic data. Moreover, there are plenty of resources of various size and generality, from the general-topic DBpedia, harvested from Wikipedia pages [2], to rather specific items like the InPhO, the Indiana Philosophy Ontology [5], or the competing PhilOnto [11]. There is, however, no overarching standard for connecting the data in these resources, to avoid incompatibilities, and to make them interoperable. As the paradigm of the life science ontologies demonstrates, the availability of interoperable ontologies could immensely improve access to cultural data and their interoperability, and ease their automated processing.

4. Case study 1: Social Entities in the NCIT

As a first case study (previously presented in [16]) I turn to an example from the biomedical domain, the National Cancer Institute Thesaurus (NCIT), a terminology database designed for the needs of the US National Cancer Institute [9]. Despite being designed for the biomedical domain, the NCIT covers a respectable selection of social entities. The following problems can be observed in the NCIT:

(1) References to social entities in the NCIT are unsystematic and have a national bias towards topics relevant for the US. E.g., the only item listed under the heading "Underrepresented Minority" is "American Indian or Alaska Native". Such **eclecticism** concerning the things to be represented is a hindrance to data-integration across national borders or topical domains.

(2) Furthermore, the NCIT is rather **parochial** in its horizon. E.g., "Underrepresented Minority" is simply defined as a group "underrepresented in cancer research". But, of course, a group can be underrepresented in many other ways, too. If the definitions ignore that the entities defined also exist outside a given domain, this affects the interoperability with other terminology databases.

(3) Often, the NCIT follows **topical associations rather than ontological guidelines**, as they are provided by the OntoClean method [12] or the GoodOD guideline [35]. E.g., the item "Business Rules" has 39 sub-items in the NCIT. An NCI business rule is, e.g., the "Improve access" rule, which is: "Support the effective dissemination, communication, and utilization of HIV/AIDS information […]." This is, of course, a formulation of the rule, not a description of it; and it is a particular rule, not a kind of rule. Nor is the plural of the term "Business Rules" appropriate if the term is to feature in the subsumption relation "subclassOf" between a class and its superclass (sometimes also called the *is_a* relation). The subsumption relation has to be distinguished from the instantiation relation between an individual thing and a class to which it belongs. We would thus have (mind the singular):

Business rule subclassOf *Rule*

Improve access rule **instance of** *Business rule*

Even worse, most of the sub-items of "Business Rules" are neither rule types nor rule instances. One sub-item of "Business Rules", e.g., is "Employment Opportunities", which are (again rather parochially) defined as "Jobs available at NCI". These are, of course, no business rules at all. Rather, the process of filling employment opportunities with suitable candidates is something that is governed by business rules. The NCIT statement

Employment Opportunities (sic!) subclassOf *Business Rule*

is not only awkward, but also false, and thus not apt for a scientific representation.

(4) The NCIT often mixes ontological categories. E.g., "Clinical Research" is given as a synonym of "Clinical Study". But this cannot be true: Clinical research is the overall activity which relates to a clinical study as uncountable stuff to countable things [8, pp. 153–156). Similarly, the NCIT does not clearly distinguish between the "Personal Medical History" of a patient and its description as part of the patient record.

(5) The NCIT entries often do not reflect actual properties of the entities classified. An item like "Other Minority" cannot refer to a property of otherness instantiated by certain minorities. It makes sense only in relation to the rest of the classification given in the NCIT. Such "other" items (there are more than hundred of them in the NCIT) are means to secure exhaustiveness of the classification on that level, but they do not reflect the structure of the world and are thus a hindrance for interoperability, as the classification will be different in other databases. Other modifiers in the NCIT are similarly troublesome, e.g., "None or Not Applicable", "Not Defined", "Not Otherwise Specified", "Not Stated" or "Unknown", as they do not describe the entities themselves but the way we know or describe them [4].

