




Systems of Intertextuality: Towards a formalization of text relations for manual annotation and automated reasoning

Jan Horstmann <jan_dot_horstmann_at_uni-muenster_dot_de>, University of Münster  <https://orcid.org/0000-0001-8047-2232>

Christian Lück <christian_dot_lueck_at_uni-muenster_dot_de>, University of Münster  <https://orcid.org/0009-0008-0649-5127>

Immanuel Normann <immanuel_dot_normann_at_uni-muenster_dot_de>, University of Münster  <https://orcid.org/0000-0003-4702-1282>

Abstract

How can intertextual relations be formalized and annotated? What would be a coherent category system of intertextuality and which formalization is suitable to make it computable while not losing its expressiveness. Against the backdrop of the most influential classical theories of intertextuality, the article does not aim for an automatic detection of intertextual relations like many other digital humanities approaches did before, but suggests a formal and expandable model of the core of intertextuality with the means of description logic, i.e. it models relations and types of entities being related by them in a machine readable RDF format. The utilization of this theory-driven model is demonstrated by several examples of intertextual relations as discussed in literary studies.

1. Everything is relational: the complexities of intertextuality

Intertextuality is a complex yet very central category in the analysis of literature. Per definition it not only concerns one (literary) artifact but at least two and describes the relation between them (cf. [Pfister 1985, 11]). These relations can be found on numerous levels and depending on the personal notion of “intertextuality,” limits can either be established or relations can be found everywhere. Leitch for example claims that “[t]he text is not an autonomous or unified object, but a set of relations with other texts. [...] Every text is intertext” [Leitch 1983, 59].

1

Several literary genres are constituted by their inherent referential character: persiflage, parody, pastiche, cento (patchwork poem), travesty, stylistic copies, and so on. This is one reason why intertextuality is one of the most central categories in literary studies—and one of the complex ones. Below the level of genre constitution, intertextuality can be detected as a more fine-grained web of references in most, if not all, texts. At this level, intertextuality is described with terms such as hypolepsis (textual responses to other texts through, for example, agreement, disagreement, continuation, correction, etc.) or allusion. In addition, references can occur here in the form of, for example, non-genre-constituting stylistic copies or in the parodic mode of a text passage. In general, a distinction is made between single text references and so-called system references, i.e. references to entire “systems” such as genres. A single text passage can (as a direct or indirect quote, as thematic or structural parallelism, etc.) (1) refer to another text passage or (2) to another artifact as a whole (by e.g. evoking the title of that artifact) or (3) to a complete system (e.g. by parodying the norms of a genre, reflecting on beliefs of a specific epoch, etc.). The manner in which these references are established (to parody is just one way) is an additional layer of categories that complexifies the universe of intertextuality.

2

This article has two aims: (1) shed light into the plurality of different conceptions of intertextuality and seek for common denominators that could form an extensible core of intertextuality. (2) Make first steps to computationally model

3

intertextual relations based on a formal logic. This formal logic (a) allows to express the complexities of different possible relations as well as (b) offers computability in order to be able to make the analysis of intertextual relations a global shared task that at some point offers researchers worldwide an extensive, extensible and standardized knowledge graph for the annotation, detecting, querying and analysis of intertextuality.

After revising some of the most central classical intertextuality theories (section 2.1) we will show the different foci that can be and were set in digital approaches towards intertextuality (section 2.2). Many of these had the aim to pave a way towards building an automatic detector for intertextual relations. This is not our goal. Text reuse, the automated finding of quotes or similar text passages in a corpus of texts—in itself very complex from a computer science point of view—frequently seems under-complex from the perspective of literary studies that traditionally deals with semantically much more complex forms of intertextual relations. We instead go the way of modeling a schema, a category system that can be applied to the description of different forms of these relations (section 4). It lies in the nature of such an approach that we focus on the schematic or typological aspects—namely the categories themselves—of the concept and only refer to instances, i.e. specific findings of intertextual references in single texts or text passages, in the form of examples. At the very basis of such an approach lies an evaluation of different formal logics and their ability to model the complexities of intertextuality effectively while at the same time offer machine readability, i.e. possibilities for standardized data acquisition, querying, and reasoning. We argue that description logic, as implemented in the context of knowledge graphs of the semantic web, is a fitting candidate for these tasks (section 3).

Focusing on what we call the core of intertextuality (sections 4.1–3), we show how different theoretical conceptualizations of intertextuality (e.g. Genette’s “hypertextuality”), that each have a different understanding of text and the level of relations between texts, can be adhered to the general model in order to expand it (section 4.4). Finally we look at a specific example of an analysis from literary studies of the intertextual references in Franz Kafka’s *Ein Landarzt* and ask how the findings can be modeled according to the specifications of our conceptualization (section 4.5).

2. Approaches in non-digital and digital research

2.1 Classical intertextuality theories

The term “intertextuality” was introduced in 1966 by Julia Kristeva when she gave a presentation in a course of Roland Barthes. She coined the term in order to overcome certain epistemological barriers structuralism had been running into since the early 1960s (cf. [Dosse 1999]). Following Michael Bakhtin’s analysis of polyphonic novels and his notion of dialogism, Kristeva suggests that the characteristic trait of writing is “both subjectivity and communication, or better, [...] intertextuality” [Kristeva 1986, 39]. The author or writer is situated in two intersubjective or dialogic dimensions: The relation to early writers Kristeva calls the vertical dimension, the relation to the reader she calls the horizontal dimension. The voices of other authors the writer is engaged with and the echos of which are somehow present in his text (polyphony), is the central point that leads her to her concept of intertextuality. She points out, “any text is constructed as a mosaic of quotations; any text is the absorption and transformation of another. The notion of intertextuality replaces that of intersubjectivity, and poetic language is read at least as a double” [Kristeva 1986, 37].

The notion of intertextuality and the mosaic metaphor in this famous passage can be taken up by digital humanities, since the “mosaic” can be considered as a network: We can see the many delimited, single parts (tesserae, texts, nodes) and a bigger picture (the mosaic, the archive, the graph). However, other parts of Kristeva’s contribution have been discussed controversially and also impose bigger challenges to digital humanities. Problems start, when the notion of text transcends from delimited texts in plural to a whole of a single text (sometimes called a “universal text,” cf. [Pfister 1985, 11f.], [Ternès 2016, 84–86]), where everything participates in or rather “is” a universal “intertext.” The linguistic marker of this transgression is the use of singular “text”—instead of “archive of texts” etc.—to denote something that is constituted of myriads of utterances. At the heart of this transgression is Kristeva’s deconstruction of the structuralist binary opposition denotation vs. connotation (cf. [Kristeva 1986, 40–43]), which delivers the foundation of inside vs. outside of a text, i.e. its delimitation to other texts (cf. [Link 1997, 41]). The transgressive (universal) notion of intertextuality can be found in neo-structuralist and deconstructionist literary theories from Barthes to Bloom (cf.

[Pfister 1985, 11–20]).

