Journal of Open Humanities Data

Representing Persons and Objects in Complex Historical Events using the Event Model F

DATA PAPER

GERARD KUYS
ANSGAR SCHERP
*Author affiliations can be found in the back matter of this article

ubiquity press

ABSTRACT

The digital representation and publishing of human history on the web has so far been stuck at the digital unlocking of collections of historical items. Those collections are described by metadata mostly for curation and for findability metadata that are put on the web in the manner of a library catalog. Usually, there is little on 'aboutness'. This is unfortunate, as modern representations of knowledge and web-based interactive presentation techniques offer ample opportunities for a more complex representation and richer interaction with digital history. In this contribution, we argue for a history in digital data that is treated for what it is: an interpretation, like all history, while remaining traceable to its information carriers.

CORRESPONDING AUTHOR:
Ansgar Scherp
Data Science and Big Data Analytics, Ulm University, Germany
ansgar.scherp@uni-ulm.de

KEYWORDS:
data modelling; events in history; linked data vocabularies

TO CITE THIS ARTICLE:
Kuys, G., & Scherp, A. (2022). Representing Persons and Objects in Complex Historical Events using the Event Model F. *Journal of Open Humanities Data*, 8: 22, pp. 1–12. DOI: https://doi.org/10.5334/johd.84

(1) CONTEXT AND MOTIVATION

A good knowledge representation approach should address a two-pronged aim. First, the complete and precise capturing of the knowledge in the use case under consideration. Second, complying with the obligation to provide flexibility when adding information or interpretations to the existing knowledge. While this is reasonable and frequently a requirement for knowledge representation, it is notoriously difficult to satisfy. The reason is that the knowledge representation approach has to lean on important domain information such as places, persons, roles, and to establish this information as "common-ground knowledge" about reality, even though not all of the information is undisputed. Today, important information such as places, persons, or roles is typically represented in relatively simple domain ontologies which are based on Linked Data principles. However, in order to express complex knowledge such as historical events, data publishers (as well as everyone using their data) need to have recourse to ontologies that allow them to represent historical events with all their sub-events, participants etc. Equally important is the requirement to express equivocality of sources or different, if not conflicting, interpretations of the same event. And, finally, if we as data publishers choose to document events from all kinds of sources, we must ensure that every type of source description (vocabularies on bibliographic, archival, iconic materials) has a feature to point to the same description of the subjects covered, if necessary, by adding a dcterms:subject property. This set of shared descriptions is the area where perspectives should meet, which we call the "common ground". In this common ground there may not be much information, but a person is granted to be a particular individual, as is also the case with relations, events and designations of time and place.

(2) APPROACH

We argue that for complex historical knowledge, we should rely on ontologies based on the Web Ontology Language (OWL).[1] While Linked Data and OWL share a foundation both conceptually and technology-wise, combining them into a common event model typically engenders a dilemma. This is the conundrum of how to reconcile canonical representations of a past state of affairs with varying interpretations, or with layered accounts of complex events. Yet, this can be done. In the Descriptions and Situations (DnS) pattern of the foundational ontology DOLCE, we believe to have found a key knowledge representation mechanism. If an ontology may be defined as "a way of showing the properties of a subject area and how they are related, by defining a set of concepts and categories that represent the subject",[2] DOLCE stands out for its epistemology-based approach (which is why it is a foundational, or "upper", ontology), and especially for its treatment of the time aspect. Over the last 15 years, DOLCE has evolved into DOLCE+DnS UltraLite (DUL),[3] adapting to the OWL and Linked-Data paradigms on the internet.

The definition of events in DUL is all but rigid: an event is an accomplishment, an achievement, a punctual event or a transition that always is related to an observable situation. In DUL, an event can have different aspects at the same time, but, for a time-oriented point of view, events that are classified by people or things that participated in it may be expected to be the most relevant.[4] The inherent fluidity of events in DUL is expressed by the DnS pattern. DnS allows users to both exploit and combine Linked Data with an expressive representation of events. Out of a number of event vocabularies,[5] Event-Model-F (Scherp et al, 2009) seems to make the most of the expressiveness of DUL.

Event-Model-F has grown out of the need to provide "comprehensive support to represent time and space, objects and persons, as well as mereological, causal, and correlative relationships between events" (Scherp et al, 2009). For each of these relationships, "F" provides a dedicated

1 https://www.w3.org/OWL/, last accessed date: 03/08/2022.

2 https://en.wikipedia.org/wiki/Ontology_(information_science), last accessed date: 03/08/2022.

3 http://ontologydesignpatterns.org/wiki/Ontology:DOLCE%2BDnS_Ultralite, last accessed date: 03/08/2022.

4 http://www.ontologydesignpatterns.org/ont/dul/DUL.owl s.v. dul:Event, last accessed date: 03/08/2022.

5 A vocabulary is a system of concepts for arranging the properties of a particular subject area in an ontology, in this case: events.

DOI: 10.5334/johd.84

object design pattern. Finally, we extend our model with Time-indexed Participations[6] to do one more thing: introducing the time dimension into the course of actions. Time-indexed Participations allow us both to capture the nature of people's involvement in some event (roles), and to express a certain level of evolution in that involvement.