These shortcomings of the NCIT reveal a lack of rigor when it comes to social ontology. If a sound underlying ontology is necessary for a coherent terminology in general, a sound social ontology is necessary for a classification of social entities in particular. The situation could be improved by four ontological guidelines for the development of an integrated ontology of the socio-cultural world:

1. Connect with standard ontology. Most formal ontological dichotomies apply to the social realm, too. In particular, this is true of the three fundamental dichotomies that also lie at the ground of the Basic Formal Ontology (BFO), i.e., continuant vs. occurrent, independent vs. dependent, particular vs. universal (Arp, Smith & Spear 2015; Jansen 2008).

2. Social *F*s are *F*s. In the history of ontology, it has been a matter of dispute whether the same categories apply to natural and social entities [21]. And indeed it can be asked whether the appropriate category of, say, *Academic degree* is a *Quality* or rather something like *Social quality*, which is then subsumed under a top-level entry *Social entity*, disconnecting the classification of social entities from the classification of other entities. Some qualifying phrases like "pseudo-…" or "bogus …" are indeed alienating phrases; something written by Pseudo-Aristotle is not written by Aristotle, and a bogus proof is not in fact a proof. The adjective "social", however, is a separable modifier phrase, like "living": A living horse is both a horse and living, and a social activity is both an activity and social. Thus, social activities should be classified as activities, a social status should be classified as a status, social circumstances should be classified as subclass of circumstances, etc.

3. Respect specific social categories. In addition to the general categories of formal ontology, there are categories specifically pertinent to the social world. A famous example of a label for a specific social category is "institution". Unfortunately, this label is ambiguous in natural language and is, in fact, used for three distinct though interrelated categories of social entities [18, Ch. 9]: (i) for institutional rules (e.g., constitutive rules of the "counts-as" type [36]), (ii) for things instituted, and (iii) for the act of instituting something. The NCIT does not always distinguish these three meanings of "institution"; e.g., the NCIT definition of "marriage" illegitimately confounds (a) the abstract institution of marriage to be found in some societies but not in others, and (b) concrete marriages.

4. Make explicit ontological relations between social entities. Social entities are not isolated, but are interconnected to other social entities as well as to natural entities. Often, these connections are formal relations that also apply to other realms of reality, like the relations **part of**, **participates in**, **has role**, and so on [10, 35, 39]. Many of these relations are not used in the NCIT; others are used, but in odd ways. Rigorous application of both a coherent set of top-level categories and such ontological relations will heavily improve the representation of social entities. E.g., the NCIT contains both "Social work" and "Social Worker", but does not relate them to each other. But of course, a social worker is someone who is trained or employed to do social work; this is the expected role-performance of a social worker.

In addition to general formal relations, specific social relations like membership will also be needed. While some have argued that membership is a variety of parthood [28], it is rather a social relation in its own right. Parthood is a transitive relation, membership is not; the same members can constitute several distinct groups, while the same parts uniquely assembly to exactly one whole; and localized parts form localized wholes, while localized members can form non-localized organizations [30, Ch. 2; 18, Ch. 3).

5. Case study 2: Improving 'The World of the Nobel Prize'

In a second case study, I turn to an example discussed by Rehbein [29]. This discussion is part of an introduction into Digital Humanities that is, to my knowledge, presently the only one that contains an introduction to digital ontologies on textbook level. The treatment of ontologies in this introductory text, however, disregards the standards developed in the domain of the life sciences during the last decades. While it is good to see ontologies discussed in a DH textbook at all, these peculiarities are shortcomings of Rehbein's account from the point of view of these standards. I will point to some of these features while discussing the toy ontology he presents as an example, which he calls 'The World of the Nobel Prize'. This ontology comes close to what one might call an application ontology, though no application is stated. Next to the upper-most class *Thing,* it contains the classes *Profession*, *Person*, *Intelligent person*, and *Nobel Prize* (Figure 1). This is far from, say, OBO Foundry good practise in several respects.

First, these classes are introduced *ad hoc*. They are not related to any top-level ontology. Using BFO, *Person* might be said to be a subclass of *Material object*, *Profession* is a subclass of *Role*, and *Intelligent Person* is a defined class, whose definition refers to the quality *High intelligence*. That is, the following axioms should be added:

Profession subclassOf *Role*

Person subclassOf *Material object*

Intelligent person equivalentTo (*Person* and **bearer of** some *High intelligence*)

High intelligence subclassOf *Quality*

It is not as obvious to which upper category *Nobel Prize* belongs. Instead of introducing this class of abstract entities of dubious categorical belonging, it would have been easier to introduce a process class *Nobel Prize Awarding*, with the Nobel committee as the (collective or institutional) active participant and the respective winners as passive participants.