The other most influential conception of intertextuality originates from structuralism, too: With his concept of architext and the metaphor of palimpsest Gérard Genette gains back a historical dimension of texts (cf. [Dosse 1999, 1.496f., 2.446f.]). His structuralist position is open for the categories of classical rhetorics and he decidedly sets focus on the notion of genre again. In his structuralist conception and as a contrast to a global conception of intertextuality, [Genette 1982] describes—also with the aim of a genre typology—“transtextual” references as transformation or imitation of hypertextual text genres. Genette is aware that every literary artifact can more or less remind readers of another artifact. However, he concentrates on the cases, where these references are obvious and asks for the different kinds of categories such relations can belong to. He differentiates five forms of what he calls “transtextuality”: 1. intertextuality through quotations, plagiarism, or allusions, 2. paratextuality, by which he means framings such as title, genre classifications, author’s name, etc., 3. metatextuality (i.e., critical commentary), 4. architextuality (i.e., external framings assigned by critics, for example), and 5. hypertextuality, in which a later text (hypertext) is inconceivable without a previous reference text (hypotext). Focussing mostly on the hypertextual conception of transtextuality, Genette on the one hand divides hypertextuality into transformation (James Joyce transforms Homer’s *Odyssey* in *Ulysses*) and imitation (Virgil imitates the *Odyssey* in the *Aeneid*) (cf. [Genette 1982, 17]). On the other hand he further differentiates playful, satirical and serious hypertexts, parody and pastiche being playful, travesty and persiflage satirical, and transposition and replication being examples of serious hypertexts.

8

While Genette is aware of the fact that intertextual relations can take many different shapes and happen on several levels (as mentioned above: text passages, whole works, and systems), in his notion of hypertextuality he limits his conceptualization in such a way that only one-to-one relations are discussed, i.e. literary texts that relate as a whole to another hypotext as a whole. An open question is, how these references relate to each other on different levels. Additionally, while Kristeva’s concept of text goes beyond the written text and potentially includes any cultural phenomena, Genette’s model of intertextuality is limited to written whole texts in the sense of works. Consequently, it is permissible to claim that Kristeva’s and Genette’s conceptions of intertextuality form the two ends of a scale: one characterized by utmost openness and global claim, the other by over-limitation and exclusion of many aspects.

9

With regards to Kristeva’s conception of (inter)text, also Pfister points out that “a concept that is so universal that no alternative to it, not even its negation, is imaginable, is necessarily of little heuristic potential for analysis and interpretation” ([Pfister 1985, 15]; our translation). His attempt to mediate between the two conceptual poles of intertextuality (global vs. specific) thus provides six degrees to scale intertextuality (cf. [Pfister 1985, 25–30]): 1. referentiality, 2. communicativity, 3. autoreflexivity, 4. structurality, 5. selectivity, and 6. (with direct reference to Bakhtin) dialogicity. (1) The more a text directly references a pre-text and thematizes it, the more intertextual it is. (2) The more intentional and obviously communicated a reference marker is in the text, the stronger the intertextuality. (3) The more a text itself reflects on its own intertextual conditionality and relatedness, the more intertextual it is. (4) The more fundamentally a pre-text forms the structural foil of an entire text, the more intertextual the text is. (5) The more concise and pointed the specific reference to a concrete pre-text, the stronger the intertextuality (i.e., a quote is more intertextual than e.g. an allusion or a systemic reference). (6) The stronger the semantic and ideological tension between the original and the new context, the stronger the degree of intertextuality.

10

Closely related to Pfister’s category of referentiality is the contribution of text linguistics who described surface markers for referencing other texts like quotations (cf. [Janich 2019, 173]). Such markers may be detected by text mining methods and thus be exploited by digital humanities. Rather studying text types than genres, text linguists have also discussed intertextual relations between types of texts, i.e. system-to-system-references (cf. [Janich 2019, 184–186]).

11

2.2 Digital approaches to intertextuality

2.2.1 Text reuse

In general, methods for automated detection of intertextuality can be classified by degree of text relationship (understood as only written language) ranging from mere syntactic similarity to semantic comparison. In its simplest form, simultaneous syntactic pattern matching—i.e. sequence alignment—is a well studied technique from

12

bioinformatics. Several tools in digital humanities are substantially inspired by this technique.^[1] Various syntactic steps in natural language processing can generalize matching—all analyzing the syntactic surface of a text. Intertextuality understood as syntactic similarity between texts can be detected automatically at large scale by means of text reuse analyses (cf. [Bär et al. 2012]; [Bär et al. 2015]). Intertextuality understood as text reuse can be defined as a “more or less objective recognizable, explicit reference on the surface of the text” ([Burghardt and Liebl 2020], our translation), which leaves out many aspects and categories of intertextuality research that go beyond direct or indirect quotations.

More elaborated digital methods towards intertextuality take semantics into account. In the context of intertextuality the notion of semantics applies on various levels of text components: words, sentences, and whole texts have their own meanings. As far as semantics is taken into account in digital text reuse methods, it so far seems to be restricted to *lexical semantics* and *distributional semantics*. In lexical semantics relations like synonyms, hyponyms, and antonyms are identified by humans and manually organized in lexical knowledge bases like wordnet. In distributional semantics word semantics is basically derived from statistical analysis of word co-occurrence of huge text corpora (e.g. Wikipedia). With the breakthrough of machine learning, this approach became very popular and led to the so-called word embedding method whose most prominent implementation was word2vec. The basic idea is to represent words as points in a vector space such that distances and directions in this space correspond to semantic relations (cf. [Burghardt and Liebl 2020]).

13

2.2.2 Semantic relations between texts

One way to go beyond this is to use latent semantic indexing (LSI), an unsupervised algorithm from the Gensim^[2] framework, to try to model semantic-thematic intertextual references between source and reference texts. This was exemplarily exercised by [Scheirer et al. 2016]. Similarly to latent dirichlet allocation, LSI models textual semantics by looking for statistical co-occurrence patterns. Differently from bag-of-word approaches however, it offers the possibility to compare definable units (i.e. singular text passages) with each other instead of whole texts which makes it particularly useful for research in literary studies beyond distant reading or macroanalysis that, e.g., aims at finding intertextual relations between singular phrases or verses of smaller text corpora. While [Scheirer et al. 2016, 215] emphasize “that LSI can be used to detect intertextual relationships of meaning where few or no words are shared by the two texts,” the approach, however, still remains under-complex from the point of view of literary studies and theories of intertextuality in that there is no specific concept or model of intertextuality applied.

14

2.2.3 Complex formalizations of intertextuality

Independent of how intertextuality could be automatically detected (if at all) is the question of how to formalize intertextuality such that computers potentially can process formalized entities of intertextuality. The mentioned approaches to “text reuse” were largely technology-driven in order to meet tool-specific requirements. The opposite approach is to design a model or formalism that comes close to the theoretical notion of intertextuality in the humanities. First steps are taken by [Schlupkothen and Nantke 2019] who conceive intertextuality as an interpretive phenomenon and attempt to represent analytic-interpretive reading practices in the formalism of situation theory. Situations are basic concepts in this theory which allows us to express predications—so-called *infons*—to hold only within a certain region of time and space. Moreover, situations themselves can be objects of infons such that one can formalize statements like “Along that conversation I recognized that Peter believed that ...” Classical logic does not provide such means of reification and hence makes it rather difficult if not impossible to express some sort of truth conveying. Schlupkothen and Nantke in their paper do not explicitly elaborate how they make use of situation theory or situation logic.^[3] The given examples, however, suggest that it is primarily used to formalize narrative structures as situations.

15

[Oberreither 2020] chooses a conceptually similar modeling approach to ours. The INTRO ontology (*Intertextual Relationships Ontology for literary studies*) he developed for a formalization in the context of the Semantic Web seems to have – similarly to Schlupkothen and Nantke – a foundation in reception aesthetics.^[4] “Receptional features” are “actualized” with each reading, and intertextuality is formalized as a relation not between text passages, but between such receptional features. However: Are all these receptional features really independent of intertextuality, so that intertextuality kind of consumes these features as pre-existing entities of self-contained texts? Isn't intertextuality rather

16

a productive thing, that adds features by relation?

