Aalto University

Ontology-based Information Access: Experiences from FinnOnto projects

UNC, Chapel Hill, April 1st 2011

Tuukka Ruotsalo
Aalto University, School of Science and Technology

Semantic Computing Research Group (SeCo)
http://www.seco.tkk.fi/

Outline

- Intro: Projects and Motivation
- Ontologies, Annotation, Retrieval
- Demonstrators: CultureSampo, SmartMuseum

Background Projects 2003-

Aalto University

- National Ontology Project in Finland (FinnONTO)
  - 2003-2007, 2008-2009, 2010-2012
  - Funded by Tekes and 38 companies and organizations
- Finnish Cultural Foundation 2008-2010
- SmartMuseum EU FP7 2008-2010
- Semantic ubicom services 2010-2012
- Joint work of some 20 researchers

FINNONTO

UNIVERSITY OF HELSINKI

SeCo semantic computing

aac global
SanomaWSOY Group

elisa

connexor
natural knowledge

ESPOON kaupunginmuseo
ESBO stadsmuseum

GTK

mmm.fi
MAA- JA METSÄTALOUSMINISTERIÖ

16 40 KANSALLISKIRJASTO

KTL
Kansanterveyslaitos
Folkhälsoinstitutet
National Public Health Institute

UNIVERSITY OF HELSINKI University
Viikin tiedekirjasto

MAANMITTAUSLAITOS

m-cult

MUSEOVIRASTO

OPETUSHALLITUS

tsk

sininen meteoriitti

STAKES

SKS
SUOMALAISEN KIRJALLISUUDEN SEURA

SUOMEN MAATALOUSMUSEO SARKA

Suomen valokuvataiteen museo
Finlands fotografiska museum
The Finnish Museum of Photography

Tekes

TietoEnator

TEKNILLINEN KORKEAKOULU
Viestintätekniikka

UNIVERSITY OF HELSINKI

Valtiovarainministeriö
Finansministeriet Ministry of Finance

YLE

TEKNILLINEN KORKEAKOULU
Kartografia ja geoinformatiikka

CEL AMANZI

grip STUDIOS

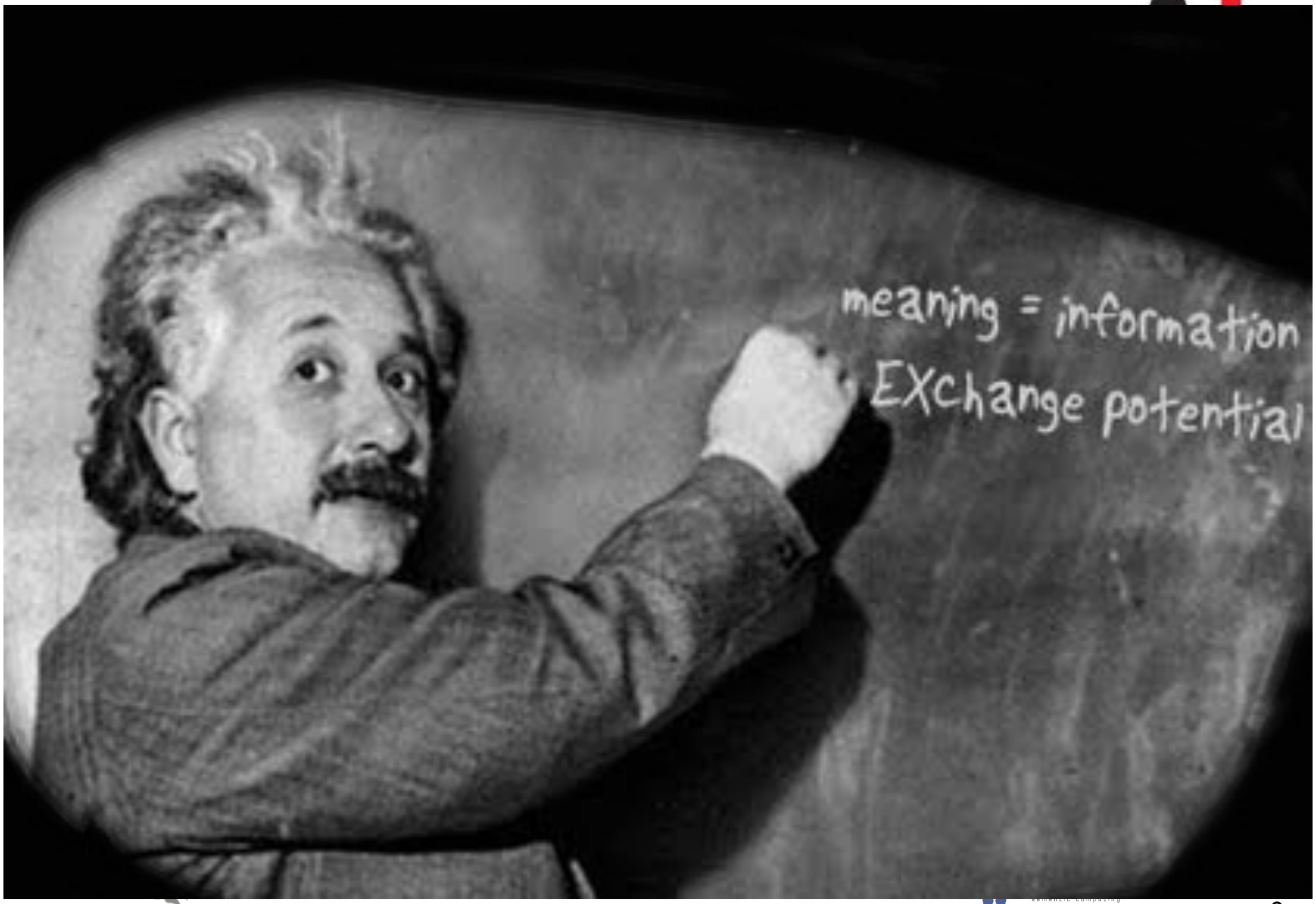
VALTION TAIDEMUSEO

HELSINGIN KAUPUNGINKIRJASTO
HELSINGFORS STADSBIBLIOTEK
HELSINKI CITY LIBRARY

UNIVERSITY OF HELSIN

SanomaWSOY

Antikvaria-ryhmä


meaning = information
EXchange potential

SEMANTIC INTEROPERABILITY

=

ONTOLOGIES
ANNOTATIONS
RETRIEVAL METHODS & APPLICATIONS

Aalto University

UNIVERSITY OF HELSINKI

SeCo semantic computing

Aalto University

ONTOLOGIES, BUT WHERE DO THEY COME FROM?