(2.1) MODELLING COMPLEX HISTORICAL EVENTS

Frequently, digitising the past is considered on a par with publishing content about cultural heritage collections on the internet. There is a friction here, though. Often, digitization projects in the cultural heritage field report that a disproportional amount of time has to be spent on the implementation of metadata for curation or publication rather than on the metadata that describes the historical situation itself. Partly, this may be ascribed to lacking experience in expressing events as data. Partly, it is because practitioners in cultural heritage projects are better at digitizing administrations than at representing real life. However, real-life events should be expressed in data to support reusability of historical knowledge. Data structures help to objectify descriptions (foreclose wildly varying interpretations by making definitions part of the data structure), to recognize the stable elements in an otherwise complex situation, and to make them better exchangeable. Evocating – that is, (re)constructing what happened by way of verifiable statements – the past in digital data on the internet should amount to telling a story or providing an analysis, the credibility of which would rely on the author's handling of the materials underpinning his or her evocation of the past. In most cases, the focus would be on the coherency of the constructed narrative, not on the carriers from which its elements came to us. The requirement to document one's evocation is a matter of methodical discipline, but it is not by itself the story to be told, nor the analysis to be presented. We believe that cultural heritage should record personal and societal accomplishments and failures, and the circumstances conditioning these. If cultural heritage were only about the material evidence that happens to exist to this day, it would have the characteristics of an art market and derive its value from rarity. On the contrary, a digital platform on history should rather help us to verify materials, assist its users in putting contentions to the test, and in suggesting relationships between phenomena or developments.

We believe that historiography is a "discussion without an end" (Geyl, 1956, pp. 59–60), and therefore that digital renderings of the human past must be structured along requirements fostering such a discussion. Digital history precludes, in our view, the predominance of the private hermeneutics of any describing person involved. The best way to make assertions transparent is to develop data structures that explicitly refer to complex phenomena like temporal perspectives and conflicting perceptions. At least, this is the broader perspective we are after.

We do realize that data-oriented reconstructions of past events are well beyond present-day practice, and that there is still the greater part to be learned when developing such an approach. This is not a reason, though, to stop there. To explore the issues involved with data-bound representations of the past, we apply what we hold to be powerful ways of modelling to a real-world example. By tackling an ancient work of historiography, we can be sure to be confronted with perspectives which are bound to differ from our own.

(2.2) CONTRIBUTIONS OF THIS WORK

We analyse the requirements of a complex data representation of events, one that covers common-ground knowledge but also leaves space to represent different and potentially contradictory interpretations. We show that the DnS-based patterns of Event-Model-F provide conceptual facilities to weave in common-ground knowledge into event representations in Linked Data. Using the DnS pattern, we apply Event-Model-F to events in the history of the Netherlands.

6 http://ontologydesignpatterns.org/wiki/Submissions:Time_indexed_participation, last accessed date: 03/08/2022.

(3) DATASET DESCRIPTION

We developed a publicly available prototype knowledge base which aims to be a foundation of further developments.

Dataset access point: https://data.pldn.nl/DBpediaAssociation-NL/PiOCHE/.

Dataset name: https://data.pldn.nl/DBpediaAssociation-NL/PiOCHE/graphs/default-1.

DOI: https://zenodo.org/record/6641891.

Format names and versions: Triply RDF triplestore, searchable by SPARQL or by ElasticSearch.

Creation dates: 2020-06-10 to 2022-05-12.

Dataset creators: Gerard Kuys (data curation, queries) and Gerald Wildenbeest (accessibility).

Language: Dutch.

License: CC-BY-SA.

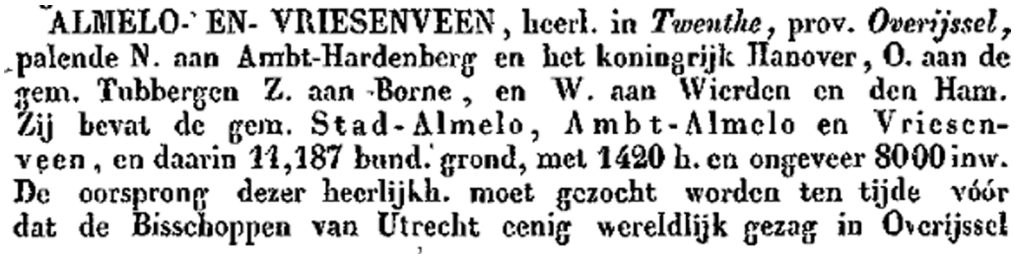
Repository name: Platform Linked Data Nederland, dataset PiOCHE.

Publication date: 2022-06-1.

Dataset size: 345.536 statements, covering three geographical regions.

(4) METHOD: STRUCTURING HISTORICAL NARRATIVES AS LINKED DATA

The basis of our project is a 19th-century *Geographical Dictionary of the Netherlands* (published 1839–1851). Its author, A. J. van der Aa, intended his work to be an overview of "all places (...) for any reason meriting to be mentioned" (Van der Aa, 1839, p. vi). This *Geographical Dictionary* (which would be called an *Encyclopedia* rather than a *Dictionary* in present-day language, as it lacks definitions), came to us in 13 volumes containing 11,146 total printed pages.



ALMELO- EN- VRIESENVEEN, heerl. in *Twenthe*, prov. *Overijssel*, palende N. aan Ambt-Hardenberg en het koningrijk Hanover, O. aan de gem. Tubbergen Z. aan Borne, en W. aan Wierden en den Ham. Zij bevat de gem. Stad-Almelo, Ambt-Almelo en Vriesenveen, en daarin 11,187 bund. grond, met 1420 h. en ongeveer 8000 inw. De oorsprong dezer heerlijkh. moet gezocht worden ten tijde vóór dat de Bisschoppen van Utrecht eenig wereldlijk gezag in Overijssel

A.J. van der Aa on the Seignory of Almelo-and-Vriesenveen, Vol. 1 (1839).

This *Dictionary* was digitized and made searchable through Optical Character Recognition (OCR) in a separate project more than 10 years ago. The reason for making the *Dictionary* our principal source is that this is a work from the first half of the 19th century, which challenged us to work out the problem of how to express a time-bound perspectives in digital data.