3. The twofold task of formalization: expressiveness and efficiency of formal methods

Whether a specific logic is suitable to formalize a certain domain of interest depends on what you aim to do with the formalized representation of your domain. What makes formalization attractive in scientific context in general is its clarity or precision due to its strict and syntactically easy language and its well defined mechanisms of reasoning: Once a domain is formalized, questions in that formalism can be posed to it and their answers do not depend anymore on meaning or vague arguments, but on proof of a sound calculus.^[5] Of course, developing an appropriate formal model of the domain of interest remains the actual challenge that always has to face criticism.

17

Concerning the suitability of the choice of a certain formalism two aspects are considered particularly relevant in practice: its *expressiveness* and its *computability*. A logic L^1 is said to be more expressive than a logic L^2 if any statement expressible in L^2 can also be expressed in L^1 but not the other way around. The computability of a logic addresses the capability of an associated calculus to compute the truth value of a given statement or to derive new valid statements from existing valid statements. In practice, though, what matters even more is *efficiency*, i.e. the question of how efficient such computations can be performed in terms of processing time and memory space.^[6] As a rule of thumb it can be said that the more expressive a formal logic is the more expensive is its reasoning. Hence, the choice of logic is necessarily a trade off between expressiveness and efficiency determined by the application requirements and their respective priorities. In the following we discuss the most relevant requirements to make explicit why and to what extent formalization matters for our current purpose.

18

Concerning expressiveness, being able to express relationships (at least for binary relations) is certainly a fundamental requirement, since the notion of intertextuality inherently indicates a notion of relationship—namely between texts (independent of the notion of text itself). We also want to be able to classify the items that are subject of intertextuality, as we want to be able to talk more specifically than just about a very general notion of text. For instance, we want to differentiate between text segments, text as a whole, or text systems like e.g. genre. Similarly different kinds of intertextual relationships will be identified and these will be defined on the just mentioned specific item classes. Moreover, classes and relationships frequently are organized hierarchically, i.e. we need a notion of subclass (e.g. epitext is a subclass of paratext) and subrelation (hypertextuality is a subrelation of intertextuality—or transtextuality in Genette's terms). These general requirements already exclude e.g. propositional logic as a formalism in that it does not provide any means to express relationships.

19

As far as expressive power is concerned, first order logic would be a sufficient formalism for our purpose.^[7] Unfortunately it is undecidable whether a first order logic formula is provable.^[8] In other words: If we express two statements A and B in first order logic and claim that B is derivable from A then there is no algorithm that could prove this claim in any case of A and B.

20

Fortunately, we can restrict ourselves to a subset of the first order logic—namely to constants, unary, and binary predicates—without sacrificing the expressive power that we need. A first order logic with this restriction corresponds directly to description logic (DL). The reader may consult standard literature for a gentle introduction into this logic.^[9] For the sake of brevity a simplistic example in such restricted logic may serve as illustration:

21

- $\text{Text}(a) \wedge \text{Text}(b) \wedge \text{partOf}(a,b)$ ^[10]
- $\forall a,b,c. \text{partOf}(a,b) \wedge \text{relatedTo}(b,c) \Rightarrow \text{relatedTo}(a,c)$ ^[11]

The important point from a practical point of view is that there is a web conform implementation of the theoretical framework of description logic called Web Ontology Language (OWL)^[12] that is machine processable by various knowledge base systems and reasoners.^[13]

22

4. In search of the core of intertextuality

4.1 A common denominator of intertextuality theories?

While there are many theoretical designs of intertextuality around, they share a common conceptual core: Intertextuality is about relations between texts. This common denominator is obviously contained in the two lexical constituents of the neologism, in “inter” and “text.” The third morphological component of inter-text-uality is a derivational morpheme which goes back to *-alitas* and *-alité*, respectively, of Romance morphology. Let’s take this morphological analysis as a starting point to crystallize the common core of the various concepts of intertextuality and to get the foundation of its formalization.

24

The analysis of lexical morphemes may seem trivial at first, but the theoretical consequences are not. “Inter,” i.e. “between” (also: “among” as in “*primus inter pares*” or “while”), is a local preposition that can be transferred to temporal and other relations. It presupposes (at least) two distinct objects. This is still true if their exact boundaries become indeterminable and distinguishing the related objects becomes problematic. E.g., if in Latin “*inter canem et lupum*” is a metaphor for the hour when dogs and wolves cannot be distinguished, that is, for dusk, (cf. [Georges 1913, 8.942]) then dog and wolf, domesticated and wild, day and night, nevertheless remain distinguished in order to set up and set in motion the game of indistinguishability in the continuum between species, characteristics, and times. If we transfer the local preposition “inter” to texts, here, too, their distinctness is the condition of the possibility of speaking of “inter.” In a phenomenological regard, distinguishing two texts is without problems: two codices, two rotula, two sheets of paper, etc. However, in a structural sense, which is of more interest here, distinguishing texts relies on processing a text’s inside-outside difference, which is described by the opposition of denotation and connotation. “Denotates are all signifiers whose signifiers belong to the set of text-signs and which fit the context of the text” ([Link 1997, 41]; our translation). Connotates, on the other hand, are not direct, not notated in the text. Unsurprisingly, Kristeva has started her deconstruction of structural fundamental oppositions right here and tries to subvert binarity through a continuum. The morpheme “inter” sets focus on that intermediate area which is not one and not the other. Genette has perhaps not coincidentally focused on a different local preposition and thus a different spatial logic: “trans,” engl. “beyond” and “across,” does not direct the focus to the intermediate area, but introduces a directionality and with the beyond also a deictic center: Here one text refers to the other. And this seems to be precisely what is often meant by the study of the intertextual relations of a text to other texts. Even if intertextuality and transtextuality, with their prepositional morphemes, are certainly not as “transparent and self-explanatory” ([Adamzik 2004, 96]; our translation) on close inspection as some authors think, both approaches have the same denominator: They describe relations between texts that can be distinguished.

25

Opinions differ on the second lexical morpheme of the neologism, “text.” All theories of intertextuality naturally speak of text. Which concept is hidden behind the word, the theories sometimes differ considerably. Even within a theory, the word is used to refer to different concepts or different aspects of a concept. According to Kristeva, “everything, or at least every cultural system and every cultural structure” [Pfister 1985, 7] can be understood as text. At the same time, she also uses the word “text” in the very concrete sense of a written text (such as a particular novel). It is this very concrete notion of text that we place at the beginning of our formalization. We simply write T, a parameter to be bound to a more theoretically refined concept later.

26

The third morphological constituent of the neologism, the derivational morpheme, *-ality*, *-alité*, *-alitas*, *-al-i-tas*, *-al* (adjectival) and *-tas* (substantival), determines what intertextuality is taken for. It indicates a passage through a twofold change of part of speech, the distortion of a property into something that exists, into substance. On the one hand, this substantializing morphology—as in *nation-al-i-ty*, *re-al-i-ty*, *relation-al-i-ty*, etc.—may be seen as the reason for the misery of intertextuality theories. On the other hand, a valuable clue to a common core emerges here as well: we use the word intertextuality in fact more like a substantive noun, and when we quantify it, we do not speak of two, five, or a hundred intertextualities, but, as with water, we add one more unit and speak of two, five, or a hundred intertextuality relations.