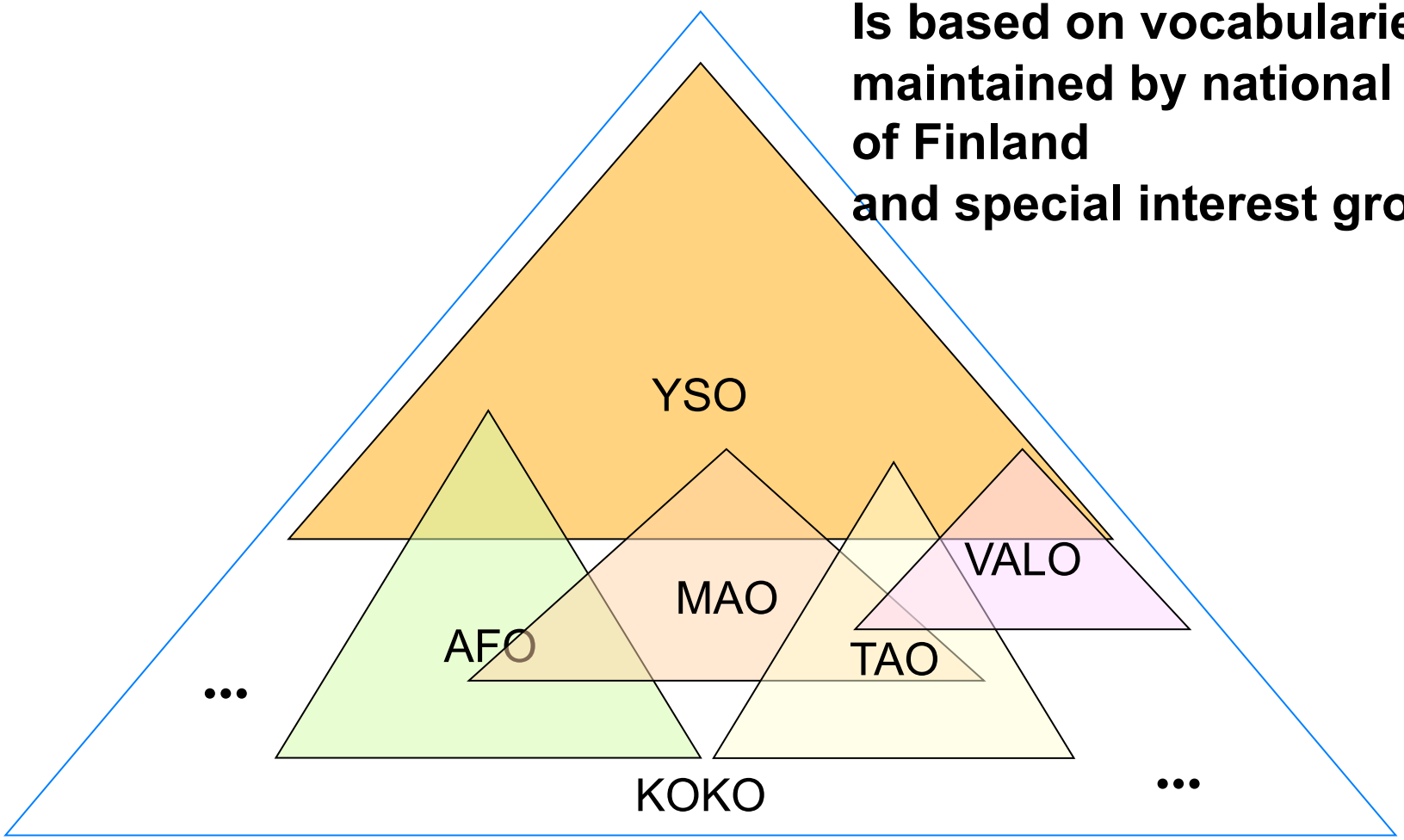
UNIVERSITY OF HELSINKI

SeCo semantic computing

KOKO: Finnish Collaborative Holistic Ontology
handcrafted: 200.000 concepts

Aalto University

KOKO Ontology
Is based on vocabularies maintained by national library of Finland
and special interest groups



UNIVERSITY OF HELSINKI

SeCo semantic computing

Aalto University

ONTOLOGIES ARE NOT COMPLETE

UNIVERSITY OF HELSINKI

SeCo semantic computing

Ontology learning from text corpora: LSA

Aalto University

Concept 1	Concept 2	Method	Human
product	production costs	1.0	0.5
Irish	immigrants	0.99	0.5
guerilla	child soldier	0.99	1.0
update	ASP	0.97	0.25
wizard	giants	0.96	0.75
flu	nutrition	0.93	1.0
drinks	chemicals	0.93	0.0
books	shelf	0.90	1.0
studios	gardener	0.82	0.0
foundations	organist	0.81	0.0
suicide attack	population group	0.79	0.5
church	plain	0.69	0.0
symbols	rose	0.68	0.0
leather	persistence	0.66	0.0
sick	disease	0.63	1.0

UNIVERSITY OF HELSINKI

SeCo semantic computing

Relation approximation from Vocabularies

Aalto University

Concept 1	Concept 2	Method	Human
parents	father	0.92	1
masonry	construction work	0.89	1
minorities	population group	0.88	1
document	application	0.84	1
expression	crying	0.84	0.5
measurement	weighing	0.84	1
sportsman	jockey	0.75	1
novel	story	0.82	1
trial	preliminary investigation	0.81	0.75
stone	marble	0.81	0.75
turkey	chicken	0.81	0.75
anecdote	fairy tale	0.77	0.5
windows	stairs	0.77	0.5
near relative	role	0.77	0
change	boiling	0.71	1

UNIVERSITY OF HELSINKI

SeCo semantic computing

Humans appreciate both methods almost equally even as a general domain knowledge

Still the ”best” approximations have less than 5% overlap!

CONCLUSION: Document collection specific relations must be learned

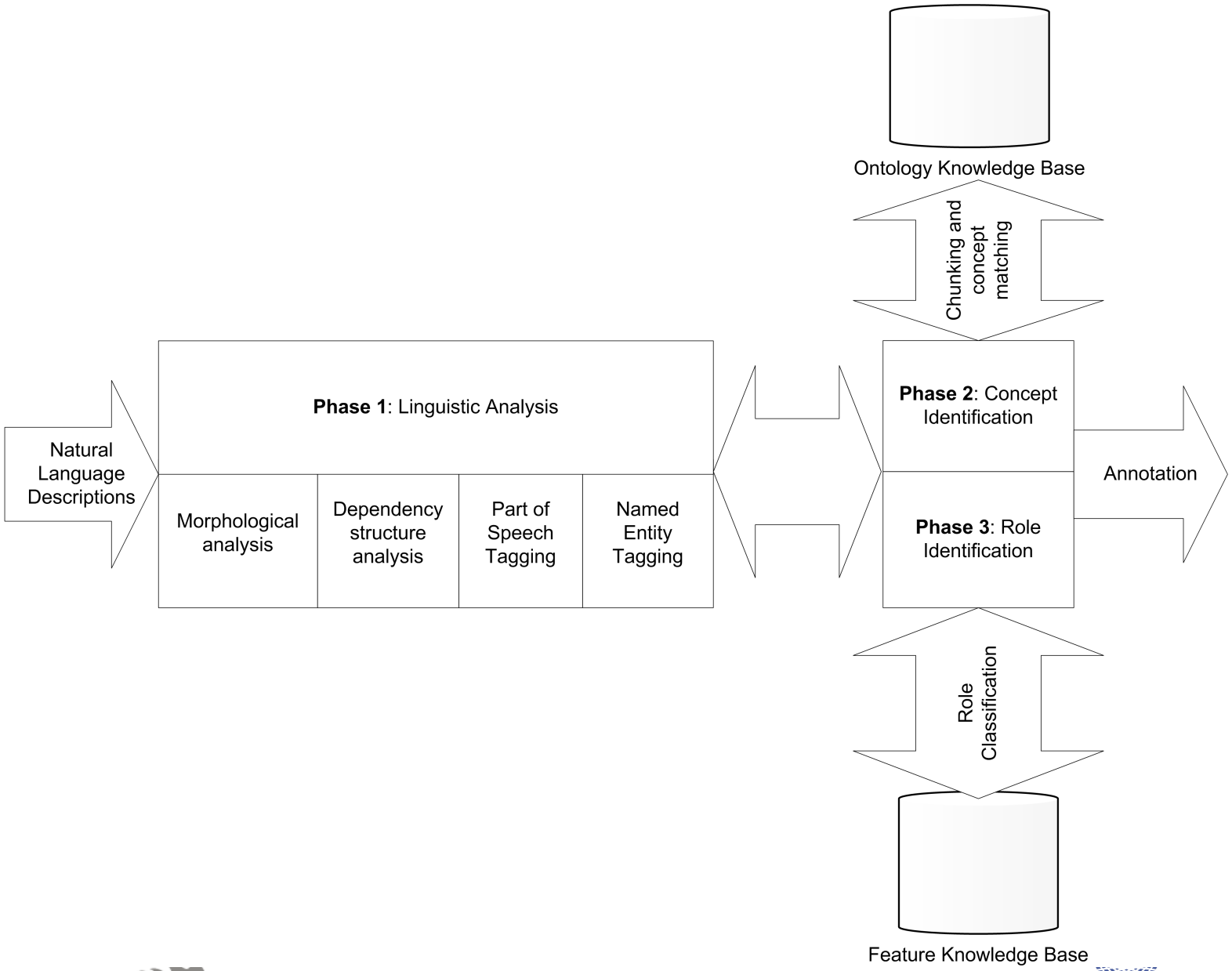
ANNOTATIONS, WHERE DO THEY COME FROM?

