Gaining Insight into Cultural Heritage Using Event Descriptions

Tuukka Ruotsalo, Thomas Häggström, Eero Hyvönen
Helsinki University of Technology, Media Laboratory
University of Helsinki, Department of Computer Science
Semantic Computing Research Group
http://www.seco.tkk.fi/

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Agenda

- CultureFinland portal as a demonstration system
  - Thomas Häggström
- Content description and search methods
  - Tuukka Ruotsalo

CultureFinland

- Virtual exhibition with the Semantic Web as a knowledgeable guide
  - An integrated publishing portal
  - All content is searchable and browsable with the same UIs
  - Content interweaved through semantic connections.

UNIVERSITY OF HELSINKI

SeCo
semantic computing

CultureFinland

- Material from dozens of museums and organizations
  - Items, works of art, photographs, video, sound, skills and processes, text documents etc.
  - Annotated with shared vocabularies
    - » YSO, MAO, ICONCLASS
- Diverse content described in different ways and on different levels of granularity
  - Still a lot of interesting links
    - » Epic poem Kalevala has inspired Gallen-Kallela's paintings
    - » Gallen-Kallela painted at Paris in 1880s

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Research problem

- Problem
  - Content not interoperable because of heterogeneous metadata schemes
  - How to integrate the content for users?
- Solutions
  - Ontology based description methods
  - Search and recommendation methods

Target users and needs

- **Professional users**
  - Museum personnel, curators, researchers
  - known item seeking, exhaustive research

- **Utilizing users**
  - School children, teachers, students
  - Historical and cultural context. The story behind the painting! Exploratory seeking, exhaustive research

- **Entertainment user**
  - People who are usually interested in museums and culture
  - Entertaining and educational. Exploratory seeking

UNIVERSITY OF HELSINKI

SeCo
semantic computing

CultureFinland

- **Faceted search**
  - with timeline and map visualisations of search results
- **Virtual exhibitions**
  - built with rule language
- **Semantic recommendations**
  - browsing the content through semantic connections
- **Theme pages**
  - edited content
- **Roundup search**
  - A roundup of one items relation to other portal content
- **Multimedia experience**
  - Video, images, sound and text

CultureFinland proto

Integration problems

1. Metadata scheme has to be general and precise at the same time

    1. Preserve precision

    2. Enable interoperability

2. One information object's data can simultaneously be another object's metadata

3. Domain specific programming and mapping between domains

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Precision and interoperability

- **Precise metadata schemes are domain dependent**
  - Painting metadata scheme might allow metadata:
    - » creatingPlace: Paris

- **General metadata schemes are not precise**
  - Biography sentence "Gallen-Kallela painted in Paris at 1880s", could be annotated:
    - » dc:subject: Paint
    - » dc:subject: Gallen-Kallela
    - » dc:subject: Paris
    - » dc:subject: 1880s
  - But we want to know how annotations are related to each other. Who did what, when and where?

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Data vs. metadata

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Creator: Gallen-Kallela

- …During his studies in Paris Gallen-Kallela was inspired by Kalevala and painted the Aino-triptych…

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Domain specific programming

- To achieve precision there is a need of enumerating attributes
  - manufacturingPlace, creatingPlace, livingPlace
- To achieve generality, an interpreting program has to know how these are related
  - Binary relations capture too much semantics
    - » CreatingPlace(x,y) of Painting(x) is actually an n-ary relation: Painting(x), Paint(x,z), Place(z,y)
    - » Why don't include this information to the ontology?