27

4.2 Formalizing the core of intertextual relations semi-formally

We can now proceed to formalize the core of intertextuality, i.e. firstly the relations between distinguishable entities and secondly a category system of entity types that can be connected by these relations. We first write the intertextual relation as a function between two T as in Frege's term logic:

```
intertextualRelation'(T, T)
```

Frege suggested such functional notations in order to formalize assertions expressed in natural language.^[14] They can express natural verbs with an arbitrary arity. We could re-translate the formula to the language pattern "text 1 and text 2 are intertextually related." However, this pattern with a plural subject is not exactly what we want. The parameters have a positional order and this means that the relation is directed. From a purely formal point of view, there might be no point in introducing direction. However, with regard to the domain of intertextuality, there indeed is: Intertextuality has been introduced not to describe any formally possible relation, but to approach the historical dimension of texts and their inherently secondary status (cf. [Kristeva 1986]; [Genette 1982]). Thus, we say that an intertextual relation points in an anti-chronological direction from a later text to an earlier text (an *avant* text).^[15] So we bind the formula to the following language patterns: "There is an intertextual relation from text 1 to text 2" or "text 1 has an intertextual relation to text 2."

Again, such an existential proposition ("there is") is not exactly what we want to come out with, since it does not give us a means to further reference the relation, the existence of which we stated. Thus, we introduce the term *i*, which denotes an instance of an intertextual relation and we rewrite the function as follows:

```
intertextualRelation(i, T, T, ...)
```

Say: "*i* is an intertextual relation from text 1 to text 2." We can also put the natural language representation in a more formal way, where we use relative clauses for every parameter following *i*: "*i* is an intertextual relation *which* refers to 'here' in text 1 *and which* refers to 'there' in text 2." This form of defining relative clauses will help us later to rewrite the statement in other formal systems.

Since this is obviously not all we want to be able to express, the formula currently allows an indefinite number of further parameters represented by points. We now try to define at least two other parameters or parameter classes.

A major requirement is to further specify the intertextual relation. One text may be a copy, a travesty, etc., of the other. We note this specification as another parameter *S*. Some kind of specification must be made by an annotator. Because a specification may be made not only in one respect, but possibly in several respects, we note *S+* for a non-empty set of such specifications.

```
intertextualRelation(i, T, T, S+, ...)
```

We also introduce the parameter set *M** for mediators of the relation:

```
intertextualRelation(i, T, T, S+, M*, ...)
```

Introducing mediators may appear a bit intransparent. Examples of mediation instances are provided by Genette's class of hypertextual relations. Here, each single intertextual relation between hypo- and hypertext is mediated by the works as a whole and by a paratextual signal which inaugurates the relation between the works as a whole. Mediators are also important for a whole class of intertextual relations: relations not marked in the referring text but established predominantly by the reader. For such a reception-aesthetically established intertextual relation, the situation of the recipient is to be regarded as a mediator. Another function of mediators is expressing transitivity: One intertextual relation can be the mediator of another intertextual relation. This way, transitive relations can be represented.

4.3 Formalizing the core

4.3.1 Formalizing in predicate logic and RDF

We used Fregeian functional semi-formal notation as an intermediate step towards a formalism with a well defined semantics. In section 3 we gave reasons in favor of description logic as a suitable formalism which in turn is the underlying formalism of the Web Ontology Language also known as OWL/RDF.

36

Since all these formalisms come up with different notations,^[16] we stick to one which we consider the most widespread—namely classical predicate logic—and only present a few translations to other formalisms for the purpose of demonstration.

37

Predicate logic symbols we need to model intertextuality (as motivated in section 4.2) are: the logical symbol “ \wedge ” for “and” and the quantifiers “ \forall ” for “all” and “ \exists ” for “some.” For constants and variables we use lower case letters optionally with an index. For unary predicates we use uppercase words and for binary predicates we use lowercase words.^[17]

38

- unary predicates: IntertextualRelation, Reference, IntertextualSpecification, Mediator
- binary predicates: here, there, specifiedBy, mediatedBy

The expression `IntertextualRelation(x)` then reads as “x is an instance of an intertextual relation” and similarly for `Reference`, `IntertextualSpecification`, and `Mediator`. The expression “`here(x,y)`” reads as “x relates here to y” and similarly for `there`, `specifiedBy`, `mediatedBy`. The sentence “i is an instance of intertextual relation that refers here to t1 and there to t2 and is specified by s and mediated by m” thus translates to the predicate logic expression:

39

```
IntertextualRelation(i)  $\wedge$  here(i, t1)  $\wedge$  there(i, t2)  $\wedge$  specifiedBy(i, s)
mediatedBy(i, m)  $\wedge$ 
Reference(t1)  $\wedge$  Reference(t2)  $\wedge$  IntertextualSpecification(s)  $\wedge$  Mediator(m)
```

We exemplify how this translates to RDF in Turtle notation:

40

```
@prefix : <https://intertextuality.org/abstract#> .
@prefix ex: <https://example.org/my-intertextual-findings/> .
ex:i a :IntertextualRelation;
    :here ex:t1 [a :Reference];
    :there ex:t2 [a :Reference];
    :specifiedBy ex:s [a :IntertextualSpecification];
    :mediatedBy ex:m [a :Mediator].
```

Not any arbitrary expression using these symbols, however, reflects our notion of intertextuality. Informally summarized constraints are: Every instance of intertextual relation must have exactly one here reference and exactly one there reference. Moreover, it must have at least one intertextual relation specification (`specifiedBy`) and it may rely on one or more mediators (`mediatedBy`). In mathematical terms^[18] this means here and there are *functions* assigning to each instance of intertextual relation an instance of reference. Similarly `specifiedBy` and `mediatedBy` are *relations* between the class `IntertextualRelation` and the class `IntertextualSpecification` respectively the class `Mediator`. The domain of `specifiedBy` equals `IntertextualRelation` whereas the domain of `mediatedBy` is a *subset* of `IntertextualRelation`.

41

This can be precisely specified in predicate logic^[19] by the notions of *left-total* and *right-unique* and of *domain* and *range* of relations. The following RDF (more precisely RDFS) formalizations expresses these constraints except for the left-total property:

42

```
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
```



```

@prefix : <https://intertextuality.org/abstract#> .

<https://intertextuality.org/abstract> a owl:Ontology .

:InterTextualRelation a owl:Class .
:Reference a owl:Class .
:InterTextualSpecification a owl:Class .
:Mediator a owl:Class .
:here a owl:FunctionalProperty;
    rdfs:domain :InterTextualRelation ; rdfs:range :Reference .
:there a owl:FunctionalProperty;
    rdfs:domain :InterTextualRelation ; rdfs:range :Reference .
:specifiedBy a rdfs:ObjectProperty
    rdfs:domain :InterTextualRelation ; rdfs:range
:InterTextualSpecification .
:mediatedBy a rdfs:ObjectProperty
    rdfs:domain :InterTextualRelation ; rdfs:range :Mediator .

```

4.3.2 Embedding intertextuality in the Web Annotation Data Model

Formalizing to RDF is not just an end in itself, but also enables practical application. A particularly desirable one from the digital humanities perspective would be a framework to describe intertextuality across resources in the world wide web. The *Web Annotation Data Model*^[20] (WADM) is a framework that fits quite well to our formalization of intertextuality. The basic idea of WADM is to express a relation between one object called body and another called target by a third object called the annotation as depicted in Figure 1.