The queen is portrayed in her full regalia. Everything here emphasises her royal status: the crown, her ermine robes and the canopy. This official portrait was made in the studio of the painter Frans Pourbus II. It is a copy of the portrait in the Louvre in Paris. The queen is Marie de Médicis (1573-164), a member of the renowned Italian family and wife of the King of France, Henry IV (1553-1610), whom she married in 1600. Henry IV was an ally of the Dutch Republic in the struggle against Spain. In 1638 Marie de Médicis, now a widow, visited Amsterdam, where she was received with great ceremony. For the occasion Joachim van Sandrart painted a militia painting in which she is also portrayed.

Role	Explanation	Value range
Work type	Work type identifies the specific type of artwork being described.	AAT, WN
Title	The title or identifying phrase given to an artwork.	Literal
Material	The substance of which an artwork is composed.	AAT, WN
Technique	The production or manufacturing processes, techniques, and methods incorporated in the fabrication or alteration of the artwork.	AAT, WN
Creator	The names, appellations, or other identifiers assigned to an individual, group, or corporate body that has contributed to the design, creation, production, manufacture, or alteration of the artwork.	ULAN, Literal
Creation Date	Date or range of dates associated with the creation, design or production of the artwork.	Literal
Repository location	The geographic location and/or name of the repository locations entity whose boundaries include the artwork.	Literal
Creation location	The geographic location and/or name of the creation locations entity whose boundaries include the artwork.	TGN
Style period	A defined style, historical period, group, school, dynasty, movement, etc. whose characteristics are represented in the artwork.	AAT, WN
Cultural context	The name of the culture, people (ethnonym), or adjectival form of a country name from which a Image originates, or the cultural context with which the artwork has been associated.	AAT, WN
Subject Term	Terms or phrases that describe, identify, or interpret the artwork and what it depicts or expresses. These include generic terms that describe the work and the elements that it comprises.	WN, AAT
Subject People	Terms or phrases that describe, identify, or interpret particular people.	ULAN, WN, Literal
Subject Location	Terms or phrases that describe, identify, or interpret geographic places.	TGN, WN
Subject Date	Terms or phrases that describe, identify, or interpret time.	Literal

Aalto University

eCo
ntic computing

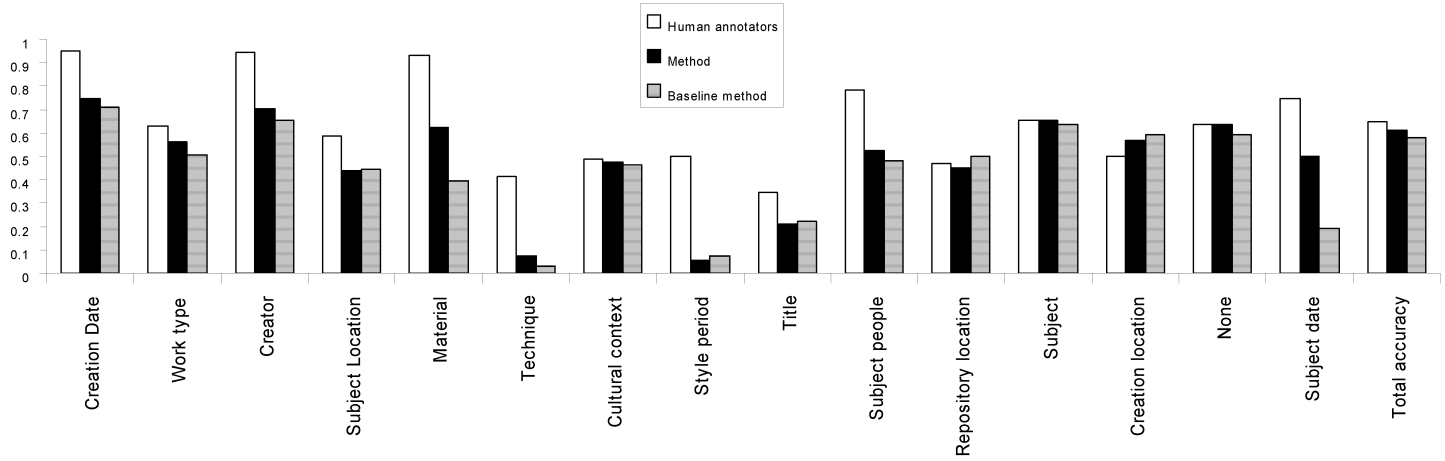

Ontology Knowledge Base
Chunking and concept matching
Natural Language Descriptions
Phase 1: Linguistic Analysis
Morphological analysis
Dependency structure analysis
Part of Speech Tagging
Named Entity Tagging
Phase 2: Concept Identification
Phase 3: Role Identification
Annotation
Role Classification
Feature Knowledge Base

- Metadata that reflect the semantics i.e. the meaning captured in the descriptions
- Metadata used in faceted search, browsing, …

The queen is portrayed in her full regalia. Everything here emphasises her royal status: the crown, her ermine robes and the canopy. This official portrait was made in the studio of the painter Frans Pourbus I. It is a copy of the portrait in the Louvre in Paris. The queen is Marie de Médicis (1573-164), a member of the renowned Italian family and wife of the King of France, Henry IV (1553-1610), whom she married in 1600. Henry IV was an ally of the Dutch Republic in the struggle against Spain. In 1638 Marie de Médicis, now a widow, visited Amsterdam, where she was received with great ceremony. For the occasion Joachim van Sandrart painted a militia painting in which she is also portrayed.

Results 1/2

- Humans 65.1% accuracy (F1) (Kappa 0.58)
- Method 61.2 % accuracy (F1) (Kappa 0.54)
- Baseline 57.8 % accuracy (F1) (Kappa 0.49)
- Differences signifigant (p<0.01), some roles signifigance (p<0.05)



Aalto University

CONCLUSION: With supervised rich linguistic models we can get over 90% accuracy compared to humans.

