Representation of Archival User Needs Using CIDOC CRM

CRMEX
TPDL 2013 (Malta)
26.09.2013

Steffen Hennicke
Berlin School of Library and Information Science
Humboldt-Universität zu Berlin
Archival “Information Potentials”


- The main means of access are **archival aids** which...
  - ... help to understand the archive and its holdings.
  - ... provide contextual information about the origins, provenance and original order of holdings.

- Research indicates that...
  - ... archival aids do not fully meet information needs of users.
  - ... we even do not know a lot about user needs.
  - ... archival research is too often determined by an archival point of view.
  - ... archival information systems do not adequately exploit the information potentials of archives.
Research Context

• **Research Interest**
  - Contribute to the design of better access and **discovery systems** for archives and other historical knowledge bases in order to improve **discovery of facts and materials** in archives.

• **Hypothesis**
  - User enquiries to archives exhibit **common patterns** in terms of the underlying research interest.
  - These common patterns can be formally represented in an **ontological model**.

• **Result**
  - An **ontological model** of common patterns which reasonably represents user needs towards archives and enables adequate answers to typical user questions.
  - working title: “Archival Knowledge Model” (AKM)
Main Research Question

• “Is there a hypothetical ontological model which can represent user enquiries and their probable interpretations as formal queries against a model of the archival target world that would adequately answer the enquiry or its implicit purpose?”
Research Data

• Archival **reference questions** from user files
  - contain an largely **unfiltered** information need
  - from the Federal Archives of Germany (Bundesarchiv)
  - focus on Contemporary German History
• 60 user files with 546 single questions
  - 260 questions explicit or implicit “resource discovery”
  - 70 questions “factual”
  - 216 questions “other”
• In total: 330 questions (“resource discovery” and “factual”) are currently being scrutinized for common patterns.
Methodological Approach

- Inspiration: Wendy Duff and Catherine Johnson (2001) [DJ01]
  - Analysis of E-mail Reference Questions to North-American Archives
  - Type of Questions and type of Wanted and Given Information
- 2-Step-Interpretation of Questions
  - Wanted Information: Explicit user perspective
  - Context of Information: Relations between information entities (expressiveness)
  - Find common patterns in inquiries and generalise and formalise them into an ontological model
- Ontological Model
  - CIDOC Conceptual Reference Model (CRM) [CRM]
  - Empirical based model for the cultural heritage domain
  - Conceptualizes information and history around the notion of events
Example 1 - Enquiry

- **Contextual Information**
  - “A source I would like to see are the police- and surveillance reports from the Weimar Republic which are about revolutionary movements. I would like to know what the surveillance agency of the Reich (or the ones of the Länder) had to say about [person name].”

- **Question 1**
  - “Do you know if the Bundesarchiv holds such documents?”

- **Question 2**
  - “Which agency of the Reich was responsible for the surveillance of revolutionary movements? The Reich or the Länder?”
Example 1 - Analysis

• **Contextual Information**
  - “A source I would like to see are the police- and surveillance reports from the Weimar Republic which are about revolutionary movements. I would like to know what the surveillance agency of the Reich (or the ones of the Länder) had to say about [person name].”

• **“Given” Contextual Information**
  - type of documents (“police- and surveillance reports”)
  - name of a period (“Weimar Republic”)
  - type of a group (“revolutionary movements”)
  - type of a legal body (“surveillance agency” [of the Reich/Länder])
  - name of a specific actor (“[person name]”)
Example 1 - 1. Interpretation Step

• Guiding Questions
  – What is the research interest (wanted information) of the question?
  – Which are probable and adequate answers to the question?

• Question 1: “Do you know if the Bundesarchiv holds such documents?”
  – Material-finding asking for pointers to documents...
    • ... about [“person name”]
    • ... about “revolutionary movements”

• Question 2: “Which agency of the Reich was responsible for the surveillance of revolutionary movements? The Reich or the Länder?”
  – Fact-finding question asking for the name of a legal body...
    • ... responsible for the surveillance
Example 1 - 2. Interpretation Step

- **Guiding Question**
  - How to translate the interpretation of the question to CIDOC CRM?
    - CIDOC CRM: Historical entities are connected through events.