The logical structure of our dataset is straightforward. It is about Abstracts providing Descriptions in a Volume of a Work. Each of these Descriptions describes a geographical feature. When the feature is a populated place, the author entertains the reader with a historical narrative about the location being discussed. This structure results in a dataset containing the entities Person for the author, Work, Edition and Abstract for the description itself and the entities ArchitecturalStructure, Place, LocationType and Jurisdiction for the spatial feature that the narrative is about. There is no history yet, but we now have a typology of Locations. Thus, we may compare the nature of the localities mentioned by Van der Aa to places we know about from other sources.

The diagram in Figure 1 presents a real-world location described by two time-bound interpretations. There is a division here between some part of the real world being described on one hand (right hand of the red diagonal, the De Haar castle near the town of Utrecht with URI AaPlts00516), and the actual description(s) on the other hand. The first description discloses the describing agent (author A.J. van der Aa), the published work, the relevant volume thereof, the year of publication (1849), and the abstract or lemma itself (upper left hand of the red

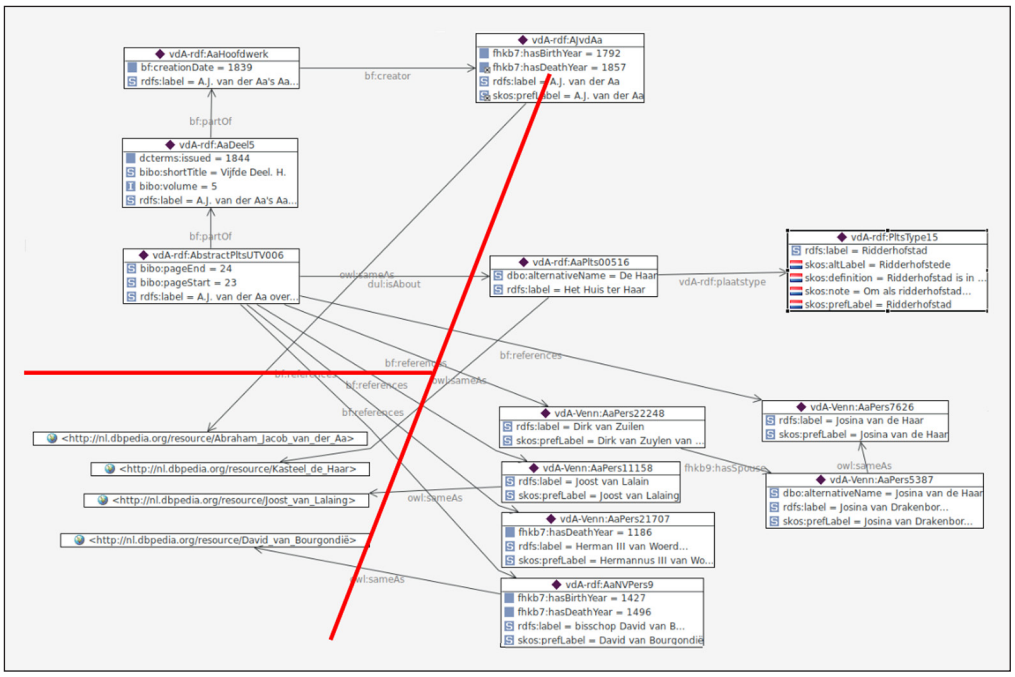


Figure 1 Kasteel de Haar in two descriptions, including people named in the narrative.

diagonal, the *Geographical Dictionary*). The other description of Diagram 1 is in present-day Wikipedia (or its datafication DBpedia). It has an article on De Haar, as well as biographies of some people who have lived there (the URIs of those are at the lower left hand of the red diagonal).

Both descriptions could easily be supplemented by other descriptions found in another source. The *Geographical Dictionary* being bibliographic material, we used the *BibFrame* vocabulary[7] to express its set-up in the data. Would another description base itself on an archival document, the vocabulary would no doubt have been the emerging archival standard, the ontology of *Records in Context*.[8] Apart from this, the physical limit to the number of sources is in the modeler's capacity to process the information that comes with every additional source.

The red diagonal in Diagram 1 separates the Description (left-hand) from what the Description is about (right-hand). At some point in the description process, we have classified the *Geographical Dictionary* as a multiple-volume work in the Non-Fiction genre. Consequently, we assume that each Abstract in the *Dictionary* provides information on (the 19th-century state of) a geographical object that may be known to us via other paths as well. Our second source, Wikipedia, may therefore provide additional information. After 1911, De Haar became famous for having been rebuilt in a grandiose Viollet-le-Duc manner, paid for by a descendant of the Rothschild family. In van der Aa's time, however, the castle was almost in ruins. Here, two sources lead to two diverging descriptions of the object concerned.

In volume 12 of his *Dictionary*, A.J. van der Aa informs us about the people who in the course of time were involved with De Haar. Based on this narrative, we can add Persons, playing a role within the context of De Haar as a location, to our data. In the right-hand section of Figure 1, we see those persons identified by one or more sources. However, there is still no way of telling what the nature of these persons' relation to De Haar is, nor by what kind of event it came about. Yet, such information may be in van der Aa's narrative, or elsewhere. In order to capture this kind of information we must integrate Event-Model-F into our data structure.

(5) MODELING EVENTS: WORKING WITH EVENT-MODEL-F

The DOLCE+DnS framework (DUL) provides the features that are vital for the layered and time-dependent approach we want to build our data structures on: time-bound relations, and shifts in perspective. Shifts that can be captured quite well by the DnS bipolarity of Situations that need to match a Description. DUL, moreover, offers a balanced model for Agents, Roles and

7 https://loc.gov/bibframe/, last accessed date: 03/08/2022.

8 https://www.ica.org/en/records-in-contexts-conceptual-model, last accessed date: 03/08/2022.