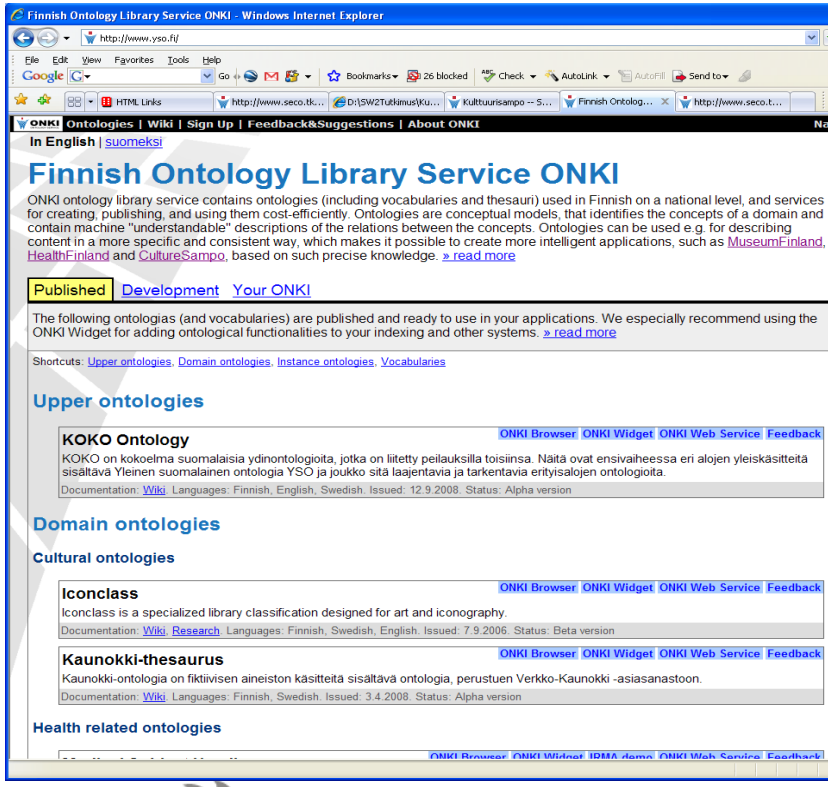
Humans inter-annotator-agreement 65%, Kappa 0.58

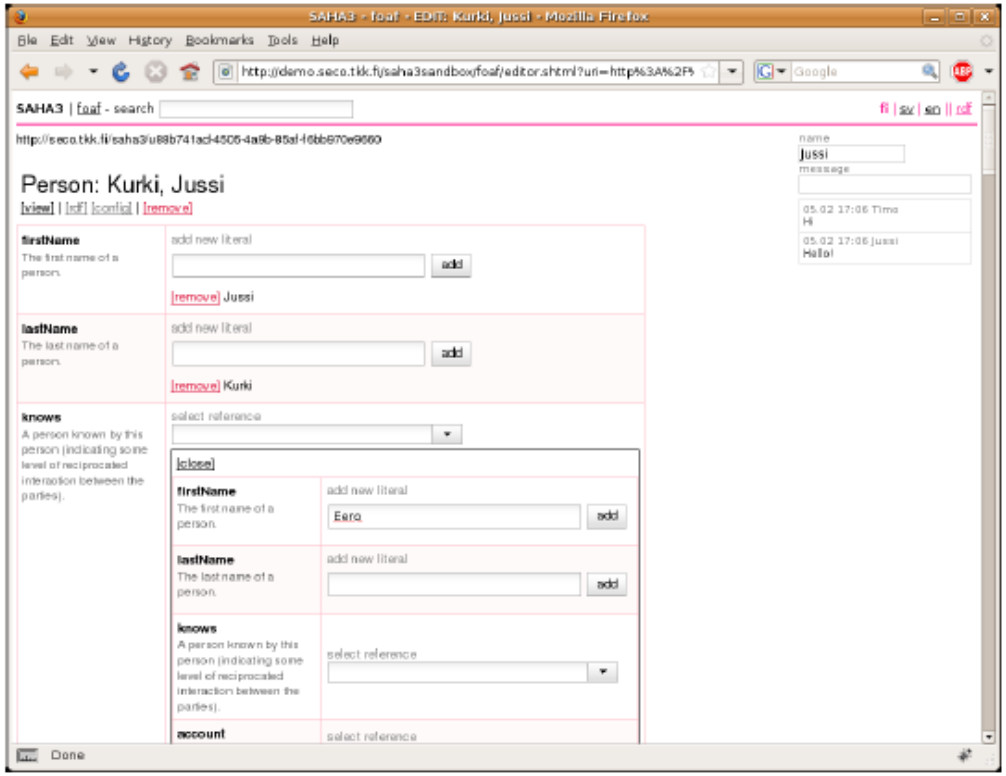
UNIVERSITY OF HELSINKI

SeCo semantic computing

Aalto University

WE ALSO DEVELOP TOOLS FOR MANUAL AND SEMIAUTOMATIC ANNOTATION





UNIVERSITY OF HELSINKI

SeCo semantic computing

ANNOTATION SCHEMAS ARE INCOMPATIBLE

Metadata as events

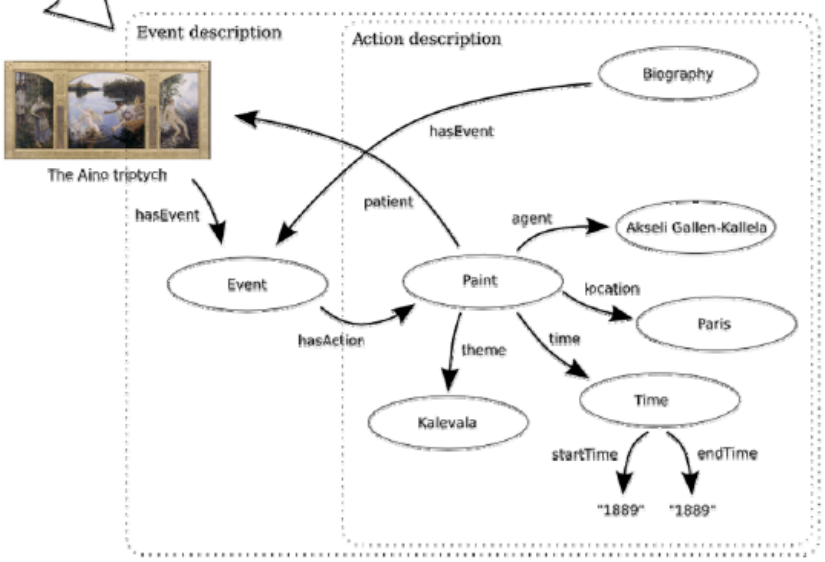
Aalto University

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 ...



Idea: Represent heterogeneous metadata in terms of homogenous real world semantics
->Interoperability
->Simple basis for semantic search and browsing

UNIVERSITY OF HELSINKI

SeCo semantic computing

Aalto University

RETRIEVAL METHODS FOR DATA WITH SEMANTICAL MARKUP

UNIVERSITY OF HELSINKI

SeCo semantic computing

Aalto University

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Nine Thematic Perspectives into Cultural Heritage

Nine perspectives

Aalto University

CultureSampo - Finnish Culture on the Semantic Web 2.0
(the portal is only partially available in English, sorry) - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.kulttuurisampo.fi/index.shtml

Most Visited | Customize Links | Free Hotmail | Windows Marketplace | Windows Media | Windows

CultureSampo - Finnish Culture o... | The 7th International Semantic Web Co... | Kulttuurisampo - Semantic Computing R...

CultureSampo

Finnish Culture on the Semantic Web 2.0
(the portal is only partially available in English, sorry)

Instructions Main Page About Feedback

Language: suomi svenska English

Quick Search: organized Map Browse and Search Relation Search Search and Organize Collections Finnish History Skills Biographies Kalevala Karelia

Map Browse and Search
Browse past and present places, areas and maps, and examine associated items.

Relation Search
How is Akseli Gallen-Kallela connected to the French emperor Napoleon I?

Search and Organize
Set the material into contexts by approaching it from different angles.

Collections
Which collections are included in CultureSampo?

Finnish History 1860-1919
Browse historical events.

Skills
Familiarize yourself with semantically described skills.

Biographies
Browse semantically enriched biographies.