Activity hierarchy

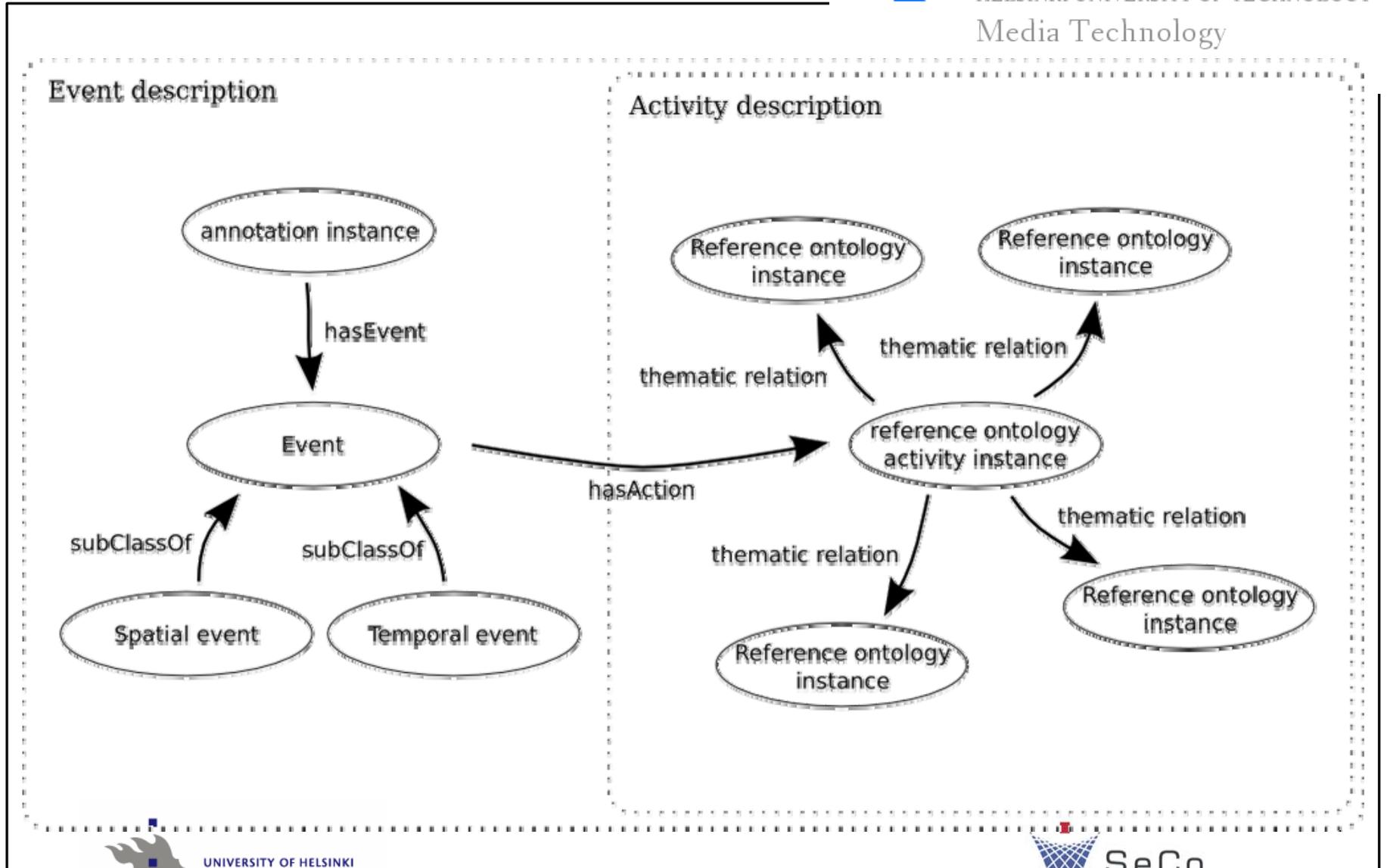- Reference ontologies