43

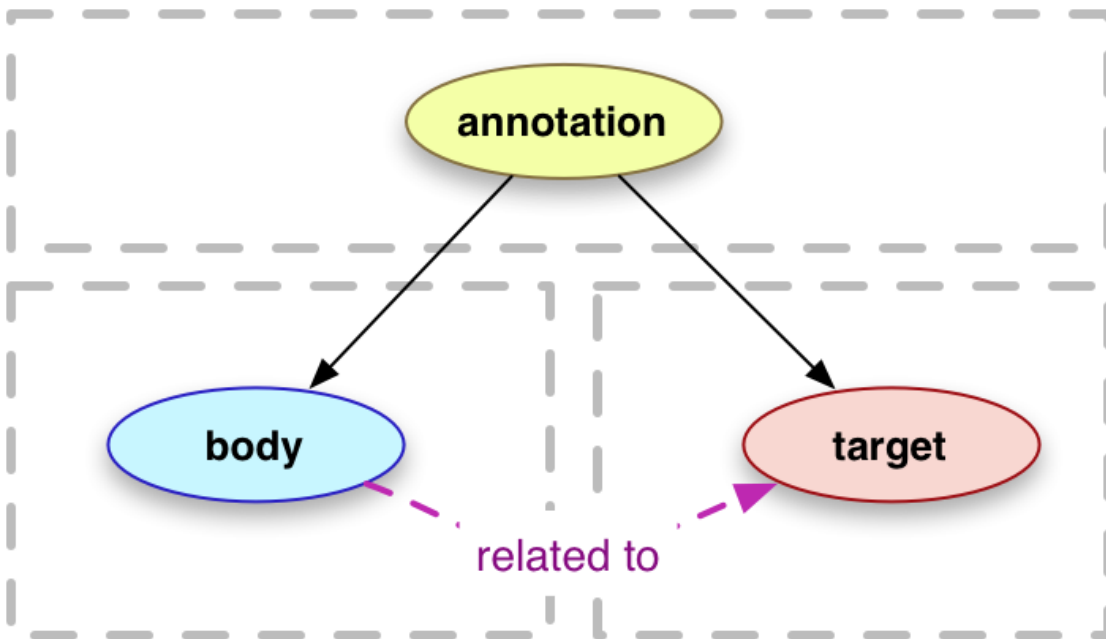


Figure 1. Web Annotation Data Model

Body and target can be any resource in the world wide web like hypertext pages, images, videos, audio streams, etc. Moreover, the model allows not just to reference a web resource like a web page as whole, but also fragments thereof. To this end the model offers the concept of *selectors* for various media types, e.g. XPath constructs for XML like documents (typically HTML, but also SVG), offset for raw text, data positions for bitstreams, etc.^[21]

44

How does WADM fit to our formalization of intertextuality?^[22] A principle about the intended relationship between body and target is: “The content of the Body resources is related to, and typically ‘about,’ the content of the Target resources.”^[23] Our understanding of the word “about” in this statement is that the body is on a meta level whereas the target is on the object level with respect to the annotation. We also find the distinction of meta and object level in our conceptualization of intertextuality: The instances of `Reference` are on the object level whereas the instances of `IntertextualSpecification` and `Mediator` are on the meta level. Hence, we interpret our `Reference` as either a target or body in WADM. More specifically this implies that our here and there predicates should be interpreted as `hasTarget` in WADM. And on the meta level we have an analogous interpretation of `specifiedBy` and `mediatedBy` as `hasBody` in WADM. Finally, corresponding to the `Annotation` in WADM our `IntertextualRelation` is the link between object and target level. We can thus formalize our interpretation in RDF-Turtle notation:

```
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix oa: <http://www.w3.org/ns/oa#> .
@prefix : <https://intertextuality.org/abstract#> .

:here rdfs:subPropertyOf oa:hasTarget .
:there rdfs:subPropertyOf oa:hasTarget .
:specifiedBy rdfs:subPropertyOf oa:hasBody .
:mediatedBy rdfs:subPropertyOf oa:hasBody .
```

4.4 Ontological specializations

The core ontology developed so far does not force its users to follow certain theoretical concepts: it is theory-agnostic in its core. However, this core ontology should be extendable with specializations of core concepts such that specific theoretical concepts can be expressed.

46

We deliberately designed our core ontology so that any theoretic specific aspects can be encapsulated in the specialization of our three core concepts: `Reference`, `IntertextualSpecification` and `Mediator`. The following three subsections demonstrate as a proof of concept how this can be done in the case of Genette’s theory of hypertextuality.

47

4.4.1 Text as specialization of reference

At its core, the formalization has not introduced a notion of text. We first used the parameter `T` to postpone the definition and later we used the class `Reference` to define the range of `here` and `there` but to keep its concept open at the same time. One might object that a user might reference a film and an image and state that they are intertextually related, but that no text is involved in this relation. However, one can think of a theory of intertextuality based on a wider concept of text that allows its application in such a case. We can now: (1) further restrict the application of `IntertextualRelation` to instances where `here` and `there` are texts in a more decided notion or (2) leave this open. The first option would make the ontology unusable for more general theories of text and intertextual relations. The latter might cause inconsistent ontologies in the light of more restrictive theories of text and its relations. A compromise is, however, to keep the definition of text out of the abstract basic ontology and put it into their own namespaces which can be used if wanted or needed.

48

The extension for text suggested here has two aspects: a) restrict the applicability of `IntertextualRelation` to relations between texts, b) further differentiate the notion of text. For either task we introduce the unary predicate `Text`. The restriction task a) is quite simple with the notations from above. We further restrict the range of `here` and `there` to `Text`:

49

```
E1: range(here, Text)
E2: range(there, Text)
```

Task b) involves more conceptual work and it is where the semantic load of this extension originates from. The research and theories around concepts of intertextuality made it possible to differentiate three different levels on which relations can appear. These levels not only apply to text, but to other artifacts to: (1) the *systemic* level of for example genres, epochs, styles, etc., (2) the level of complete *works*, i.e. texts, music pieces or art works as a whole, and (3) the level of *segments*, i.e. text passages, music sequences or sections of art works. Before formalization, we should briefly hint upon difficulties that may come with conceptions of system, work, and segments in their own rights. 50

The differentiation between single text relations and systemic relations stems from theoretical considerations by [Pfister 1985, 53f.]. The notion of system, however, is conceivably unrestricted. As problematic as notions of genre, epoch, literary style or school, etc. themselves, so is a conception between (textual) systems. It is undeniable, however, that single text passages, texts as a whole or even literary systems can and do relate to (other) literary systems in one way or the other. It thus seems the right approach to face the vagueness and also integrate the system level of intertextuality into our formalization in order to make it possible to record these kinds of connections, too. 51

With regards to the notion of complete works, the question of unity that especially arises from the possibility of compiling texts can be problematized, too: Is Proust's *À la Recherche du Temps Perdu* a text, or is the volume *Du Côté de chez Swann* a complete text in its own right? We do not make any decisions in this regard, but leave it up to the specific model usage to decide which units are taken as a text. In doing so, we only presuppose that agreement can be reached about the self-containedness. 52

In contrast to system and work, we speak of a text passage or segment if the beginning and end of this passage lie within a text, but are not identical with the text itself. Equivalent to the vague notion of text as a whole, the question as to what constitutes a segment is fluid: *Du Côté de chez Swann* can be considered to be a work or a segment, depending on individual decisions, theoretical background, or annotation guidelines. 53

To formalize this, we introduce three unary predicates *System*, *Work*, and *Segment*. They are a partition of the set of *Artifacts*, i.e. an artifact is either a system, a work, or a segment. 54