Semantic Kalevala
Browse Kalevala on the semantic web.

Information about Karelia
Material about Karelia on Wikipedia, semantically enriched.

Last Viewed:

novel or short story : Pianonsoittaja

historical event : Kansainvälinen Meteri-sovinto

Aerial photo : Alahärmä, Härmän Parantola

photograph : Karjalaisen osakunnan kesäjuhla, L. Nissisen (isännän) vieraina

museum item : Sirppi

Last Commented:

Aerial photo : Turku, 1960

Aerial photo : Lauritsala, Kaukas Fabrik Ab

Aerial photo : Lauritsala, 1968

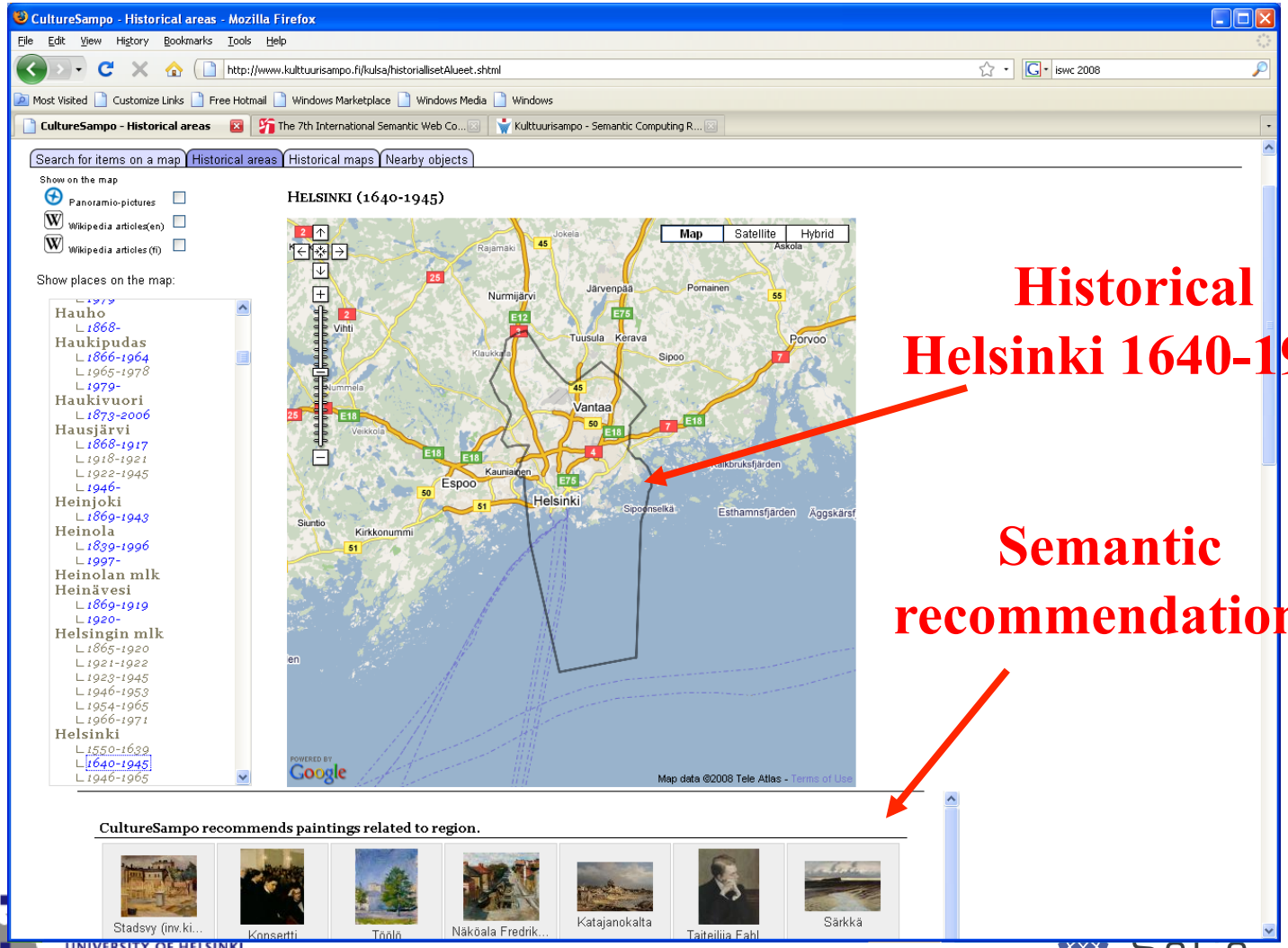
Aerial photo : Lauritsalan saha, Lauritsala, 1958

