

AN ANALYSIS OF THE CHALLENGE OF DATA COMPLETENESS OF EVENTS FOR CULTURAL HERITAGE INSTITUTIONS

DAAN RAVEN (2591548) - VICTOR DE BOER



Problem Definition. Databases of cultural heritage institutions currently often contain limited object-related event data. Providing this data is necessary for adding historical context to objects. However, the current structure of the data makes it difficult to measure the data completeness.

Research Objective. In a cooperation with the Rijksmuseum Amsterdam, we attempt to enhance the structure and completeness of the historical context which cultural heritage institutions add to their objects within a data model.

Dijkshoorn et al. (2018) van den Akker et al. (2011) de Boer et al. (2017)

1. FINDING A DEFINITION OF THE HISTORICAL EVENT

Challenges.
Lack of historical consensus on events.

Need for both a theoretical and practical foundation.

Works regarding Event & Models



Table 1. Prototypical Data Model

| Information Features | Domain | Range |
|----------------------|--------|---------------|
| hasName | Event | Name |
| hasDescription | Event | Description |
| hasParticipant | Event | Actor |
| hasLocation | Event | Place |
| hasRegion | Event | Place |
| hasEventDate | Event | Date |
| isObservableAt | Event | Time Interval |
| hasType | Event | Type |
| knownAs | Event | Event/Name |
| wasInfluencedBy | Event | Event |
| fallsWithin | Event | Event |
| consistsOf | Event | Event |

An event is (1) a unique interpretation of a phenomenon which has (2) occurred at a place or region in (3) a limited extent in time. An event is (4) described in causes which result in change in the course of history. This interpretation needs (5) to be acknowledged by a contemporary.

2. ANALYSING DATA COMPLETENESS W.R.T THIS DEFINITION

Challenges.
Mapping to current RMA database.

Extracting values from Adlib/XML.

Events stored in Database Rijksmuseum Amsterdam

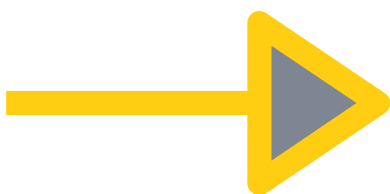
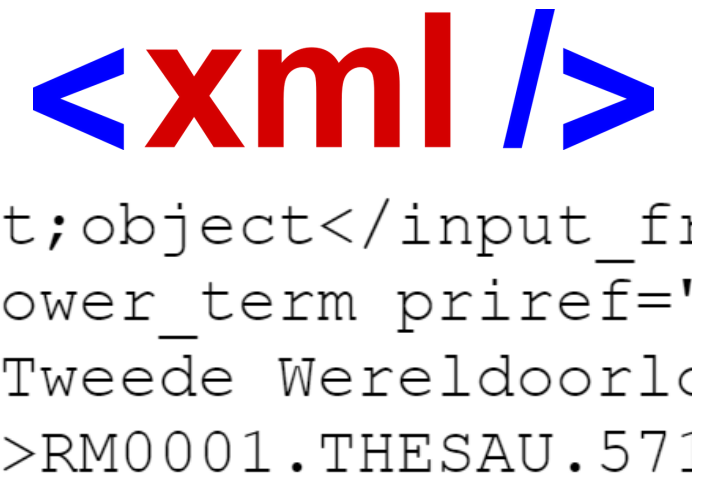


Table 2. Current Data Completeness (32 Events)

| Relation | Number of Occurences |
|----------------|----------------------|
| hasDescription | 18 |
| hasName | 32 |
| fallsWithin | 29 |
| consistsOf | 31 |
| hasParticipant | 0 |
| hasLocation | 0 |
| hasRegion | 0 |
| hasEventDate | 0 |
| isObservableAt | 0 |
| hasType | 0 |
| knownAs | 0 |
| influencedBy | 0 |

3. ENHANCING COMPLETENESS USING NLP TECHNOLOGY

Challenges.
Extracting labels from event descriptions using NLP Technology.

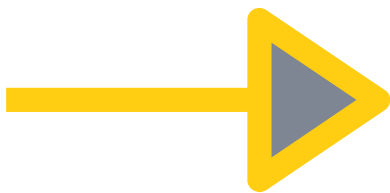


Table 3. Comparison of labelled entities by FROG NLP and TextRazor (32 Events)

| Label | FROG NLP | Correct Label | TextRazor | Correct Label |
|---------------|----------|---------------|-----------|---------------|
| Location | 60 | 17 | 45 | 19 |
| Time Interval | | | 22 | 12 |
| Participant | 22 | 15 | 20 | 17 |
| Event | 11 | 8 | 15 | 15 |

FUTURE STEPS ACKNOWLEDGEMENTS

- 1. Visualizing the model.
- 2. Increasing the number of analysed events.
- 3. Interpreting the results

Chris Dijkshoorn & Trineke Kamerling - Rijksmuseum Amsterdam
Chiel van den Akker - Vrije Universiteit Amsterdam

POSTER & REFERENCES