DOI: 10.5334/johd.84

Events that is more expressive than event model SEM (van Hage et al, 2011),[9] more fine-grained than SEM and less collection-oriented than CIDOC-CRM.[10] The main instrument in the DnS pattern to deal with both complexity and conflicting perspectives is a conceptual framework for ordering notions about social situations, among others. This framework has been designed in the following way. Events — in texts identifiable by the occurrence of verbs like "happen", "do", etc. — need to be arranged in Situations, and the latter need to satisfy the view expressed in a Description. In this way, we end up with different configurations of Events (which may be the same events) in Situations. Of these, each and every Situation must comply with the leading perspective of its Description.

With such modelling potential, it makes sense to build modelling events on top of this framework, and extend it with Event-Model-F by Scherp et al. (2009). Event-Model-F defines seven patterns. A pattern models how people or things participate in events (Participation pattern), while others show how one event may be part of another event or set of events (Mereology pattern), how one event is the cause of one or more other events (Causality pattern), how one event relates to another (Correlation pattern), and how descriptions of events may be related to particular views (Interpretation pattern). Since our data model of van der Aa's *Dictionary* leaves every possible room for multiple sources, the Interpretation pattern of Event-Model-F is of central importance. Sources are there to make assertions traceable to the description they are derived from. However, the relative weight of an assertion depends on the credibility of its source, on the relevance to the current perspective, and on the quality of the correlations to other events. Therefore, we cannot treat the descriptions in sources as if they were exchangeable at will. There must be a corpus of acknowledged historical facts to refer to, a touchstone both for a statement by A.J. van der Aa and for the one by an additional source. This is the "common ground" knowledge we introduced in section (1).

Event-Model-F refers to precisely such a "common ground", naming it the "Domain Ontology". It figures in several patterns of "F", and is represented by the large cloud on the left in the diagram in Figure 2. This "F" Domain Ontology is the foothold in reality that is the foundation for any interpretation in any time window. For this reason, we must be strict on private interpretations. If a non-standard reading of some situation might require non-standard data in order to be made logically consistent, there ought to be a separate space for them, such as the smaller clouds in Figure 2. Data underpinning any private interpretations by van der Aa (or by others) should be confined to an RDF graph or namespace of their own.

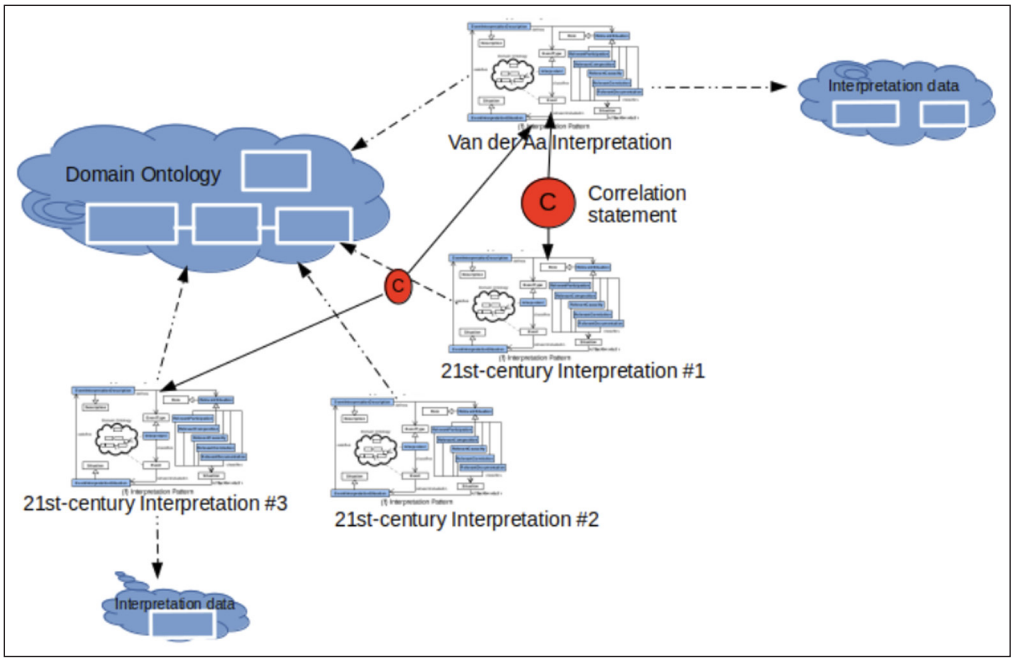


Figure 2 The "F" Interpretation Pattern applied to multi-temporal interpretations.

9 SEM was designed in the first place to meet the requirements involved in recognizing events in possibly oblique internet content.

10 CIDOC-CRM (see https://cidoc-crm.org/, last accessed date: 03/08/2022), defines itself as "a theoretical and practical tool for information integration in the field of cultural heritage". For how we chose our event model, see the document 2_PersonLink How to deal with events.pdf at https://data.pldn.nl/DBpediaAssociation-NL/PiOCHE/assets, last accessed date: 03/08/2022.

In Figure 2, every Interpretation points to the Domain Ontology. The smaller clouds with Interpretation Data each must not contain any other data than required to sustain the concomitant interpretation. They convey information that is not consistent with the information in the Domain Ontology. We see four instances of the "F" Interpretation Pattern, of which one is from the 19th century (the *Geographical Dictionary*), and three of our own time. Apparently, there are slight differences between the description by van der Aa, on one hand and the 21st-century interpretations #1 and #3 on the other. The instances of the Correlation Pattern of Event-Model-F (the "C" in the red circle) may have been used to document how an object in the first description still corresponds to a certain present-day object in one of the other two.