E3: $\text{Artifact} = \text{System} \cup \text{Work} \cup \text{Segment}$

In RDF/Turtle this extension of the TBox reads as follows: 55

```
@prefix a: <https://intertextuality.org/extensions/artifacts#> .
@prefix : <https://intertextuality.org/abstract#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .

<https://intertextuality.org/extensions/artifacts#> a owl:Ontology .

a:Segment a owl:Class .
a:Work a owl:Class .
a:System a owl:Class .
a:Artifact owl:disjointUnionOf (a:Segment a:Work a:System)
```

We can specialize the concept of artifacts and its partition to the domain of *Text* as *Text* is a subset of *Artifact*, too. A work can be a text, but it need not be. The partition of the set of artifacts into systems, works and segments applies to texts, too. We define: 56

E4: $\text{SingleText} = \text{Text} \cap \text{Work}$
E5: $\text{TextSegment} = \text{Text} \cap \text{Segment}$
E6: $\text{TextGenre} = \text{Text} \cap \text{System}$

Since *Text* is a subset of *Artifact* and *System*, *Work* and *Segment* are a partition of *Artifact*, the following is 57

true:

```
E6: Text = SingleText ∪ TextSegment ∪ TextGenre
```

In RDF/Turtle this extension of the TBox reads as follows:

58

```

@prefix t: <https://intertextuality.org/extensions/text#> .
@prefix a: <https://intertextuality.org/extensions/artifacts#> .
@prefix : <https://intertextuality.org/abstract#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .

<https://intertextuality.org/extensions/text> a owl:Ontology ;
    owl:imports <https://intertextuality.org/extensions/artifacts#> .

t:TextSegment owl:intersectionOf (t:Text a:Segment) .
t:SingleText owl:intersectionOf (t:Text a:Work) .
t:TextGenre owl:intersectionOf (t:Text a:System) .
t:here rdfs:subPropertyOf :here ; rdfs:domain :IntertextualRelation ;
rdfs:range :tText .
t:there rdfs:subPropertyOf :here ; rdfs:domain :IntertextualRelation ;
rdfs:range :tText .

```

Analogously we could specialize the domain of artifacts to the domain of music, art, and further cultural domains. We call these cultural domains the *horizontal dimension* artifacts and the partition into segment, work, and system its *vertical dimension*. For each combination of these two dimensions one may investigate relations. We have looked for prototypical examples and present them in our project repository on github.^[24]

59

4.4.2 Specific intertextual specifications

While the previous section was about a specialization of our general notion of Reference into the vertical dimension (segment, work, and system) and horizontal dimension (text, music, art, ...), this section is about the specialization of our general notions of IntertextualSpecification.

60

Intertextuality theories differ in this aspect even more than in different notions of Text (as discussed above). Difference does not necessarily imply contradiction, though. More often, it means a more or less independent refinement or granularity of concepts where different theories focus on different aspects of their subjects. To test the adequacy of our formalization approach, we formalize a tiny part of a single theory as a sample. To this end we choose Genette's concept of hypertextuality ([Genette 1982]; our translation):

61

	playfully	satirically	seriously
transformation	parody	travesty	transposition
imitation	pastiche	persiflage	replica

Table 1.

We consider these terms as types of intertextual relations or more formally as subclasses of our IntertextualSpecification concept and call them all together HypertextualRelation whereof we call *transformation* and *imitation* RelationalType and *playfully*, *satirically*, and *earnestly* ModalType. We can straightforwardly formalize this in RDF:

62

```

@prefix t: <https://intertextuality.org/extensions/text#> .

```

```

@prefix g:
<https://intertextuality.org/extensions/genette/hypertextuality#> .
@prefix : <https://intertextuality.org/abstract#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .

```

```

<https://intertextuality.org/extensions/genette/hypertextuality> a
owl:Ontology ;
    a.

```

```

g:HypertextualRelation rdfs:subClassOf :IntertextualSpecification ;
    owl:disjointUnionOf (g:RelationalType g:ModalType) .
g:RelationalType owl:ObjectOneOf (g:transformation g:imitation) .
g:ModalType owl:ObjectOneOf (g:playfully g:satirically g:earnestly) .

```

This is only a basic formalization. An exhaustive one would have to declare or at least discuss constraints for these six categories since Genette introduces them as relations between whole texts (instances of `SingleText`) which are established and mediated by paratextual signals like the title of a work.

63

4.4.3 Specific mediators

We also want to give some hints towards the formalization of mediators for intertextual relations. At the time of writing, there seems to be no system of mediators, but a variety of mediator types. To define such a mediator type, we declare it to be a *subset* of `Mediator`. For example following Genette's notion of hypertextuality, intertextual relations between Joyce's *Ulysses* and Homer's *Odyssey* are mediated by the title of Joyce's novel, which is a paratextual signal. Another example may be motives or even subject matters or forms. The motive "horses in a pigsty," for example, can serve as a mediator for an intertextual relation between a text by Kafka and others.^[25] The formalization task for such a case is (a) to declare a `Motive` to be a subset (subclass) of `Mediator` and then make `HorsesInAPigsty` an element (instance) of this set. Let's write some such definitions down in RDF/Turtle:

64

```

@prefix m: <https://intertextuality.org/extensions/motives#> .
@prefix pt: <https://intertextuality.org/extensions/genette/paratext#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

m:Motive rdfs:subClassOf :Mediator .
m:HorsesInAPigsty a m:Motive .
pt:ParatextualSignal rdfs:subClassOf :Mediator .
pt:Title a pt:ParatextualSignal .

```

We should emphasize that we start putting assertions about individuals into the ontology^[26] here: Making `HorsesInAPigsty` an instance of `Motive` is an assertion. `Title` is a similar case which one could say a lot about, e.g. how to annotate the title as a text segment with web annotations or about its relation to Dublin Core's title.

65

We also add a mediator type, that may look strange at first:

66

```

: IntertextualRelation rdfs:subClassOf :Mediator .

```

Regarding the structure of our data, this makes intertextual relations a recursive datatype. Regarding the implied semantics, it enables us to record sets of transitive intertextual relations, where one relation is mediated by another intertextual relation or a set of other relations, that are again mediated by other intertextual relations.

67

4.5 An example

We choose a rather complex example to demonstrate the expressive power of our formalizations. In an article about intertextuality in Franz Kafka's *Ein Landarzt*, Thomas Borgstedt analyzes a text with a kind of paradoxical intertextuality: it was classified as intertextually poor on the one hand [Borgstedt 2010]. On the other hand obvious intertextual relations have been perceived for a long time, but scholars did not provide interpretations. One of these well known intertextual relations Borgstedt deals with is the scene with the horses in the pigsty. The scene is said to be reminiscent of the scene in Heinrich von Kleist's *Michael Kohlhaas*, where Kohlhaas' horses were ill-treated and housed in a pigsty. The relation has long been described as a taking up of the motive of horses in a pigsty. This can be recorded straight forward:^[27]