already have hierarchies for activity
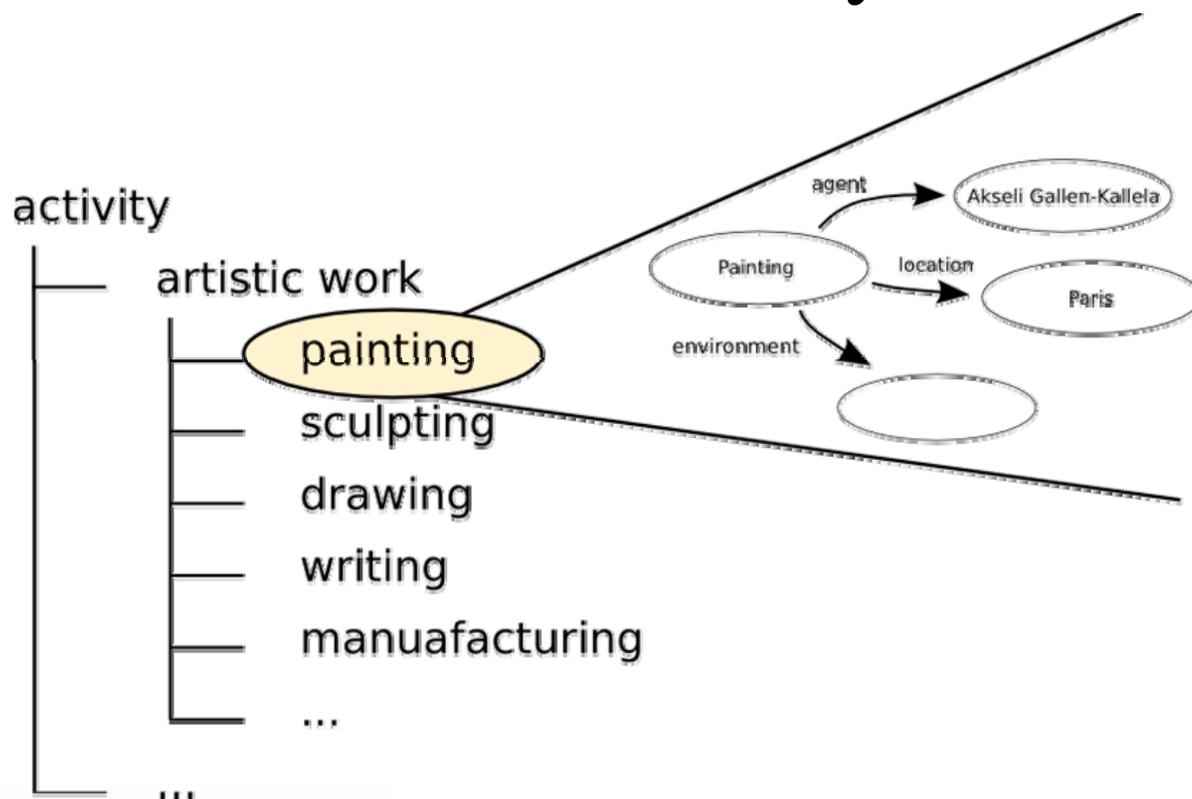