- **Question 1**
  - The documents in question are the result of a “surveillance activity” targeted at a specific type of group or a specific actor.

- **Question 2**
  - The legal body whose name is in question was involved in an “surveillance activity” which was targeted at a specific type of group.

- The common denominator is a “surveillance activity”.
- Such activities result in documents.
- Such activities are based on “mandates”.
- The “documentation activity” is the most general abstraction of a “surveillance activity.”
Example 1 - Question 1

Do you know if the Bundesarchiv holds such documents?

E39 Actor
"Surveillance Agency of the Reich"

E4 Period
"Weimarer Republik"

E31 Document
?

E29.1 Mandate

type: to police
type: to keep under surveillance

has mandate

E39 Actor
"[person name]"

has target group

"revolutionary movements"

P14 carried out by

E7.1 Documentation Activity

type: policing / surveillance

documents

P94 has created

E5 Event

P11 had participant

P129 is about
Example 1 - Question 2

Which agency of the Reich was responsible for the surveillance of the revolutionary movements? The Reich or the Länder?
Why extending?

- **E65 Creation**
  - “events that result in the creation of conceptual items or immaterial products”

- **E29 Design and Procedure**
  - “documented plans for the execution of actions in order to achieve a result of a specific quality, form or contents”

- **P17 was motivated by**
  - “describes an item or items that are regarded as a reason for carrying out the E7 Activity”

- **P20 had specific purpose**
  - “identifies the relationship between a preparatory activity and the event it is intended to be preparation for”

- **P21 had general purpose**
  - “describes an intentional relationship between an E7 Activity and some general goal or purpose”
Example 2

• Contextual Information
  – In 1980, a delegation of the FDGB lead by Harry Tisch laid down a wreath of flowers in Oradour. The visit was part of a trip of the FDGB to France (demonstration in Limoges, reception and meeting with the FKP and CGT in Paris). At this time, Tisch was also a member of the Politbüro of the ZK of the SED.

• Question 1
  – Where can documents be found about the planning [of this trip]...

• Question 2
  – ...and the report on this trip?

• Question 3
  – In your opinion, has such a trip been discussed or, at least, been approved in the ZK?
Example 2 - Question 1 and 2

E29.1 Mandate
- type: to document / to report

E39 Actor
- has mandate

E7.1 Documentation Activity
- type: documenting / reporting

E31 Document
- P14 carried out by

E5 Event
- "Trip to France"
- documents
- refers to

E5 Event
- "planning activities"
- documents

P94 has created

P129 is about

Where can documents be found about the planning [of this trip] (...)?

Where can (...) the report on this trip be found?
Example 2 - Question 3

In your opinion, has such a trip been discussed or, at least, been approved in the ZK?
Documentation-Activity Pattern

E29 Design or Procedure

E7 Activity

E65 Creation

EAD

E29.1 Mandate

E39 Actor

E4 Period

E52 Time-Span

E31 Document

ARCHIVE

Intentional Level

Archival Reality

has mandate

P14 carried out by

P10 falls within in

P4 has time span

P94 has created

P129 is about

following mandate

target function

Activity Type

target actor

Actor Type

has type

has type

documents

has type

E5 Event

P11 had participant
“Archival Knowledge Model” (AKM)

- The AKM comprises several common patterns found in user enquiries to archives.
  - Documentation-Activity pattern
  - Document-Relation pattern
  - Actor-Context pattern
  - etc.

- Constitutes a formal model of user needs towards archives in covering and extending the CIDOC CRM.
Value and Potential

• “Archival Knowledge Model” (AKM)
  - as a formal model (CIDOC CRM)
    • constitutes a conceptual and empirical reference of (archival) user needs
    • supports the design and extension of (archival) metadata schemas
    • derive new archival cataloguing rules
  - as an implementation (e.g. in RDFS/OWL)
    • an unobtrusive access layer to archival holdings
    • an extensible context layer for archival holdings
  - as methodology to formalise (archival) user needs
    • 2-Step-Interpretation of user questions
References

