Bio CRM – elämätarinoiden kuvaaminen linkitetynä datana

Jouni Tuominen
HELDIG – Helsinki Centre for Digital Humanities, University of Helsinki, http://heldig.fi
jouni.tuominen@helsinki.fi
Biographical Conceptual Reference Model

“Create a systematic way to represent biographical information for prosopography, including basic data of people, personal relations, professions, and events with participants in different roles.”
Bio CRM: principles

- Harmonizing model for different datasets
- Linked Data approach to combine with other cultural heritage datasets
- Separation of the general data model and dataset-specific extensions
- Support for principal query types for prosopographical research
- Intuitive information representation and query writing
- Support for reasoning and validation
Prosopographical use cases

- **Information retrieval.** Find a set of people that share selected characteristics. For example: men who were born in England 1800–1850 and had certain occupations.
- **Network analysis.** Find a network of people based on some criterion. This network can then be analyzed further by network analysis tools or visualizations.
- **Knowledge discovery.** Identify (automatically) a hidden group of people in the data that share some characteristics of interest that are not known before hand.
- **Dynamic analysis.** Analyze the structure and changing composition of a group and the changing roles of individuals or subgroups.
CIDOC CRM

- Event-based model for cultural heritage data
- Natural basis for representing biographical data
- Person’s life can be seen as a sequence of spatiotemporal, possibly interlinked events from birth to death
CIDOC CRM events (or temporal entities)
What is an event?

- Are professions events, such as being a bishop or a president, because holding an office occurs in time and space with an agent involved?
- Are family relationships events, e.g., getting and being married?
- Modeling everything as an event gets quite complicated
Role-based modeling as a solution

- Attributes are properties of a person that are assumed to characterize her independently of time and space, e.g., profession
- Relationships are established between people and are assumed to characterize the people independently of time and space, e.g., father-of
- Events take place in time and space and involve participants in different roles
- All these are expressed using roles
Existing options for representing roles in events

- W3C PROV (used in EMLO Linked Data pilot)
- CIDOC CRM (CRMpc proposal)
- VIVO/BFO
- Simple Event Model (SEM)
- Custom role specific-properties
- schema.org

Further discussion:
https://docs.google.com/document/d/1XplgUBjCz16nYwhmLwxKV5av6dn3QHz5o8pMN6jTUbq
Roles as instances (VIVO/BFO-inspired)

- **Unary roles (attributes):** Female, British, Bishop, ...
  - John F. Kennedy has the role President
- **Binary relationship roles:** Father, Teacher, Member, ...
  - John F. Kennedy has Spouse Jacqueline Kennedy Onassis
- **Roles in events:** Baptismal candidate, Participant, ...
  - John F. Kennedy has the role Assassination target in the Assassination event in Dallas, Texas on Friday November 22, 1963
Core classes of Bio CRM

The role modeling