(6) USE CASES

In order to demonstrate the applicability of the DnS pattern and of Event-Model-F, we must now identify the use cases suitable to underscore the potential of this kind of data modelling. The point to be made is that, by using the DnS pattern and "F", we can capture events that as data structures can stand on their own when compared to an inventory or a catalog.

The principal source in this project, A.J. van der Aa's *Geographical Dictionary*, has plenty of event descriptions. Several of those are challenging, not only due to a meandering course, but also because of their implicit or explicit causality. For the purpose of accommodating possibly conflicting views, we have to dispose of a data structure allowing to capture and express different perspectives between sources, or even conflicting perspectives within one and the same source document. At least we should be able to differentiate between van der Aa's 19th-century perspective and more or less accepted readings of the same historical event today as may be found in sources such as Wikipedia. For the example of seven events described in the *Geographical Dictionary*, this requirement would lead to the view in Table 1.

In Table 1, the distinction between the readings by different sources might suggest that there can be no common notions between a 19th-century perspective and one of our own times. This is not the case. What any perspective shares with another is the presence of a geographical

Table 1 Use cases by source perspective and by causality type.

USE CASE	LOCUS	A.J. VAN DER AA'S READING	WIKIPEDIA READING	CAUSALITY TYPE VAN DER AA	CAUSALITY TYPE WIKIPEDIA
1	Vol. 11	The catching of a whiting (a salt-water fish) just before the town walls of Utrecht during the All-Saints Flood of November 1–2, 1170	High and low tide near Utrecht, a pout whiting was caught	Flood	Flood
2	Vol. 1	Six persons named to be bom into the Almelo dynasty of counts, 12th – midst 14th century	Four persons named to be born into the Almelo dynasty of counts, 12th – midst 14th century	No male heir	Marriage policy no male heir
3	Vol. 1	Hendrik van Almelo joins alliance of regional nobles against bishop-elect Hendrik van Vianden, 1249–1250	–	Political strife of feudal nobles against their liege lord	–
4	Vol. 5	The beheading of the nobleman Frederik III Uten Ham, a vassal to the Utrecht bishop, in the town of Utrecht, April, 16, 1482	–	Political factions in a late medieval town	–
5	Vol. 7	The death of Steven van Zuylen van Nijeveld, land commander of the Teutonic Order in Utrecht, in 1527. Van Zuylen was killed as a group of soldiers invaded his manor house in Maarssen. They did so after one of their companions had been shot on Van Zuylen's orders while seen angling in the moat	–	Tragic incident	–
6	Vol. 3	The replacement of the catholic parish priest by a preacher of calvinist persuasion in the town of Diepenheim, 1601–1618	–	Local patriciate tending towards local autonomy, the regional powers insisting on the proper doctrines of preachers	–
7	Vol. 11, vol. 2	Skirmishes and cannonades in the vicinity of Utrecht, between Patriot and Orangist troops, in May and July, 1787	General account of defeat Patriot movement, 1787	Military suppression of civil uprising	Military suppression of civil uprising

space that, although there is change in that landscape, remains in essence the same over time. Given that, on a time scale of approximately 180 years, a location changes certainly in outlook, sometimes also in character, and in rare cases even in geographical position (a place may be rebuilt elsewhere after a flood or a sand drift), we must find a way to transmit this evolution to our reference data, that reference data often pointing to our present-day image of this place. The evolution itself is signalled by the red "C" in Diagram 2, the Correlation statements of Event-Model-F, stating that interpretations have drifted apart in the course of time. But evolution also has had its influence on the quality of the participating entities. For the purpose of introducing timelines for all entities partaking in an event, we took recourse to a modelling technique that records just as much information as is needed to express a state of affairs at a certain point in time. This technique is known as "snapshots" (D'Souza & Wills, 1998, pp. 50–52). Snapshots are used to express that a change has occurred in a particular situation. In so doing, they introduce the aspect of movement in what would otherwise risk to be a static model. And movement we need in order to stress the fact that a person, situation or place cannot be considered to have remained exactly the same over time.

For the sake of identifying use cases, we have singled-out seven events within their source text and placed them into the framework of Table 1. We classified the attributed causes of these events on the basis of a provisional typology of causality. The next thing to be done is to model the unfolding of the event itself and to be specific about how one event may relate to another. To this end, we will elaborate on three use cases out of our list of seven. Use case 6 may be deemed complex enough to show how to handle events intermingling one with the other, whereas the Birth, Marriage and Death events implied by the genealogy of use case 2 compel us to deal with an incomplete description. In order to deal with conflicting views on a state of affairs, we conclude with use case 3.

(7) MODELING EVENTS: EXAMPLES

Section (7.1) serves as an exercise in modelling an event of more than average complexity. Section (7.2) is about how to respect, but put into proper context, an imprecise description. Section (7.3) discusses how to deal with contradictions and incongruent readings of a social situation.

(7.1) COMPLEXITY: INSTALLING A CALVINIST PREACHER IN DIEPENHEIM, 1601–1618

In the third volume of his *Dictionary* published 1841, A.J. van der Aa discusses the Diepenheim parish priest who, in 1601, refused to cross the line to the new Calvinist civil order and was sent away. In search of a Protestant preacher to take his place, the local patriciate failed to find a livelihood for the forthcoming clergyman. Apparently, there were no monasteries to be sold in Diepenheim. For a while, Diepenheim shared a preacher with the neighbouring town of Goor, but the Diepenheim patriciate wanted a minister of their own. Over this, they fell out with the regional authorities (both worldly and religious), who had sanctioned the solution of a single parson for two local churches.

The search for a livelihood was interrupted when, in August 1605, the nearby town of Oldenzaal was conquered by the Spanish army. In the countryside, panic spread among the Calvinist preachers who fled the region. Only after the conclusion of the Twelve-Year's Truce with Spain in 1609, efforts to install a parson had a chance again. In 1616, Diepenheim churchmen had called upon Hugo Gallus, a preacher perhaps of Huguenot origins. Gallus, however, aroused the mistrust of the regional classis for alleged Remonstrantism. The approval required for his nomination was refused, and Gallus had to go.