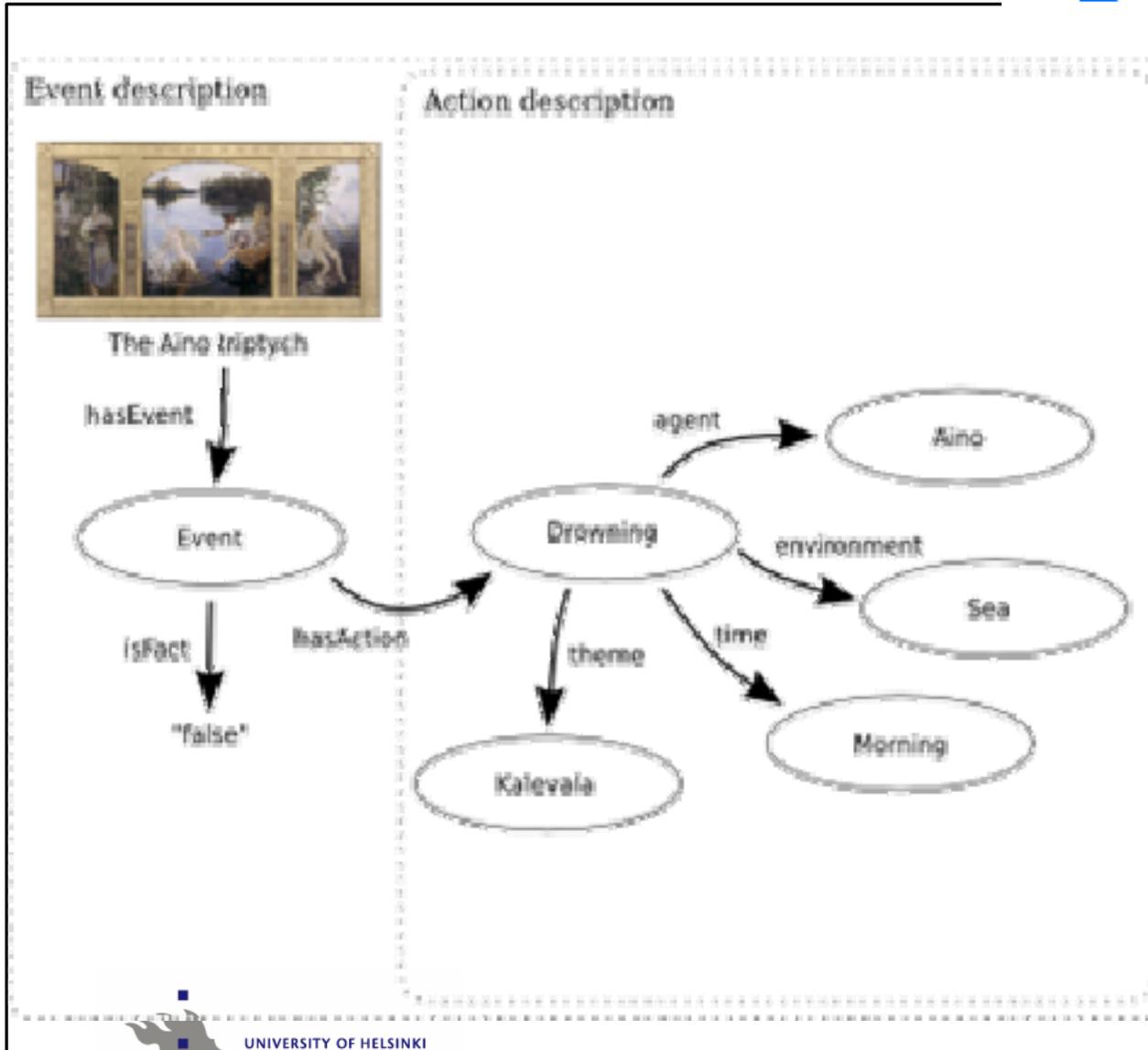
Event descriptions and thematic relations