68

```
@prefix ex: <http://example.com/> .
@prefix : <https://intertextuality.org/abstract#> .
@prefix g:
<https://intertextuality.org/extensions/genette/hypertextuality#> .
@prefix m: <https://intertextuality.org/extensions/motives#> .
@prefix pt: <https://intertextuality.org/extensions/genette/paratext#> .

ex:itr1 a :IntertextualRelation ;
      :here ex:landarztPigsty ;
      :there ex:kohlhaasPigsty ;
      :specifiedBy _undefSpec ;
      :mediatedBy m:HorsesInAPigsty .
ex:landarztPigsty oa:hasSource
  <https://textgridlab.org/1.0/tgcrud-
public/rest/textgrid:qn00.0/data>;
  oa:hasSelector
  [a oa:TextQuoteSelector;
  oa:exact "zerstreut, gequält stieß ich mit dem Fuß an die brüchige
Tür des schon
seit Jahren unbenützten Schweinestalles. Sie öffnete sich und
klappte in den Angeln auf
und zu. Wärme und Geruch wie von Pferden kam hervor"]

ex:kohlhaasPigsty oa:hasSource
  <https://textgridlab.org/1.0/aggregator/text/textgrid:r0th.0> ;
  oa:hasSelector
  [a oa:TextQuoteSelector;
  oa:exact "so zeigte er mir einen Schweinekoben an ... in welchem
... ich nicht
aufrecht stehen konnte. ... Sie guckten nun wie Gänse aus dem Dach
vor." ] .

_undefSpec a :IntertextualSpecification .
```

But that is only the well known part of the relation. Borgstedt also adds a relation to Homer who's echo he thereby includes in Kafka's text. The analogy between horses and goats reminds him of the verse of the *Iliad*, where Automedon, with the horses of Achilles, plummets into the goats (Iliad XVII.460). We record this without writing down the text selectors from the Web Annotation ontology:

69

```
m:HorsesAndGoats a m:Motive.
```

```
ex:itr2 a :IntertextualRelation ;
  :here ex:kohlhaasPigsty ;
  :there ex:iliadAutomedonsPlummet ;
  :specifiedBy _undefSpec ;
  :mediatedBy m:HorsesAndGoats .
```

Finally, there also is the transitive intertextual relation from Kafka to Homer, at least according to Borgstedt:

70

```
ex:itr3 a :IntertextualRelation ;
  :here ex:landarztPigsty ;
  :there ex:iliadAutomedonsPlummet ;
  :specifiedBy _undefSpec ;
  :mediatedBy ex:itr2 .
```

This shows that even complex chains of intertextual relations, chains of readers' reminiscences, can be recorded with the formalization presented here.

71

5. Conclusion and outlook

The goal of the current article was not to offer an algorithmic detector of intertextual relations in corpora of texts, but to suggest a theory- as well as domain-driven formalized category system that allows intertextuality researchers to annotate relations between text passages, works, or systems in a structured, general as well as specific and expandable way. The formal core we developed is based on predicate and description logic and was modeled with regards to the Web Annotation Data Model. This allows for (a) connections to web-based RDF knowledge graphs and thus possibilities of automated reasoning on the level of categories, i.e. the TBox in semantic web vocabulary, as well as (b) extensions towards specific theories of intertextuality which usually come with a certain definition of the notion of "text." Reasons for this form of modeling the concept were computability, expressiveness, and efficiency on one side, and the impression on the other side that the global applicability and the allowance of unary and binary predicates of description logic make it a suitable candidate for the formalization of a concept that genuinely deals with relations between at least two objects. Two main conceptual aspects of what we framed as the core of intertextuality are the directedness of intertextual references and with it the inherent chronologic dimension: later texts refer to earlier texts. Leaving out the chronology of intertextual references (e.g. texts being related by a common motive, topic or structure independent of their temporal situatedness) would make it impossible to have a "here" and "there" and the relations would not be directed anymore. As a consequence, the whole formalization would have to be transformed quite fundamentally.

72

The suggested formalization serves as a starting point for further research and elaboration: theorists of intertextuality on the one hand can sharpen their concepts through it by further refining and specializing the proposed model or by clearly delineating their own concepts. To have a core of intertextuality as a common denominator of intertextuality theories moreover is able to connect the different theories and shed light on possible paths between them. Practitioners of intertextuality on the other hand are provided with an (extensible) model to identify intertextuality in a structured way and to analyze it (e.g., by means of network/graph visualization and analysis, or by a synoptic comparison of text passages with intertextual reference).

73

The theoretical foundation should subsequently serve as the basis for the architecture and design of a research environment that allows researchers both to further develop or adapt the model in specific intertextuality research projects as well as to explore intertextual relations in a corpus that has been annotated on the basis of this model. We believe that a core formalization of intertextuality has great potential—once it will be embedded into a low-threshold work environment—to be actually taken up by many researchers worldwide so that a global knowledge graph of manually selected and established intertextual relations can emerge as a result of a shared research task.

74

Notes

[1] Cf. e.g. TextPAIR (<https://github.com/ARTFL-Project/text-pair>; access: October 24., 2022) or Tracer (<http://www.etrp.eu/research/tracer/>; access: October 24, 2022).

[2] Cf. <http://radimrehurek.com/gensim/intro.html> (access: October 14, 2022).

[3] *Situation theory* was initiated in “Situations and Attitudes” by [Barwise and Perry 1981] and further elaborated as *situation logics* (including the notion of *infor*) in “Logic and information” by [Devlin 1991].

[4] Cf. <https://github.com/BOberreither/INTRO> (access: December 19, 2023, commit 46cfd12).

[5] As Leibniz’s famous phrase goes: “[...] if controversies were to arise, there would be no more need of disputation between two philosophers than between two calculators. For it would suffice for them to take their pencils in their hands and to sit down at the abacus, and say to each other: *Calculemus.*” [Leibniz 1965, 200]; the translation has been adopted from https://en.wikiquote.org/wiki/Gottfried_Leibniz.

[6] Whereby the difference is not in the range milliseconds, but between a few milliseconds and hours.

[7] The well-defined semantic of predicate logic is the so-called model-theoretic Tarski Semantic. It should be mentioned, though, that modeling natural language in predicate logic is not the only approach to provide it with a well defined semantic. There have been various formalisms developed with that goal in mind. One of the most prominent and elaborated is the Montague grammar. However, these formalisms are too elaborated in the sense that they are not tractable for computation.

[8] As worked out in Alan Turing’s groundbreaking article *On Computable Numbers, with an Application to the “Entscheidungsproblem”* in 1936.

[9] For this correspondence see [Baader et al. 2003].

[10] This statement says that there exist two texts, a and b , and a is part of b . \wedge is the “and” operator in logic.

[11] This statement says that if a is part of b , and b is related to c , then a is also related to c . \forall (pronounced “for all”) is a quantifier in logic.

[12] Cf. latest W3C recommendation: <https://www.w3.org/TR/owl2-rdf-based-semantics/> (access: December 05, 2023).

[13] For a list of knowledge base systems see e.g.: <https://www.w3.org/wiki/SparqlImplementations> and reasoners see e.g.: <http://owl.cs.manchester.ac.uk/tools/list-of-reasoners/> (access: December 05, 2023).

[14] See [Frege 1879] and [Frege 1893/1903]. An introductory example for the functional notation is the formal analysis of the sentence “John is happy.” It is an assertion about the individual John using the general pattern “(x) is happy,” which is a unary function. In functional notation the sentence thus is written as $H(j)$ or $happy(\text{John})$. See [Zalta 2022].

[15] The anti-chronological assumption even holds true in cases, where it is undecidable which text comes earlier. We further acknowledge – but are for the current philological case not interested in – other possible worlds, such as contra-factual or imaginary ones, in which a future text influences a prior one.

[16] There exist various alternative notations even within RDF: RDF/XML, JSON-LD, Turtle, Manchester Syntax, Functional Syntax etc. We will opt sometimes for the so-called turtle syntax as it is one of the best human readable or JSON-LD as it is very common nowadays e.g. in the Web Annotation Data Model.

[17] The rationale behind this convention is the corresponding best practice in OWL.

[18] The mathematically less inclined reader may skip this formalization without losing the basic idea of constraint formulated in the sentence above.