It was not until 1618 that the Diepenheim patriciate called upon another clergyman, Elbertus Westenberg. But he, too, was suspected of Remonstrantism. This dispute was not so much about Remonstrantism as about pecking order. For, as soon as Westenberg had duly repented and admitted that the regional classis had had to be consulted, the classis withdrew its decision and approved of the nomination of Westenberg. The wordly authorities quickly followed suit. Westenberg was installed as a minister in Diepenheim, and stayed there until his death in 1624.

This chain of events we unravel into four consecutive episodes, each of which in Figure 3 is a pair of a Description and a Situation:

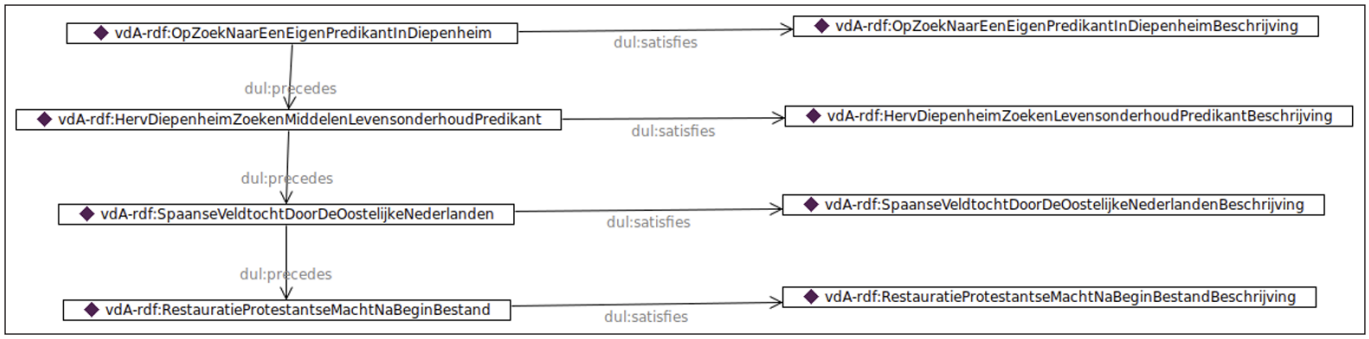


Figure 3 Time – a plain Correlation between Situations.

The diagram in Figure 3 uses a plain type of relation between the episodes: sequence in time. The order of events within one and the same Situation may be just time-bound. These would be either concurring or subsequent events, depending on the procedure observed or on a perceived course of events. In this use case, we might also want to identify more complex relations. Of those, there is one, causality, that stands out for having its own pattern in Event-Model-F.

When applying the causality pattern of Event-Model-F, the shift of responsibility for the livelihood of clergymen from the catholic church to the local patriciate would seem to be a major cause. Here, we may lay down two pairs of Cause and Effect, notably that:

A. If the catholic church is to no longer provide religious services, it is the local church community that pays for their Calvinist replacement.
B. Those of the local patriciate who have inherited or acquired the right to acclaim or veto the nomination of a clergyman should bear responsibility for granting a livelihood to the new preacher called under their jurisdiction.

In Event-Model-F, both of these pairs can be expressed as individuals of class f:Cause and class f:Effect, each a subclass of f:EventType. Like every event type in the patterns of "F", causes and effects are defined by the Description, that in this case is meant to clarify (a view on) the causes of a particular event.

(7.2) CONFUSION: REPRESENTATIONS OF A DYNASTY IN ALMELO

In the first volume of his *Dictionary* published 1839, van der Aa treats the medieval history of Almelo in a rather slipshod manner. His account of the House of Almelo, of how this lineage of counts rose and fell between the 12th and 15th centuries, is at odds with present-day knowledge. The latter is all but "common ground": on the website www.genealogieonline.nl, at least five deviating pedigrees on the House of Almelo may be found. Still we can construct our own "common-ground" genealogy on the basis of Wikipedia and of an additional source, a professional commentary written to introduce the archives of the House of Almelo (Historisch Centrum Overijssel, 1933/2022, archive 0214).[11]

If A.J. van der Aa's genealogy of the counts of Almelo is projected onto our own pedigree, we can clearly see that Van der Aa's interpretation is less precise. Whereas Van der Aa mentions a single latin-named Arnoldus ab Almelo, present-day genealogies claim that there were four of those. There is also a problem with van der Aa's contention, that Egbert, Arnold (II)'s son, had a heiress, Beatrix van Almelo, who, being married to a descendant of the van Hekeren family, transferred the rights over the Almelo seigniory to her husband's family. There were actually two Beatrices van Almelo, of whom only the youngest was married to a van Hekeren. The latter Beatrix, moreover, was born about the year 1340, almost 40 years after the death of Egbert. Either version must find its place within our model.

Since we chose to have the Domain Ontology of Event-Model-F reflect the present-day state of knowledge, the confronting view, in this case van der Aa's, must be modelled as an Interpretation. The Interpretation pattern of Event-Model-F defines an f:Event (which may or may not have taken place) as being classified by an f:Interpretant to the background (defined by the f:EventInterpretationDescription) of an f:RelevantSituation. It is against the latter that the classification of the event ought to be rated. As mentioned by the Interpretant,

11 *Heerlijkheid Almelo* (https://nl.wikipedia.org/wiki/Heerlijkheid_Almelo, last accessed date: 03/08/2022).

any situation that cannot be corroborated directly from the Domain Ontology is to be relegated to a separate set of Interpretation data like the ones shown in Figure 2. The events brought out by an Interpretation will be arranged, as in any DnS pattern, in a Situation. An f:EventInterpretationSituation distinguishes itself by accommodating (among others) events of a type that is marked to be part of an Interpretation.