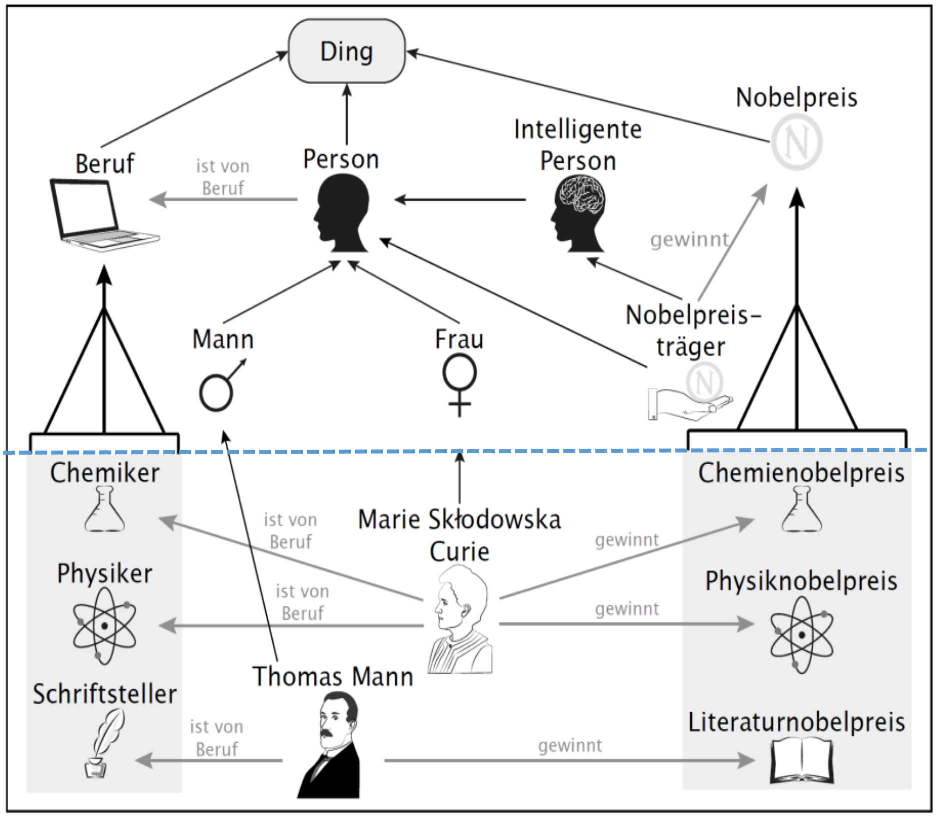

Figure 1. The 'World of the Nobel Prize'. The toy ontology together with some instances below the dotted line. For a translation of terms, see the discussion in the text. Source: Rehbein 2017, 167 (modified).

Second, *Intelligent person* is characterized as a subclass of *Person*, as are *Man* and *Woman*. Doing so would treat these classes like subclasses of *Person* in the same way as *Dog* and *Cat* are subclasses of *Mammal*. But this blurs the important contrast between rigid and non-rigid attributes (highlighted in, e.g., the OntoClean method [12]; cf. also [35]): Being a cat is a rigid feature: any animal that is a cat at some time is a cat at all times of its life. Intelligence and, as we now know, the sex–gender complex, is not rigid in this sense. Someone who is very intelligent at one time, might become not-so-intelligent later, e.g., due to an accident. Similarly, Thomas Mann might have decided to become a woman at some time in his (or her) life, had surgical intervention for this been available during his life time. Here an alignment to a top-level ontology is very helpful. BFO makes it clear, e.g., that roles are non-rigid by their very definition.