- Event descriptions represent the activity and the related objects
  - Related objects are connected to the activity with thematic relations
  - Thematic relations specify the type of role the object has in the activity
- State what the content actually is (not just keywords)
- Integrate the metadata and content description

Content

Event description

Action description

The Aino triptych

hasEvent

Event

isFact

"false"

hasAction

Drowning

agent

Aino

environment

Sea

theme

time

Kalevala

Morning

A section of epic poem Kalevala:

…The maiden Aino Drowned herself into the sea in the morning…

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Painting annotation

| Type	Painting
Title	Aino-triptych
Creator	Gallén-Kallela
Date	1889
Location	Paris
Subject	Kalevala

Biography in text format

... Gallén started depicting Kalevala, which had fascinated him for many years. First he painted the three parted Aino-painting, which he finished in 1889. It won the second place medal in the Paris worldfair the same year ...

Event description

The Aino triptych

Action description

Biography

hasEvent

patient

agent — Akseli Gallen-Kallela

hasEvent

Event

Paint

location — Paris

hasAction

theme — Kalevala

time — Time

startTime — "1889"

endTime — "1889"
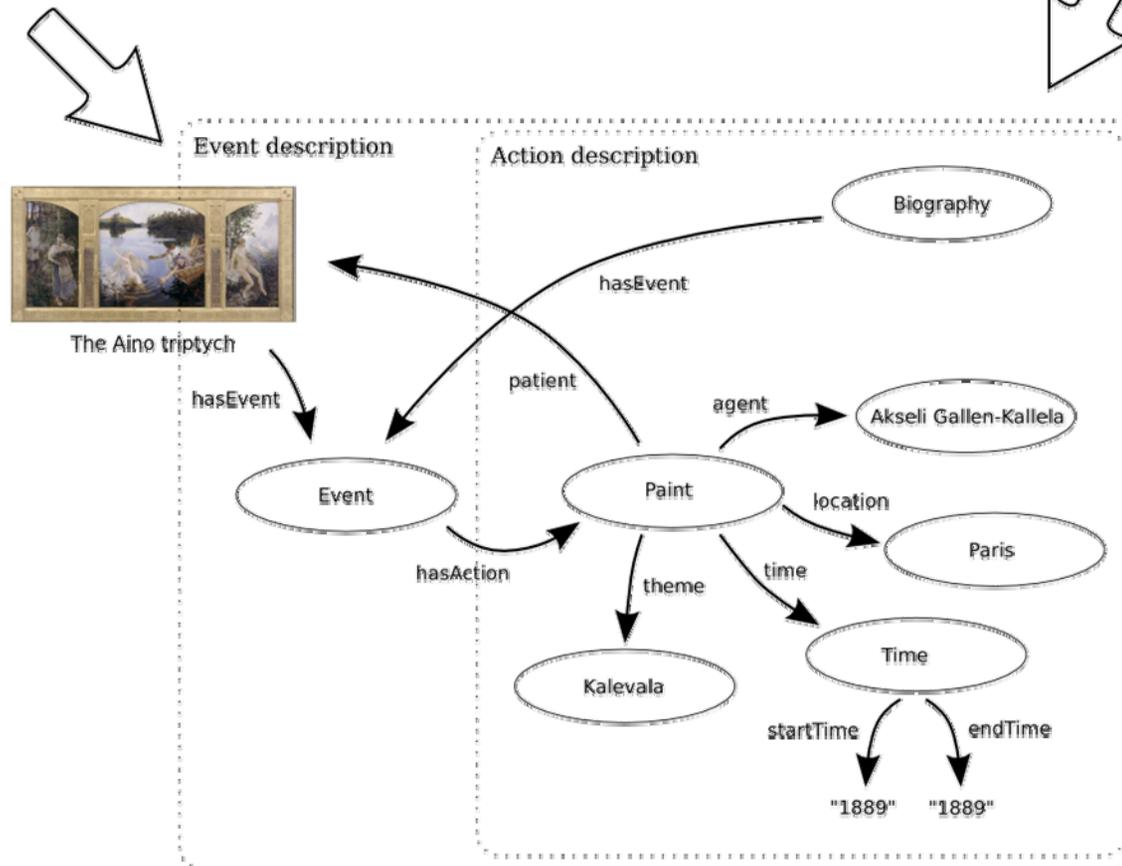
Processes and narratives

- Event descriptions are connected to other to formulate a narrative
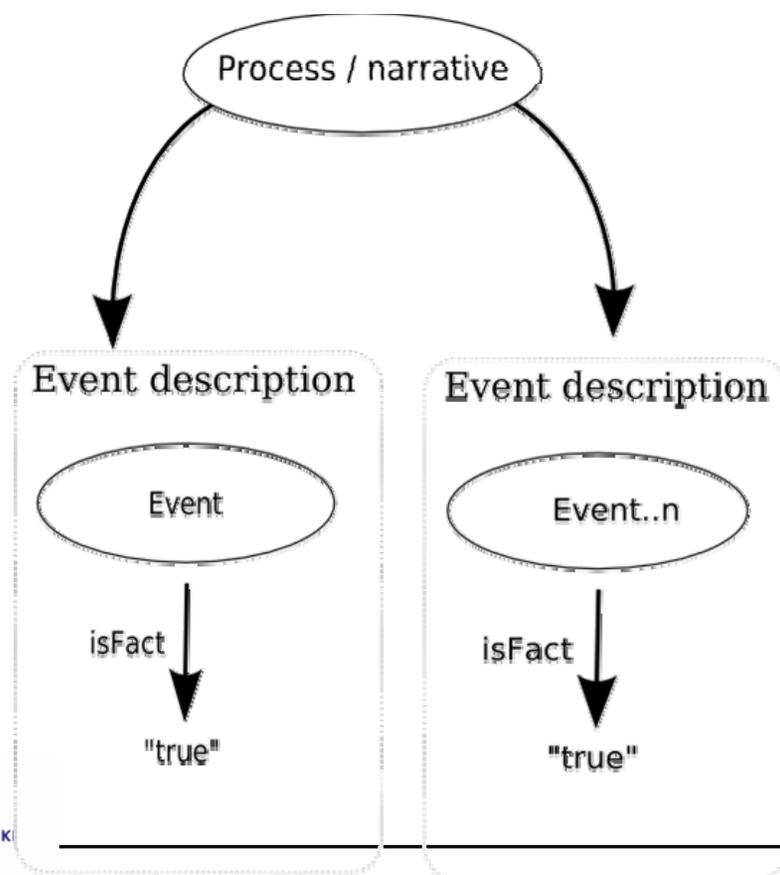
Methods

