

Co-Reference, CRM, and people or: Next Steps: Assertions

Øyvind Eide

Lehrstuhl für Digital Humanities
Universität Passau

<http://www.oeide.no/>

What is co-reference?

- “The table by the window was...”
- “The beautiful table could be...”



What is co-reference

- The fact that two or more signs in a context and a reading refer to the same entity
- Exist in most if not all documents
- Co-references exist independently of our information systems
- Based on an identity definition for the types of entities in question
- Based on context, meaning, not string matching

Co-reference statements

- Implicit co-references exist in the world and can be detected
- This is not enough, we need explicit co-reference statements. They should include:
 - Two entities which can refer
 - The date of the statement
 - The responsible entity (person, software, etc.)
 - A third entity which is the target they refer to

Co-reference in CRM

- Co-references exist independently of our information systems
- Co-reference has always been in CRM
 - Example: E48 Paris P87 identifies E53 Place A
 - E48 Lutetia P87 identifies E53 Place A

If Place A in the two examples is the same, then E48 Paris co-refers with E48 Lutetia in the context of a specific reading of the CRM model.
- Explicit co-reference assignement new in 5.1.2

E91 Co-Reference Assignment (snap:merges)

The event of asserting co-reference:

“This class comprises actions of making the assertion whether two or more particular instances of E89 Propositional Object refer to the same instance of E1 CRM Entity. The assertion is based on the assumption that this was an implicit fact being made explicit by this assignment. Use of this class allows for the full description of the context of this assignment.”

E91 properties

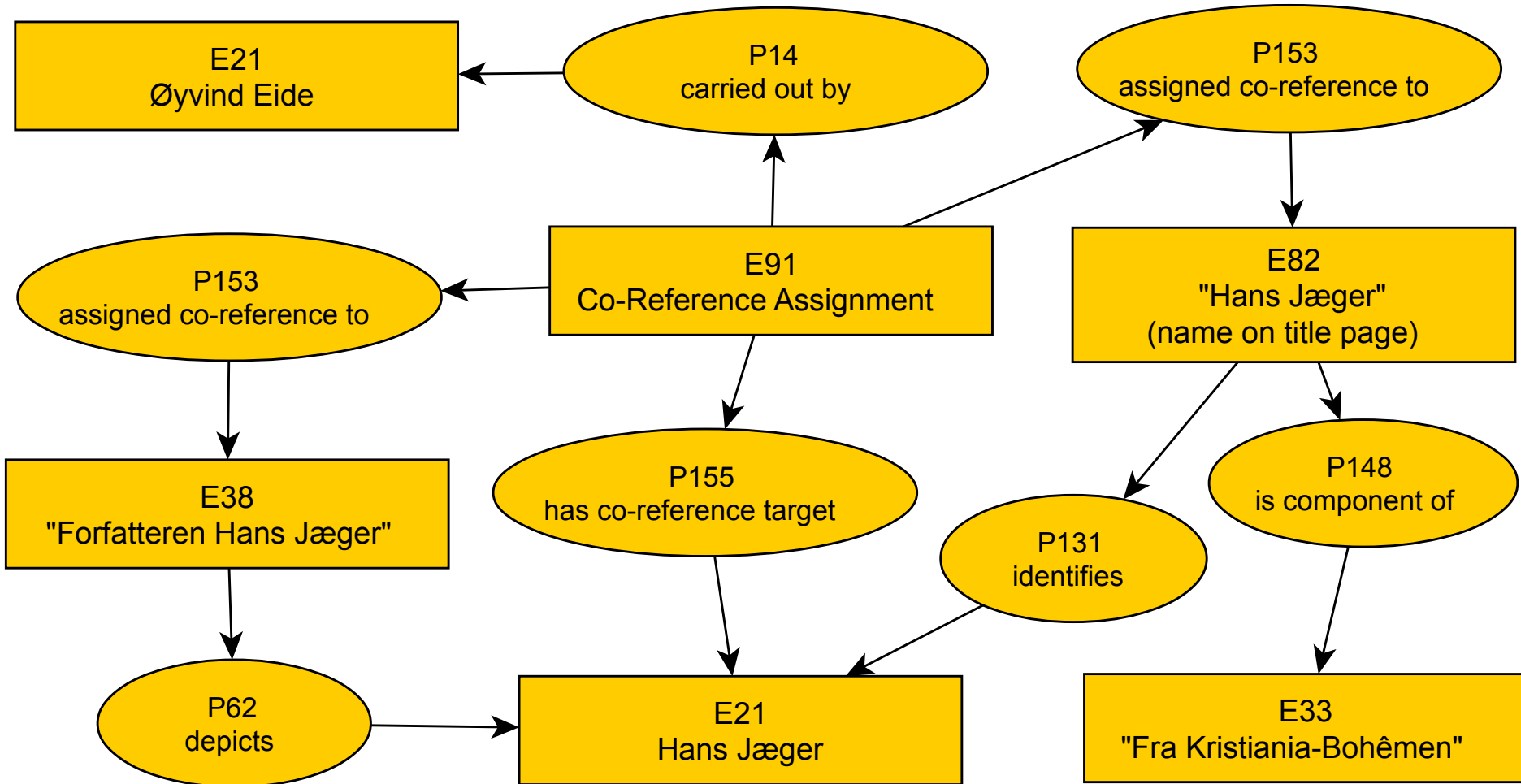
- P153 assigned co-reference to (was regarded to co-refer by): E89 Propositional Object
- P154 assigned non co-reference to (was regarded not to co-refer by): E89 Propositional Object
- P155 has co-reference target (is co-reference target of): E1 CRM Entity (*details still under discussion for this property*)
- Inherited properties: carried out by, have time span, etc.