Another problem lies in the delineation of classes from instances. Surprisingly, while *Nobel Prize* is introduced as a class, the Nobel Prizes for chemistry, physics and literature are introduced as instances. But aren't there *many* Nobel Prizes for, say, Literature – each year a new one? In a way, Thomas Mann and Bertrand Russell can have said to have won 'the same prize'. But the same phrase could be applied to Thomas Mann and Madame Curie. So this is no conclusive evidence whether the sameness here is based on token identity or type identity. Things are much clearer with the alternative modelling strategy suggested above. For sure, the process of awarding the Nobel Prize of literature to Thomas Mann is a different token of the same class as awarding the Nobel Prize of literature to Bertrand Russell.

Finally, the toy ontology uses its own idiosyncratic relations: **has profession** and **wins**. OBO Foundry practise suggests not to introduce relations in an ad hoc manner, in order to improve interoperability and semantic consistency. Instead, the more generic relations **has role** and **participates in** could be used, in combination with the classes or top-level alignments suggested above. Currently, the toy ontology models in the following way:

ThomasMann has profession Writer

ThomasMann wins Nobel Prize of Literature

Instead, it would be advisable to follow the following patterns:

ThomasMann has role some *Writer role*

ThomasMann passive participant in some *Nobel Prize of Literature Awarding*

The toy ontology of Rehbein has, of course, been developed to illustrate certain elements of digital ontologies, like the distinction between classes and instances, the combination of classes and relations (both subsumption and other relations), and the possibility of designing a complex semantic network by these means. It is designed, however, in a rather *ad hoc* way, and thus fails to illustrate certain good-practise rules that have been established in ontology development in the life sciences. By sticking to these good-practise rules, an ontology becomes interoperable with other ontologies following the same standards. There will be higher precision and semantic consistency.

6. Special Challenges for an Ontology of the Socio-Cultural Domain

6.1. Can there Be an Ontology of the Social?

Luhmann [25] and others famously claim that an ontology of socio-cultural phenomena is not possible for principled reasons. On closer examination, however, Luhmann does operate on his own ontological assumptions and only seems to reject essentialist or eternal social categories [18, Ch.1] – for in the socio-cultural domain nothing seems to be 'fixed', and everything is constantly changing and developing. This does, however, not prevent the possibility of a useful ontological category scheme being developed for the socio-cultural domain. In the end, the cosmic and biotic domains of our world have also developed in time, and due to evolution nothing is 'fixed' in the biological domain either.

6.2. Constructivist Arguments

Another argument that is often adduced against the possibility of an ontology of the social is based on the observation that social entities are 'constructed', and hence, it seems, not of the same dignity as the eternal essences that sometimes are supposed to be the topic of ontology. Nothing, however, prevents an ontological analysis of non-eternal beings. To the contrary: the fact that an entity is a human construction allows us, first, to infer that it does in fact exist, and, second, gives hope that it is more easily analyzed than a biomedical phenomenon [19, 43].

6.3. Arguments from Vagueness and Ambiguity

It might also be argued that the socio-cultural domain is intrinsically vague and ambiguous, and, thus the argument would go, that the exactness of a formal ontology is misplaced in any treatment of the socio-cultural domain. However, this assumption holds only for parts of the socio-cultural domain. Friendship might be vague and ambiguous, but marriages are crisps and legally regulated [17]. Moreover, the biomedical domain displays the same features of vagueness and ambiguity, as the world of life is full of continuous transitions and exceptions. In fact, 'life' itself is a word that is both intensionally and extensionally vague – both in the diachronic and the synchronic perspective [3]. Finally, even if a domain is vague, the research vocabulary used to describe the domain should be controllable and researchers should be able to account for possible ambiguities and be able to disambiguate their statements.

6.4. Arguments from the Diversity of Languages

Also, humanities are more language sensitive than the natural science. In the sciences, English is more or less the uncontested *lingua franca* used for the lion share of all publications, for conferences and correspondence, describing a language-independent domain. Cultural studies and the humanities, or so it seems, feature much more linguistic variety, and many cultural phenomena are language-dependent. Still an ontologically structured common terminology and classification is possible for cultural phenomena, though it will be more complex. This can be seen, e.g., in the terminology of grammar: Descriptive linguistics can find syntactic phenomena like tense, number or person across a wide variety of languages, while not every language has to feature each of these phenomena. (The plain language of propositional logic, e.g., does not feature any of these three phenomena.) Thus, an ontological analysis of cultural phenomena is possible; and the case studies discussed above show that it is also extremely helpful to organize data.