Less precision is not the same thing as having a different perspective. In order to respectfully treat van der Aa's reading, we also may complement our own kinship statements with less circumscribed relations, like hasAncestor etc., in the dataset pointing to Van der Aa's interpretation.[12]

(7.3) CONTRADICTION: CONFLICTING REPRESENTATIONS OF POLITICAL STRIFE

A. J. van der Aa (1839) states that count Hendrik van Almelo associated with four nobles against the Utrecht bishop-elect Frederik van Vianen, in a year that must have been 1249 (Van der Aa, 1839, p. 101). The latter part of this contention is clearly incorrect. In the bishopric of Utrecht, there has never been a bishop with that name. What, then, do we actually know to be accepted as true?

The new bishop who was to succeed to the deceased bishop Otto van Holland (died 1249, uncle to count Willem II of Holland), was pivotal in the political direction the bishopric would take. Was he to toe the line of Holland's interests again, or would the local Utrecht patriciate have their way? Tellingly, the Utrecht church chapters elected one of their own to the bishop's see. This bishop-elect was named Gozewijn van Randerath but, because of chronicler Johannes de Beka, he became known to posterity under the name Gozewijn van Aemstel (Beka, ca. 1346, lines 4–5). Against Gozewijn van Randerath the Welfs (count Willem II was also Roman king and enjoyed full support by the pope) pressed the Utrecht chapters to counter-elect Hendrik van Vianden, whose name was mangled as well. Hendrik van Vianden came to be known, via "Henric van Vigennen" with Beka, as "Frederik van Vianen" with Van der Aa.

Hendrik van Almelo colluding to the benefit of "Gozewijn van Aemstel" is plausible. The counts of Almelo tended to seek shelter with the Utrecht bishop, their liege lord, as soon as any worldly competitor got too influential. If necessary, this could also be the other way round (Historisch Centrum Overijssel, 1933/2022, archive 0214). Judging by the company he joined, Hendrik's move was directed against the count of Holland this time, if only because his fellow count of nearby Goor had just recently been chased away by Willem II's uncle, Otto van Holland.

In the situation described by the narrative, we mark five episodes. To some of these episodes, interpretations differ. The first interpretation is by Johannes de Beka (ca. 1346), and another is by A.J. van der Aa (1839), generally following Beka's narrative. When making data structures out of these interpretations, the following contradictions have to be dealt with:

- There is no Gozewijn van Amstel, except as a character in dramaturgy. Gozewijn van Randerath was the Utrecht chapters' favourite after a spell of Holland-oriented episcopal activity;
- There is no Frederik van Vianen, either. In 1249–1250 near-namesake Hendrik van Vianden was pushed onto the scene, notably by his uncle, Konrad I, the archbishop of Cologne;
- The counts of Almelo, vassals to the Utrecht bishop, sometimes counterbalanced their liege lord seeking the backing of a rival grandee. But in 1249 regional nobles supported the anti-Holland line within the Utrecht bishopric;
- Although bishop Hendrik van Vianden had been consecrated due to the support of the Welf party in 1252, he repositioned against Holland as soon as he had assumed his full worldly powers.

12 We modelled family relations after the tree of personal relations in Robert Stevens' Family History Knowledge Base. Cf. https://robertdavidstevens.wordpress.com/2010/05/04/the-family-history-knowledge-base/, last accessed date: 03/08/2022.

DOI: 10.5334/johd.84

Interpretations find their natural connecting points in the places or persons they are about. It is our intuition that we should not overstate the credibility of one interpretation against another. There are descriptions of events deserving more credit than others, but, still, as descriptions, they stand on an equal footing. This is because, in the DnS approach, it is the context that matters most. After all, the impact of events may easily shift, depending on their context.

The diagram in Figure 4 represents the scramble for the bishop's see in 1249. This phase in the conflict is represented as a dul:Process (on top) that has two constituents. The first is a Situation grounded on "common-ground" knowledge (the f:EventParticipationSituation on the left), whereas the other, a f:EventInterpretationSituation derived from van der Aa's *Dictionary*, bases itself on the chronicle of Johannes de Beka and is obviously mistaken. Both Situations have one event in common: the undisputed death of bishop Otto van Holland in 1249. The events included in the Situation stemming from Beka must be fictional and are clustered into an Interpretation. As a consequence, the participants in this Beka Situation have to be typecast as FictionalCharacters and not as Persons who have lived a real life. Moreover, different Situations may have a different number of Events. We added one: the alleged "Gozewijn van Amstel" was said to have died in 1250, whereas Gozewijn van Randerath died in 1262.

