





OpenAtlas An Open Source Application to Map Historical data with CIDOC CRM

Alexander Watzinger

June 18th, 2021

CC-BY 4.0









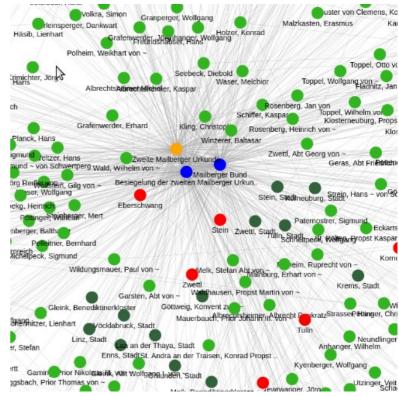




OpenAtlas

https://openatlas.eu

- Open source, browser based database software
- Acquire, edit and manage research data
- Historical, archeological and prosoprographic projects
- Developed by a small core team

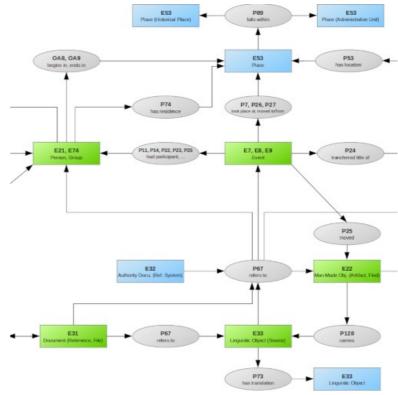


https://demo.openatlas.eu/overview/network/

Model

CIDOC Conceptual Reference Model

- International standard (ISO)
- Developed by CIDOC CRM Special Interest Group
- Entities with classes like actor, source, event, place connected via links
- Stored in an object oriented network



https://demo.openatlas.eu/overview/model

Challenges

- Different or uncertain spelling
- Name parts with meaning
- Ambiguous names
- Possible duplicates
- Changes over time

Falsehoods Programmers Believe About Names

- 1. People have exactly one canonical full name.
- 2. People have exactly one full name which they go by.
- 3. People have, at this point in time, exactly one canonical full name.
- 4. People have, at this point in time, one full name which they go by.
- 5. People have exactly N names, for any value of N.
- 6. People's names fit within a certain defined amount of space.
- 7. People's names do not change.
- 8. People's names change, but only at a certain enumerated set of events.
- 9. People's names are written in ASCII.
- People's names are written in any single character set.
- 11. People's names are all mapped in Unicode code points.
- People's names are case sensitive.
- People's names are case insensitive.
- 14. People's names sometimes have prefixes or suffixes, but you can safely ignore those.
- 15. People's names do not contain numbers.

Patrick McKenzie: Falsehoods Programmers Believe About Names https://www.kalzumeus.com/2010/06/17/falsehoods-programmers-believe-about-names/

Names in Forms

- Classical approach
 - Title
 - First Name
 - Last Name
- This approach gets quickly problematic if:
 - A former name exists
 - A middle name exists
 - It's uncertaint what was the first and what was the last name
- Would be adding more fields be a viable solution?

Names in OpenAtlas

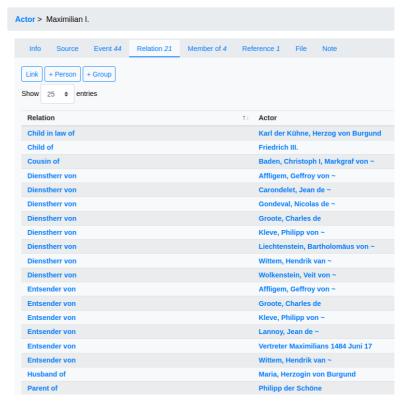
- Only one free text field
- Unlimited aliases
- Supports a wide range of Unicode characters
- Title/function/profession via relations
- Tools to manage names



https://demo-dev.openatlas.eu/entity/10743

Title, Function, Profession

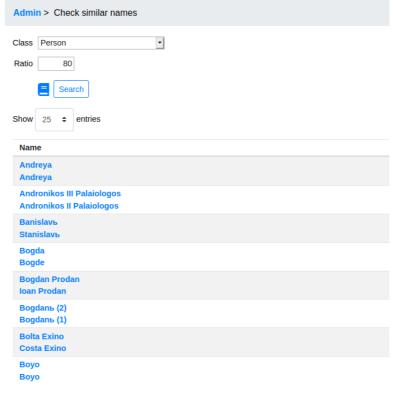
- Always in a relation to a group or institution
- Multiple relations possible
 - Can differ in time
 - Can differ in context
 - Can differ in function
- Allows more detailed data and analysis



https://demo.openatlas.eu/entity/144#tab-relation

Tools for names in OpenAtlas

- Links to external reference systems
- Display aliases in table overviews
- Also search in aliases
- Find name duplicates
- Find orphaned data



https://demo-dev.openatlas.eu/admin/similar

Open Challenges

- There isn't one right solution, no silver bullet
- (Many) anonymous persons
- Track uncertainty if one name identifies one or more persons

Thank you for listening



https://openatlas.eu