6.5. Disputed Entities

Another objection to a general ontological framework for the Digital Humanities could be adduced from entities of dubious existence. There is no consensus in religion, myth or spiritual beliefs as to which entities exist or do not exist. How can we deal with this problem within an ontological framework? One suggestion to deal with this situation is to use defined classes for contentious entities, as the standard semantics (e.g., of OWL [45]) for digital ontologies comes with an open-world assumption and does not assume that all classes are instantiated – whereas there is an ontological commitment to all individuals mentioned [32]. Another strategy would be to deal with those entities about whose existence there is a consensus [33]. Hence, while it might be contentious that there are gods, demons, spirits, and revelations, there is a consensus that there are narrations about gods, actions that aim to influence demons, texts that are believed to root in acts of revelation, and so on. A DH ontology would thus not be compelled to introduce a contentious *God* class or a God individual, but could instead talk about religious texts and narrations. These could be dealt with by means of the Information Artefact Ontology (IAO). As the IAO has primarily been developed for the biomedical domain, there is the problem how to model contentious content of an uncontentious *Information Content Entity* (ICE) [42]. One solution would be to introduce a class *God character* to which religious texts ascribe certain properties – in analogy to literary characters, which have been characterized as theoretical entities of literary criticism [44]. All this shows that there are challenges in the humanities which do not normally arise in the life sciences. However, it also indicates that there are suggestions on how to deal with these challenges.

7. Discussion and Conclusion

In this paper I have argued that ontology development in the Digital Humanities can learn from ontology development standards developed for the domain of the life sciences. For this claim, I have first collected evidence from two case studies, namely by analyzing the social entities in the NCIT and by re-engineering a toy ontology from a DH textbook. The case studies revealed a plethora of inaccuracies, factual errors, or disregard of basic modelling rules. Second, I discussed (and dismissed) some standard objections to the use of general ontological frameworks in the humanities. It is, thus, not only possible, but also highly desirable that DH ontology development learn from the life sciences, and develop standards for domain modelling. This would include the use of top-level ontologies and standardized relations.

The case studies show that social ontology and principled ontology development guidelines can help to improve terminologies and classifications used to represent and structure the socio-cultural domain. There are, however, some intriguing open questions for DH ontologies: What are the up-most categories for social entities? What fundamental formal relations are needed to give an adequate picture of the socio-cultural domain? How can these means be used to analyze important socio-cultural categories? And, last but not least: How can existing data standards in digital humanities and social sciences be integrated to make them consistent and interoperable by means of these analyzes? In order to have a coherent approach to answering these questions, it is very much desirable to have an open platform for DH ontologies, comparable to the OBO Foundry. It would also be desirable to have an upper-level ontology for the socio-cultural domain to ease DH ontology development – in analogy to the upper-level ontology BioTop for the biomedical domain [31]. This upper-level ontology should allow integrating existing standards, like CIDOC CRM, and RDA. Like in the life sciences, such an effort could only be achieved by means of a coordinated collaborative project. It would thus be very much desirable to initiate a network of ontology developers and stakeholders with this very goal.

8. References

[1] R. Arp, B. Smith, A. Spear, *Building ontologies with basic formal ontology*, MIT Press, Cambridge MA, London, 2015.

[2] S. Auer, DBpedia: A large-scale, multilingual knowledge base extracted from Wikipedia, *Semantic Web* **6** (2015), 167–195.

[3] M. A. Bedau, What is Life?, in: S. Sarkar, A. Plutynski (Eds.), *A Companion to the Philosophy of Biology*, Blackwell, Malden/Oxford.