Aerial photo : Lauritsala, Kaukas Fabrik Ab

semantic computing

Map Browse and Search

Aalto University



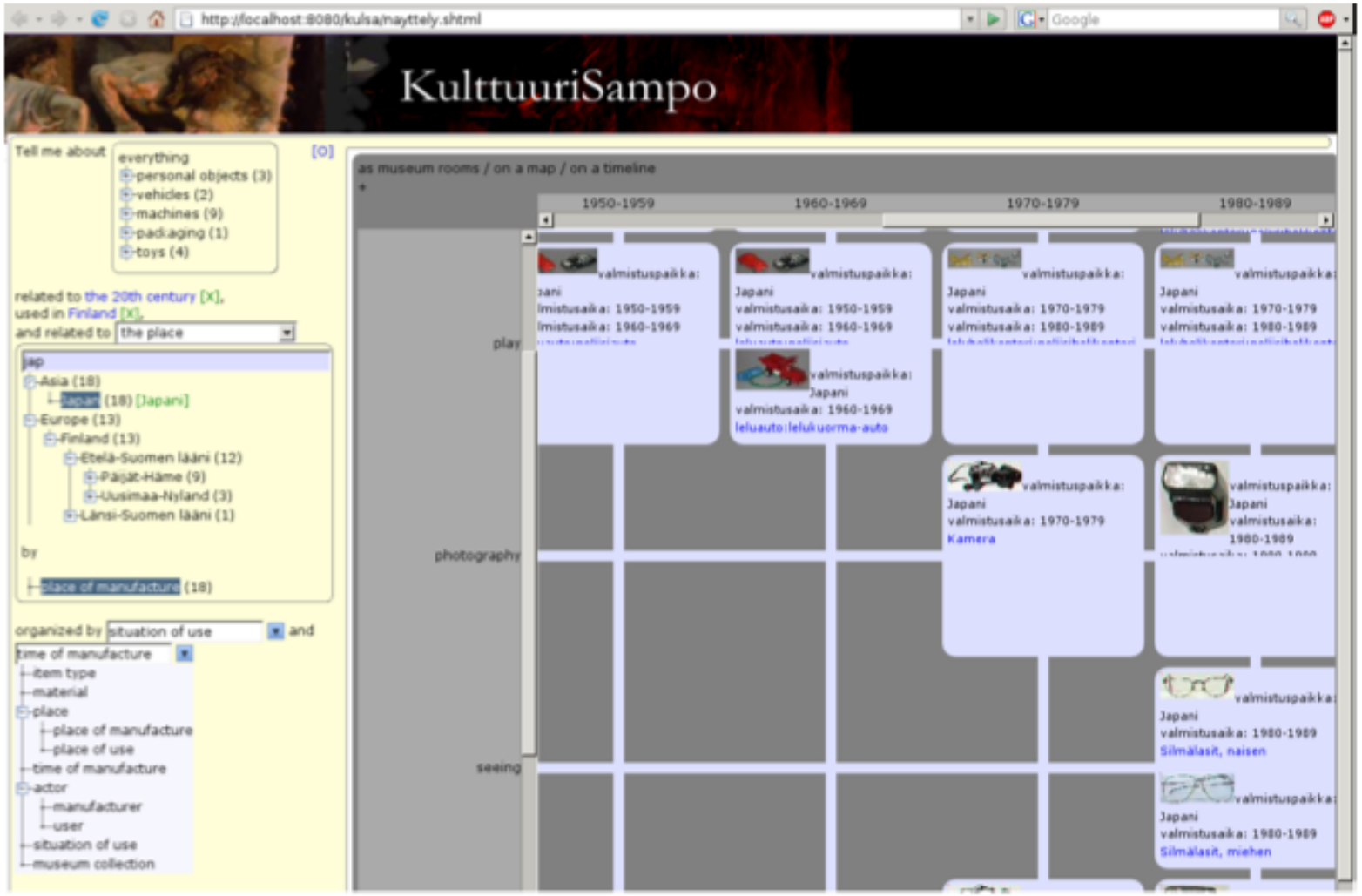
Historical Helsinki 1640-1945

Semantic recommendations

UNIVERSITY OF HELSINKI

SeCo semantic computing

Dynamic facets and matrix view

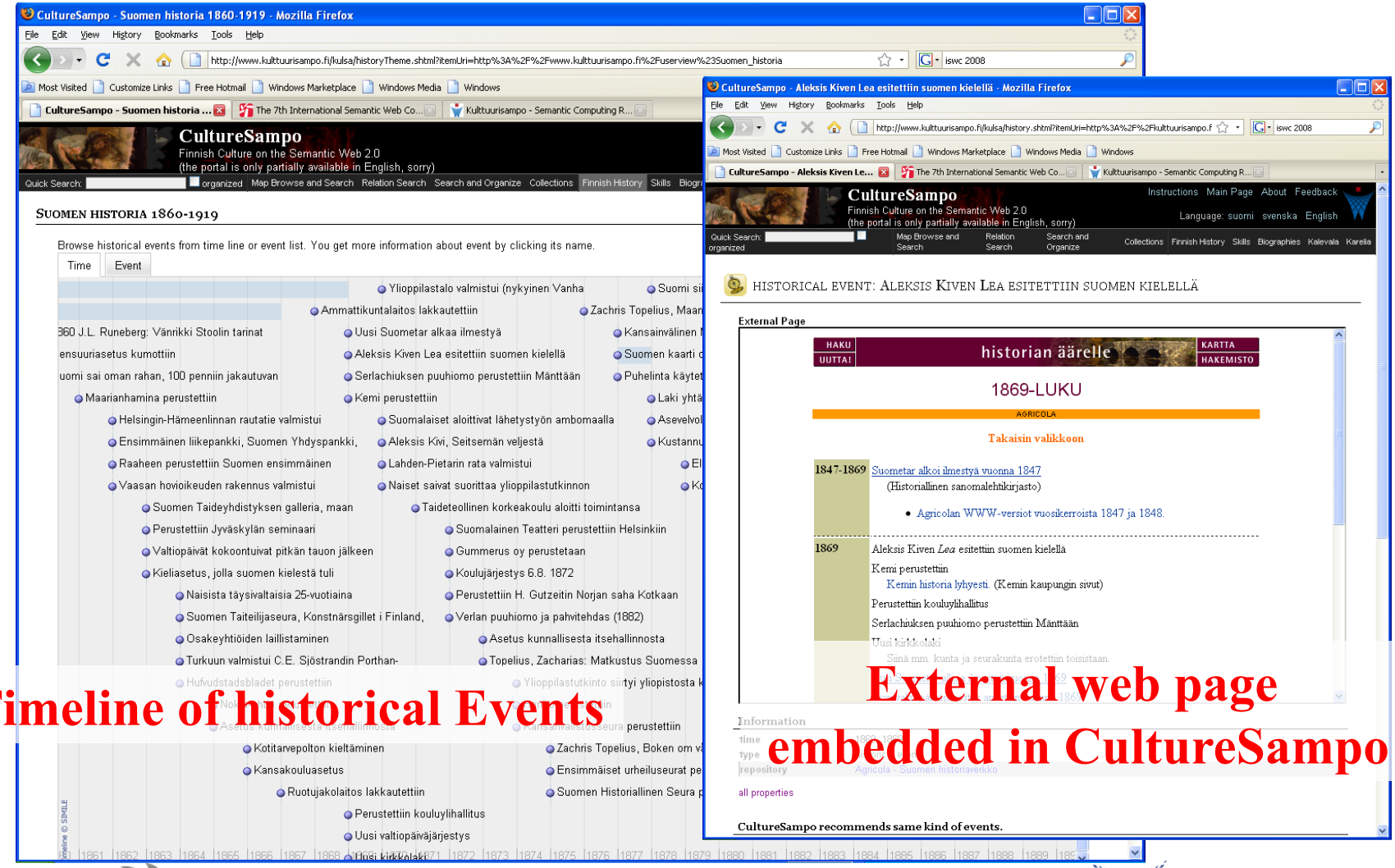


UNIVERSITY OF HELSINKI

SeCo semantic computing

Finnish History
Co-operation with History Researchers

Aalto University



Timeline of historical Events

External web page embedded in CultureSampo

UNIVERSITY OF HELSINKI

SeCo semantic computing

Recommender systems and semantic information retrieval

Aalto University

KulttuuriSampo

Navigointi

Takaisin hakusivulle

OSAKUNTALAISIA MATKALLA KOLILLE

Kohdetiedot

mitat:	8 x 11 cm
type:	valokuva
kuvaajat:	Wartiainen, I. M.
kokoelmassa:	Kansatieteen kuvakokoelma
valmistuspaikka:	Pielinen
tekijät:	Wartiainen, I. M.
nimi:	osakuntalaisia matkalla Kolille

Asiasanat

osakunnat järvet veneet matkailu

Narratiivinen kuvailu

Toiminta
retkeily
liittyy **järvet**
liittyy **osakunnat**

Toiminta

KulttuuriSampo suosittelee

Karjalaisen osakunnan jäseniä Kaarnesaaren laella

Perustelut

retkeily
liittyy
osakunnat

retkeily
liittyy
järvet

osakunnat
liittyy
osakunnat

järvet
liittyy
järvet

matkailu
liittyy
matkailu

Valokuvaus
tekijä
Wartiainen, I. M.