E91 example 1

Positive co-reference:

‘The assertion that the author name “Hans Jæger” on the title page of the novel “Fra Christiania-Bohêmen” refers to the same historical person as the motive of the painting “Forfatteren Hans Jæger” by Edvard Munch.’

Example 1



E91 example 2

Negative co-reference:

‘The assertion that the author name “Hans Jæger” on the title page of the novel “Fra Christiania-Bohêmen” does not refer to the same historical person as the author of the collection of drawings “Til Julebordet : ti Pennetegninger / af H.J.” incorrectly attributed to Hans Jæger in the Bibsys database.’

Intra-textual co-reference

- Register to printed 18th century historical sources
- Identify persons in text based on page number and name (various spelling)
- Co-reference links the group of references to a single historical person
 - reference in text at same level as registry record
- Also: anaphor resolution

Extra-textual co-reference

- Link to other sources
 - Wikipedia
 - Church books
 - Swedish border protocols
 - Historical Person Authority List (not yet existing)
- Distributed
- Authority list just another reference

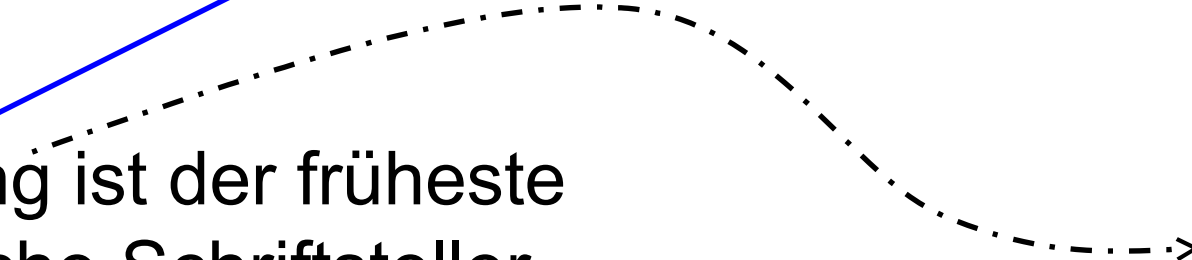
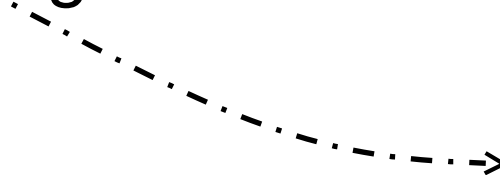
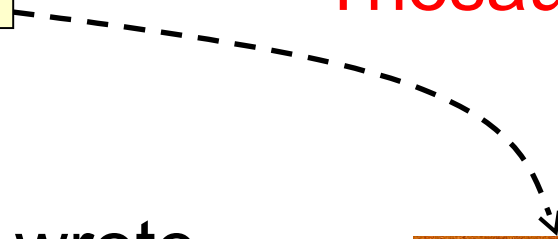
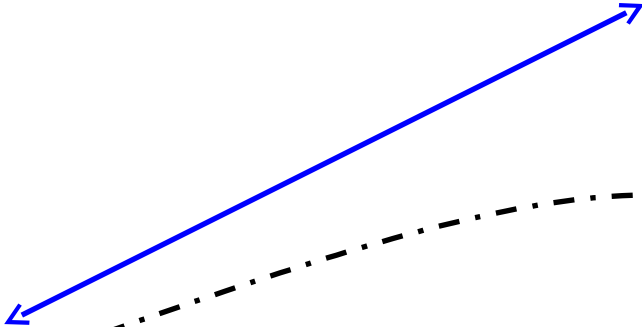
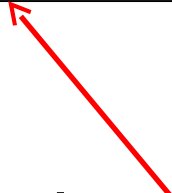
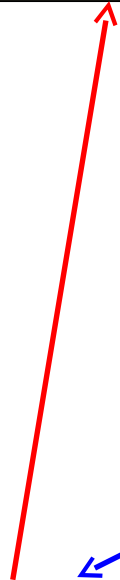
Thesauri and co-reference

Thesaurus record: Gotthold
Ephraim Lessing

Co-reference
Thesaurus link

When Lessing wrote...

Lessing ist der früheste
deutsche Schriftsteller...



Use of co-reference assignment

- E91 Co-Reference Assignment is part of CIDOC-CRM
- Can also be included as a module in other formalisms
- Map snap:merges to E91
- Implementations can be influenced by it

The co-references exist

We can make them explicit or not