[4] O. Bodenreider, B. Smith, A. Burgun, The Ontology–Epistemology Divide: A Case Study in Medical Terminology, in: A. Varzi, L. Vieu (Eds.), *Formal Ontology in Information Systems. Proceedings of the Third International Conference (FOIS 2004)*, IOS, Amsterdam, 2004, 185–195.

[5] C. Buckner, M. Niepert, C. Allen, From encyclopedia to ontology: toward dynamic representation of the discipline of philosophy, in: *Synthese* **182** (2011), 205–233.

[6] M. Courtot, F. Gibson, et al., MIREOT: The minimum information to reference an external ontology term, *Applied Ontology* **6** (2011), 23–33.

[7] D. Fiormonte, T. Numerico, F. Tomasi, *The Digital Humanist. A Critical Inquiry*, Punctum, Brooklyn 2015.

[8] A. Galton, *The Logic of Aspect*, Clarendon Press, Oxford, 1984.

[9] J. Golbeck, The National Cancer Institute's Thesaurus and Ontology, *Journal of Web Semantics* **1** (2004), 75–80.

[10] P. Grenon, B. Smith, The Cornucopia of Formal-Ontological Relations, *Dialectica* **58** (2004), 279–296.

[11] P. Grenon, B. Smith, Foundations of an ontology of philosophy, *Synthese* **182** (2011), 185–204.

[12] N. Guarino, C. Welty, An Overview of OntoClean, in: S. Staab, R. Studer (Eds.), *The Handbook on Ontologies*, Springer, Berlin, 2004, 151–172.

[13] IFLA, *Functional Requirements for Bibliographic Records. Final Report*, September 1997, as amended and corrected through February 2009, http://www.ifla.org/VII/s13/frbr/.

[14] F. Jannidis, H. Kohle, M. Rehbein (Eds.), *Digital Humanities. Eine Einführung*, Metzler, Stuttgart 2017.

[15] L. Jansen, Categories: The Top Level Ontology", in: K. Munn, B. Smith (Eds.), *Applied Ontology*, Ontos, Frankfurt/Lancaster, 2008, 173-196.

[16] L. Jansen, Four Rules for Classifying Social Entities, in: R. Hagengruber, U. Riss (Eds.), *Philosophy, Computing and Information Science*, Pickering & Chatto, London, 2014, 189–200.

[17] L. Jansen, Gern helf' ich dem Freunde? Pflichten in informellen Sozialbeziehungen, in: K. Mertens, J. Müller (Eds.), *Die Dimension des Sozialen. Neue Philosophische Zugänge zu Fühlen, Wollen und Handeln*, de Gruyter, Berlin/Boston 2014, 333–349.

[18] L. Jansen, *Gruppen und Institutionen. Eine Ontologie des Sozialen*, Springer VS, Wiesbaden 2017.

[19] L. Jansen, Constructed Reality, in: C. Kanzian, J. Mitterer, K. Neges (Eds.), *Realismus – Relativismus – Konstruktivismus*, de Gruyter, Berlin et al., 2017, 255–267.

[20] L. Jansen, B. Smith (Eds.), *Biomedizinische Ontologie. Wissen strukturieren für den Informatik-Einsatz*, vdf, Zürich, 2008.

[21] T. Kobusch, *Die Entdeckung der Person. Metaphysik der Freiheit und modernes Menschenbild*, 2nd, extended edition, Wissenschaftliche Buchgesellschaft, Darmstadt 1997.

[22] S. Kurz, *Digital Humanities. Grundlagen und Technologien für die Praxis*, Springer Vieweg, Wiesbaden 2015.

[23] F. Lazarinis, *Cataloguing and Classification. An Introduction to AACR2, RDA, DDC, LCC, LCSH and MARC 21 Standards*, Chandos, Amsterdam et al., 2015.

[24] S. Leonelli, *Data-centric biology. A philosophical study*, University of Chicago Press, Chicago, London, 2016.

[25] N. Luhmann, *Soziale Systeme. Grundriß einer allgemeinen Theorie*, Suhrkamp, Frankfurt am Main, 1984.

[26] G.H. Merrill, Ontological realism: Methodology or misdirection? *Applied Ontology* **5** (2010), 79–108.