Sijaita
paikka
Kansatieteen kuvakokoelma

Karjalaisen osakunnan jäseniä Kaarnesaaren laella

Perustelut

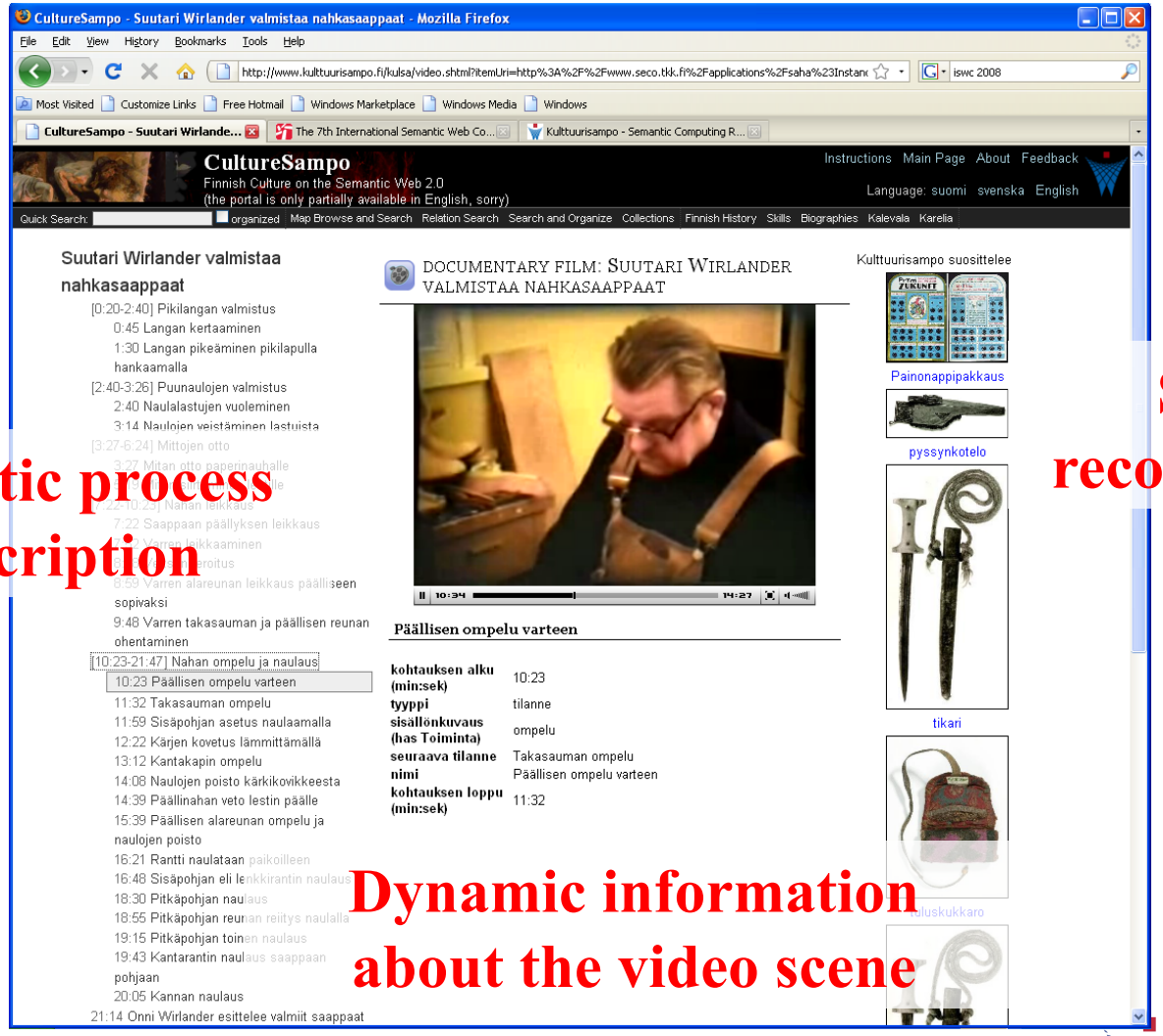
Metadata of the original annotation

Explicated thematic event descriptions

Recommendation explanations using thematic event descriptions

Skill Documentations
A Semantic Video Viewer

Aalto University



Semantic process description

Semantic recommendations

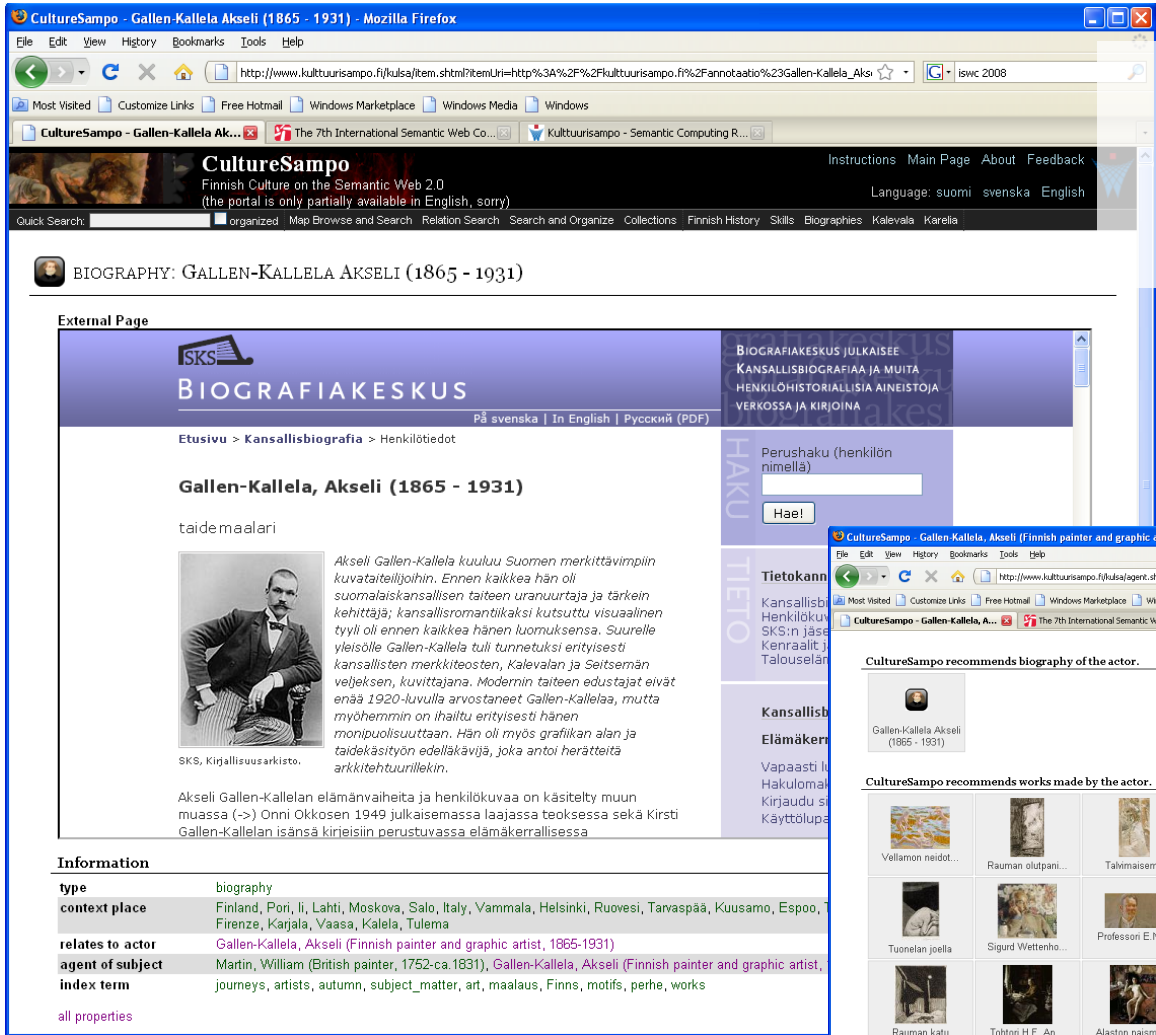
Dynamic information about the video scene

UNIVERSITY OF HELSINKI

SeCo semantic computing

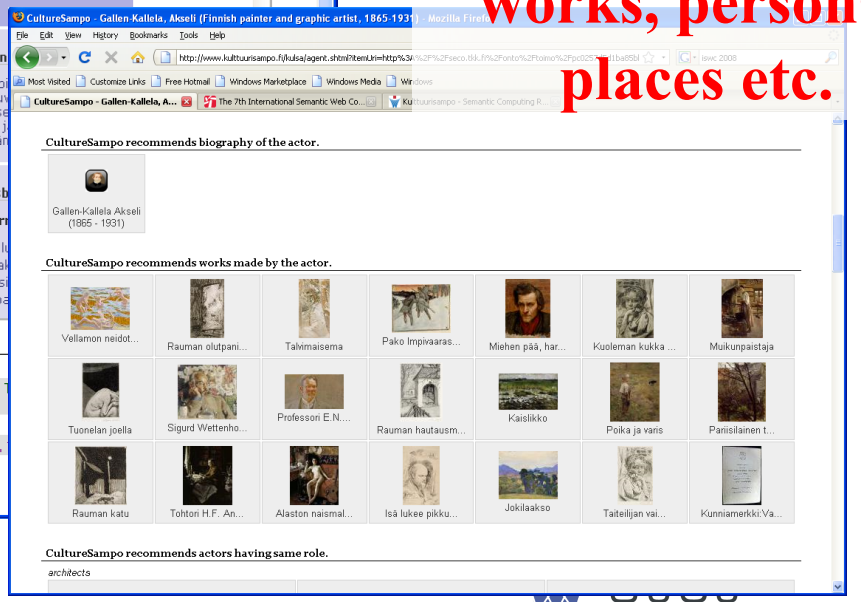
Narrative Perspective
National Biography on the Semantic Web

Aalto University



Semantically annotated biographies

Links to related works, persons, places etc.