- Content and metadata can be annotated with a single event description scheme

- The methods can be applied to event description scheme instead of creating domain specific applications
  - Thematic relations are a closed set of relations that the application / agent has to analyse
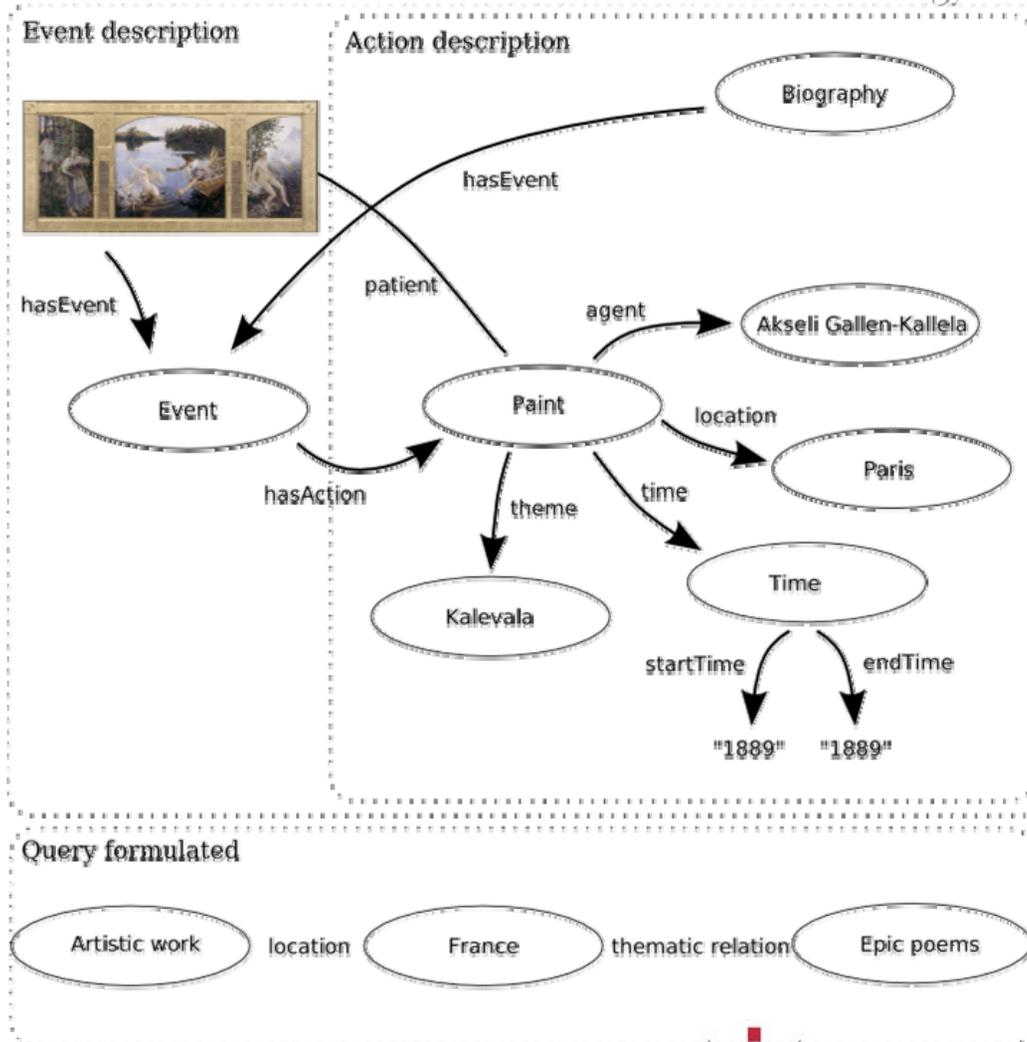
Methods

- Match the query to event frames instead of blind graph search
- Search
  - Exact match
- Recommendation
  - Match with a amount of variance
- Virtual exhibitions
  - Queries following each other

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Search example

- Tell me about Artistic work in France that is somehow related to Epic Poems

Summary

- Heterogeneous data from heterogeneous domains need to be integrated to provide insight into the content
- Need for precise but general description method enabling cross-domain interoperability
- A clear separation between data and metadata can not and should not be made
- Traditional annotation schemes use binary relations and force to create domain specific programs to preserve precision
- Event descriptions with activity as the central concept
    - Activity is the glue between cultural objects
    - Processes and narratives are continuums of activities
- Domain independent event description based search and recommendation methods can be developed

UNIVERSITY OF HELSINKI

SeCo
semantic computing