[27] K. Munn, B. Smith (Eds.), *Applied Ontology. An Introduction*, Ontos, Frankfurt am Main, 2010.

[28] A. Quinton, Social Objects, *Proceedings of the Aristotelian Society* **76** (1975/76), 1-27.

[29] M. Rehbein, Ontologien, in: F. Jannidis, H. Kohle, M. Rehbein (Eds.), *Digital Humanities. Eine Einführung*, Metzler, Stuttgart 2017, 162–176.

[30] D.-H. Ruben, *The Metaphysics of the Social World*, Routledge & Kegan Paul, London 1985.

[31] S. Schulz, M. Boeker, C. Martinez-Costa, The BioTop Family of Upper Level Ontological Resources for Biomedicine. In: R. Randell, R. Cornet et al. (Eds.), *Informatics for Health: Connected Citizen-Led Wellness and Population Health*, IOS, Amsterdam 2017, 441-445.

[32] S. Schulz, M. Brochhausen, R. Hoehndorf, Higgs bosons, Mars missions, and unicorn selusions: How to deal with terms of dubious reference in scientific ontologies, in: O. Bodenreider, M.E. Martone, A. Ruttenberg (Eds.), *Proceedings of the 2nd International Conference on Biomedical Ontology, Buffalo, NY, USA, July 26-30, 2011. CEUR-WS* **833** (2011), http://ceur-ws.org/Vol-833/paper24.pdf.

[33] S. Schulz, L. Jansen, Towards an Ontology of Religious and Spiritual Belief, in: S. Borgo, P. Hitzler. O. Kutz (Eds.), *Formal Ontology in Information Systems. Proceedings of the 10th International Conference (FOIS 2018)*, IOS, Amsterdam, 2018, 253–260.

[34] S. Schulz, D. Seddig-Raufie, N. Grewe, J. Röhl, D. Schober, M. Boeker, L. Jansen, *Guideline on Developing Good Ontologies in the Biomedical Domain with Description Logics*. Version 1.0 (2012), http://www.purl.org/goodod/guideline.

[35] U. Schwarz, B. Smith, Ontological Relations, in: Munn & Smith (Eds.) 2008, 219–234.

[36] J. R. Searle, The Construction of Social Reality, Free Press, New York, 1995.

[37] B. Smith, Ontology, in: Luciano Floridi (Ed.), *Blackwell Guide to the Philosophy of Computing and Information*, Wiley, Oxford, 2003, 155–156.

[38] B. Smith, Beyond Concepts. Ontology as Reality Representation, in: Achille Varzi, Laure Vieu (Eds.), *Formal Ontology and Information Systems*, IOS, Amsterdam, 2004, 73–84.

[39] B. Smith, W. Ceusters et al., Relations in Biomedical Ontologies, in: *Genome Biology* **6** (2005), R46.

[40] B. Smith, M. Ashburner et al., The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration, in: *Nature Biotechnology* **25** (2007), 1251–1255.

[41] B. Smith, W. Ceusters, Ontological Realism as a Methodology for Coordinated Evolution of Scientific Ontologies, in: *Applied Ontology* **5** (2010), 139–188.

[42] B. Smith, W. Ceusters, Aboutness: Towards foundations for the Information Artifact Ontology, in: F.M. Couto, J. Hastings (Eds.), *Proceedings of the International Conference on Biomedical Ontology, Lisbon, Portugal, July 27-30, 2015*, *CEUR-WS* **833** (2015), http://ceur-ws.org/Vol-1515/regular10.pdf.

[43] A.L. Thomasson, Realism and Human Kinds, *Philosophy and Phenomenological Research* **67** (2003), 580–609.

[44] P. van Inwagen. Creatures of Fiction, *American Philosophical Quarterly* **14** (1977) 299–308.

[45] W3C, *OWL 2 Web Ontology Language. Document Overview (Second Edition)*, W3C Recommendation 11 Dec 2012, http://www.w3.org/TR/owl2-overview-20121211.

[46] Z. Xiang, M. Courtot, et al., OntoFox: web-based support for ontology reuse, BMC Research Notes **3**:175 (2010)