UNIVERSITY OF HELSINKI

semantic computing

Narrative Perspective: The national epic Kalevala

Aalto University

CultureSampo - Neljäs runo - Mozilla Firefox

CultureSampo
Finnish Culture on the Semantic Web 2.0
(the portal is only partially available in English, sorry)

POEM: NELJÄS RUNO

Kalevala
Runot
Ensimmäinen runo
Toinen runo
Kolmas runo
Neljäs runo
Viides runo
Kuudes runo
Seitsemäs runo
Kahdeksas runo
Yhdeksäs runo
Kymmenes runo
Yhdestoista runo
Kahdestoista runo
Kolmastoista runo
Neljästoista runo
Viidestoista runo
Kuudestoista runo
Seitsemästoista runo
Kahdeksastoista runo
Yhdeksästoista runo
Kahdeskymmenes runo
Yhdeskolmatta runo
Kahdeskolmatta runo
Kolmaskolmatta runo
Neljäskolmatta runo
Viideskolmatta runo

Neljäs runo

Tuopa Aino, neito nuori,
sisar nuoren Joukahaisen,
läksi luutoa lehosta,
vastaksia varvikosta.
Taittoi vastan taatollensa,
toisen taittoi maammollensa,
kokoeli kolmannenki
verevälle veijollensa.

Jo astui kohin kotia,
lepikkoä leuhautti.
Tuli vanha Väinämöinen;
näki neitosen lehossa,
hienohelman heinikössä.
Sanan virkkoi, noin nimesi:
"Eläpä muille, neiti nuori,
kuin minulle, neiti nuori,
kanna kaulanhelmiöitä,
rinnanristiä rakenna,
pane päätä palmikolle,
sio silkillä hivusta!"

Neiti tuon sanoiksi virkki:
"En sinulle enkä muille
kanna rinnanristiöitä
päätä silkillä sitaise.
Huoli en haahen haljakoista,
vehnän viploista valita;
asun kaioissa sovissa,
kasvan leivän kannikoissa
tykönä hyvän isoni,
kanssa armahan emoni."

Riisti ristin rinnaltansa,
sormukset on sormestansa,
helmet kaulasta karisti,
punalangat päänsä päältä,
jätti maalle maan hyviksi,

Aino lähtee metsään taittamaan kotiväelle vastoja (1-8)

Hae suositukset painamalla tekstiä

CultureSampo recommends objects related to poem.

Aino lehdossa, untitled, Haavoittunut en..., Esikaupunkia, Kelkanvetäjät, Aino, untit, Kotitallilla: L..., Johanneksen kir..., Merelle katsova..., Aino-triptyykki..., Mäntykangas Uud..., Kieliportti, ra...

Translation into modern Finnish

Links to related art etc.

Semantically annotated 50 poems of the national epic - Events and narratives

WSA
WORLD SUMMIT AWARD
e-CONTENT+
CREATIVITY

Aalto University

1st Price Winner!

smartmuseum

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Pastoral Scene
Creator: Boucher, Francois
Created: 18th century
The painting depicts a romantic, freely rendered and a monochrome, pastoral scene. The centre of the composition is dominated by a tree, near which a herd of cows and a flock of sheep are
Read more...
iPAQ

ObjectPresentation
3:25
Hans Tucher
Terrestrial telescope
John Marshall [attr.]
Demonstration model of Oersted's experiment
Carlo Dell'Acqua
Thunder fa\u00E7ade
Objective lens
120 KB
Select an artifact
History
Type N#
Settings
Exit

SmartMuseum Browser - smartmuseum - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://localhost:8080/kulsa/smartmuseumInside.shtml

Most Visited Smart Bookmarks Getting Started Latest Headlines Kymmenen päivän s...

SmartMuseum Browser - sm... SmartMuseum Browser - sm...

Single-handed dividers

Single-handed dividers

Single-handed dividers

-- objectType angle measuring devices -- objectType distance measuring devices -- ageGroup adults -- tyyppi Instrument -- subject science

Sundial

Theodolite

Theodolite

Surveying compass

Surveying compass

-- ageGroup adults -- subject science -- tyyppi Instrument

Galileo's objective lens

Galileo Galilei, "Sidereus

Subject matter

Location

Temporal

Context

Done

Applications Places System

Thu Sep 3, 1

Aalto University

ACCURACY WITH ONTOLOGY-BASED QUERY EXPANSION AND FEATURE BALANCING INCREASED ACCURACY AROUND 10%

CLUSTERING HYPOTHESIS ON RETRIEVAL RESULTS CONFIRMED TO BE USEFULL. USERS APPRECIATE DIFFERENT POINTS OF VIEW.

USERS REPORTED 80-90% SATISFACTION TOWARD THE SYSTEM IN FIELD EXPERIMENTS IN TWO MUSEUMS (SUS survey)

UNIVERSITY OF HELSINKI

SeCo semantic computing

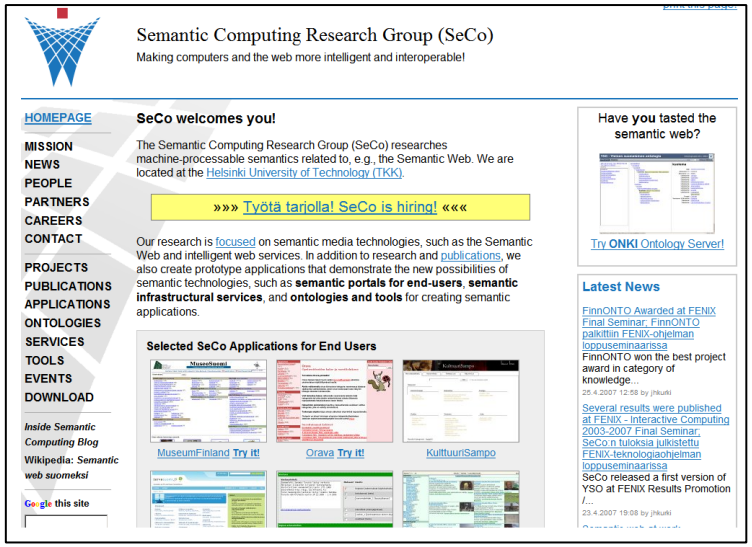
Current interests

- Domain independent information extraction using Frame Semantics
- Narrative information composition and retrieval
- Search Results Clustering using blind source separation
- Dynamic faceting
- Combining Web of Things, Semantic Web and Linked Open Data

Questions?

Aalto University

?!



Demos, Publications, Services, Content, Projects:
http://www.seco.tkk.fi/

www.smartmuseum.eu
www.kulttuurisampo.fi