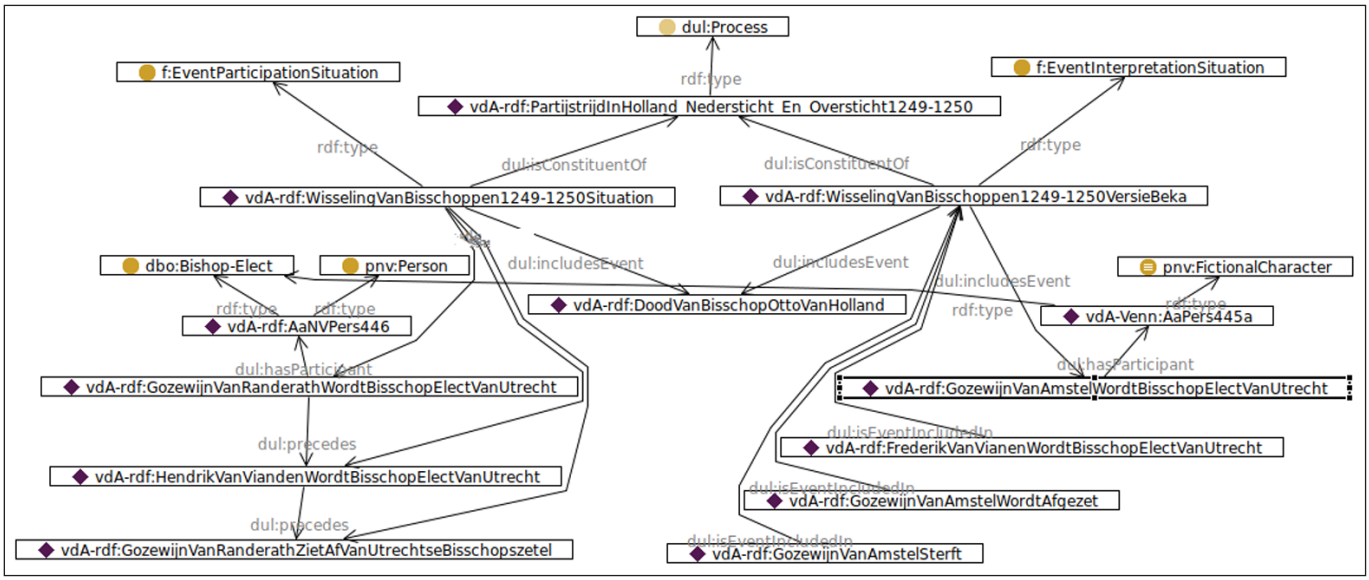


Figure 4 Political strife in the bishopric of Utrecht, 1249–1252, common ground and interpretation.

(8) REFLECTIONS AND CONCLUSIONS

In this work, we have discussed how to transform history-book information into digital representations of events. Representations that are to stand the same tests as text criticism applied to any interpretation of a facet of history. We started from the wish to have better digital data about history than just tables and names. Data that, although still very much traceable, have come loose from the information carriers in collections of cultural heritage items. Also, we wanted to make data express the same complexities which we encounter in everyday life. If we were to succeed, a view of the past would be created which, because of its digital form, is available always and everywhere and does not need to be reconstructed every time a new reading is proposed. With a source like van der Aa's *Dictionary* set as data, cultural heritage institutions would not need to restructure their database each time a narrative has been adapted. A huge amount of historical information would be ready to be reused in order to express the ideas behind a new exposition etc. Furthermore, in an educational context, this kind of complex knowledge representation structure may help us to discuss with pupils the notion of a set of entities considered as accepted knowledge versus their subjective interpretation, allow rich visualizations of the knowledge, and interaction with the knowledge, and finally enable the manipulation and editing of the knowledge to form alternative views, possibly based on the same material. One conclusion that we can be sure of is that in the near future, if we are going to do this on a bigger scale, we need to develop a standard for the identification, typology, and granularity of events to be represented through digital channels.

ACKNOWLEDGEMENTS

The authors wish to thank Gerald Wildenbeest (Dutch-language DBpedia), Ed de Heer (Dutch National Archives), and Lizzy Jongma (NIOD Institute for War, Holocaust and Genocide Studies) for their support and their comments.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR INFORMATION

Gerard Kuys: conceptualization, investigation, data curation, original draft – writing

Ansgar Scherp: validation, original draft – writing

AUTHOR AFFILIATIONS

Gerard Kuys orcid.org/0000-0001-7776-3776
Dutch-language DBpedia, Netherlands

Ansgar Scherp orcid.org/0000-0002-2653-9245
Data Science and Big Data Analytics, Ulm University, Germany

REFERENCES

de Beka, J. (ca. 1346). *Croniken van den Stichte van Utrecht ende van Hollant* (ca. 1346); http://resources.huygens.knaw.nl/pdf/beke/bekeNL02deel1.pdf, last accessed date: 03/08/2022.

D'Souza, D. F., & **Wills, A. C.** (1998). Objects, Components and Frameworks with UML. *The Catalysis Approach*. Reading, Mass.: Addison-Wesley.

Geyl, P. (1956). *Gebruik en misbruik der geschiedenis (de Terry lectures, gehouden aan Yale University, Newhaven, Conn., U.S.A., in October 1954)*. Groningen: Wolters.

Historisch Centrum Overijssel. *Huis Almelo. 1236–1917 (1933). Archief van Huis en heerlijkheid Almelo;* https://www.archivesportaleurope.net/ead-display/-/ead/pl/aicode/NL-ZlHCO/type/fa/id/0214/, last accessed date: 03/08/2022.

Scherp, A., Franz, T., Saathoff, C., & **Staab, S.** (2009). *F--a model of events based on the foundational ontology DOLCE+DnS Ultralight, K-CAP* 2009, 137–144. DOI: https://doi.org/10.1145/1597735.1597760

van der Aa, A. J. (1839–1851). *Aardrijkskundig Woordenboek der Nederlanden, bijeengebragt door A.J. van der Aa,* 13 vols. Gorinchem: Jacobus Noorduijn.

van Hage, W. R., Véronique, M., Segers, R., Hollink, L., & **Schreiber, G.** (2011). Design and use of the Simple Event Model (SEM). *Journal of Web Semantics*, 9(2): 128–136. DOI: https://doi.org/10.1016/j.websem.2011.03.003

TO CITE THIS ARTICLE:
Kuys, G., & Scherp, A. (2022). Representing Persons and Objects in Complex Historical Events using the Event Model F. *Journal of Open Humanities Data*, 8: 22, pp. 1–12. DOI: https://doi.org/10.5334/johd.84

Published: 02 September 2022

Journal of Open Humanities Data is a peer-reviewed open access journal published by Ubiquity Press.