[19] Such formalizations can be found in standard introductions to mathematics, see e.g. https://en.wikipedia.org/wiki/Finitary_relation (access: December 4, 2022).

[20] Cf. <https://www.w3.org/TR/annotation-model/> (access: December 2, 2022).

[21] The INTRO ontology mentioned above is rather restricted in this respect. With selectors, intertextuality can be represented as a network of

texts instead of “only” as a network of excerpts. One of the strengths of INTRO, however, is its integration into foundational ontologies such as CIDOC-CRM and FRBR.

[22] For the sake of coherence, we do not adopt the JSON-LD notation used in the specification of the Web Annotation Data Model, but translate them to RDF-Turtle notation. In fact the specification does not enforce the usage of JSON-LD, but just prefers this format as serialization to show examples of the model. For our purpose we prefer RDF-Turtle notation. Note that both are just serialization formats of RDF.

[23] Cf. <https://www.w3.org/TR/annotation-model/#web-annotation-principles> (access: December 04, 2022).

[24] <https://github.com/janhorstmann/intertextuality> (access: December 04, 2022).

[25] The example stems from a non-digital analysis of intertextuality in literary studies by [Borgstedt 2010]; cf. section 4.5, too.

[26] In description logic it is common to call the set of all statements about individuals (e.g. instances of a class) *ABox* and those about classes and properties *TBox*. In less formal contexts “ontology” means “TBox.”

[27] We use a quick and dirty hack to give a specification for the intertextual relation, which we cannot really define (`_undefSpec`) due to lack of information in [Borgstedt 2010]. If Borgstedt had used our ontology, he would have had to be more explicit about the type of the relation.

Works Cited

- Adamzik 2004** Adamzik, Kirsten. (2004) *Textlinguistik. Eine einführende Darstellung*. Tübingen: Niemeyer.
- Baader et al. 2003** Baader, Franz et al. (2003) *The Description Logic Handbook*. Cambridge University Press.
- Barwise and Perry 1981** Barwise, Jon, and John Perry. (1981) “Situations and Attitudes.” In *Journal of Philosophy*, 668–691. URL: <https://www.jstor.org/stable/2026578> (access: October 11, 2022).
- Borgstedt 2010** Borgstedt, Thomas. (2010) “Kafkas Kubistisches Erzählen. Multiperspektive Und Intertextualität in Ein Landarzt.” In *Kafka Verschieben*, edited by Irmgard Wirtz, 53–96. Göttingen, Zürich: Wallstein, Chronos.
- Burghardt and Liebl 2020** Burghardt, Manuel, and Bernhard Liebl. (2020) “‘The Vectorian’ – Eine Parametrisierbare Suchmaschine Für Intertextuelle Referenzen.” In *DHd 2020 Spielräume: Digital Humanities zwischen Modellierung und Interpretation. 7. Tagung des Verbands “Digital Humanities im deutschsprachigen Raum” (DHd 2020)*, Paderborn. DOI: 10.5281/ZENODO.4621836.
- Bär et al. 2012** Bär, Daniel, Torsten Zesch, and Iryna Gurevych. (2012) “Text Reuse Detection using a Composition of Text Similarity Measures.” In *Proceedings of COLING 2012*, 167–184, Mumbai, India. The COLING 2012 Organizing Committee. URL: <https://aclanthology.org/C12-1011> (access: October 11, 2022).
- Bär et al. 2015** Bär, Daniel, Torsten Zesch, and Iryna Gurevych. (2015) “Composing Measures for Computing Text Similarity.” Report. Darmstadt, Germany. URL: <https://tuprints.ulb.tu-darmstadt.de/4342/> (access: October 11, 2022).
- Devlin 1991** Devlin, Keith. (1991) *Logic and Information*. Cambridge University Press.
- Dosse 1999** Dosse, François. (1999) *Geschichte des Strukturalismus*. 2 Volumes. Frankfurt/M.: Fischer.
- Frege 1879** Frege, Gottlob. (1879) *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle: Nebert.
- Frege 1893/1903** Frege, Gottlob. (1893/1903) *Grundgesetze der Arithmetik*, vol. I and II. Jena: Pohle.
- Genette 1982** Genette, Gérard. (1982) *Palimpsestes: la littérature au second degré*. Collection Poétique. Paris.
- Georges 1913** Georges, Karl Ernst. (1913) Ausführliches lateinisch-deutsches Handwörterbuch. Hannover: Hahnsche Buchh., Reprint 1998. URL: <http://www.zeno.org/georges-1913> (access: Jan. 29, 2024).
- Janich 2019** Janich, Nina. (2019) “Intertextualität und Text(sorten)vernetzung.” *Textlinguistik. 15 Einführungen und eine Diskussion*. Ed. by Nina Janich. 2nd ed. Tübingen: Narr Francke Attempto. 169-188.
- Kristeva 1986** Kristeva, Julia. (1986) “Word, Dialogue and Novel.” *Julia Kristeva. The Kristeva Reader*. Edited by Toril Moi. New York: Columbia UP, 34–61.
- Leibniz 1965** Leibniz, Gottfried Wilhelm. (1875–1890) *Die philosophischen Schriften von Gottfried Wilhelm Leibniz*, volume 7. Berlin: Weidmann. Reprint Hildesheim, New York: Olms 1965.

Leitch 1983 Leitch, Vincent B. (1983) *Deconstructive Criticism: An Advanced Introduction*. Hutchinson University Library. London: Hutchinson.

Link 1997 Link, Jürgen. (1997) *Literaturwissenschaftliche Grundbegriffe*. 6th Edition. München: Fink.

Oberreither 2020 Oberreither, Bernhard. (2020) "Zwei Überlegungen zur Konzeption einer Linked-Data-Ontologie für die Literaturwissenschaften." In *Digital Humanities Austria 2018. Empowering Researchers*, edited by OEAW, 134–39. University of Salzburg, Austria: DOI: 10.1553/dha-proceedings2018s134.

Pfister 1985 Pfister, Manfred. (1985) "I. Konzepte der Intertextualität." In *Intertextualität*, edited by Ulrich Broich and Manfred Pfister. Berlin, Boston: de Gruyter, 1–30. DOI: 10.1515/9783111712420.1.

Scheirer et al. 2016 Scheirer, Walter, Christopher W. Forstall, and Neil Coffee. (2016) "The Sense of a Connection: Automatic Tracing of Intertextuality by Meaning." *DSH Digital Scholarship in the Humanities* 31.1, 204–217. DOI: 10.1093/llc/fqu058.

Schlupkothén and Nantke 2019 Schlupkothén, Frederik, and Julia Nantke. (2019) "FormIt: Eine multimodale Arbeitsumgebung zur systematischen Erfassung literarischer Intertextualität." In *DHd 2019 Digital Humanities multimedial und multimodal. 6. Tagung des Verbands "Digital Humanities im deutschsprachigen Raum" (DHd 2019)*, Frankfurt am Main and Mainz. DOI: 10.5281/zenodo.4622106.

Ternès 2016 Ternès, Anabel. (2016) *Intertextualität*. Wiesbaden: Springer Fachmedien. DOI: 10.1007/978-3-658-12792-3.

Turing 1937 Turing, A. M. (1937) "On Computable Numbers, with an Application to the Entscheidungsproblem." In *Proceedings of the London Mathematical Society*. Series 2.

Zalta 2022 Zalta, Edward N. (2022) "Gottlob Frege," *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.). URL: <https://plato.stanford.edu/archives/fall2022/entries/frege/> (access: October 11, 2022).



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.