

# Semanttisten annotaatioiden tuottaminen selainpohjaisella annotaatioeditorilla

16.11.2005

Olli Alm & Onni Valkeapää

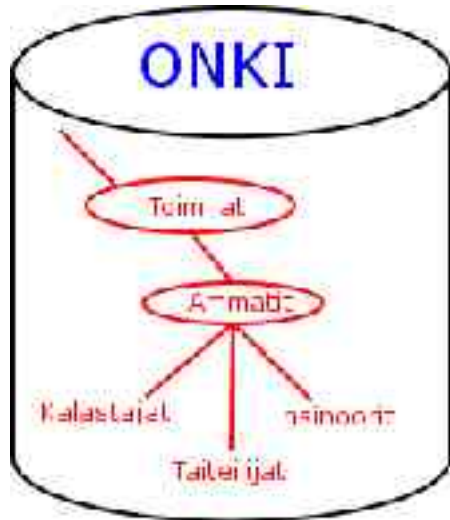


# Annotointi

- annotointi on metadatan, eli tietoa kuvailevan tiedon tuottamista
- metadataa tarvitaan semanttisen webin palveluiden tuottamisessa
- ontologiaperustaisen metadatan avulla tietokoneet pystyvät jäsentämään aineiston älykkäästi

# Annotointiprosessi

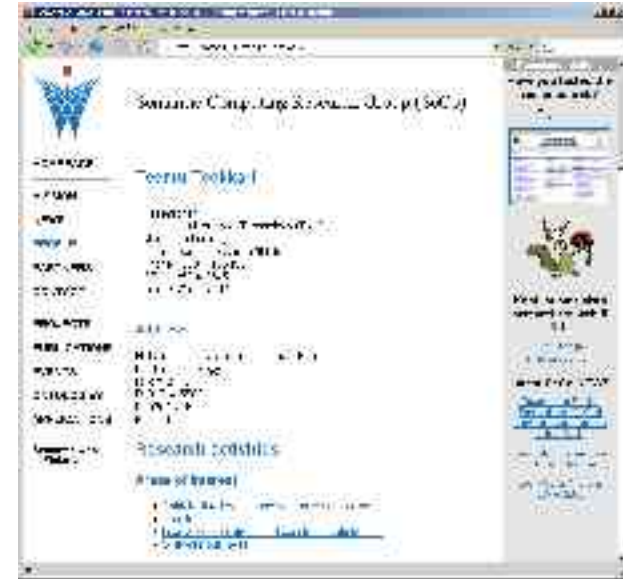
Ontologiat



Skeema



Dokumentti



Metadata

# SAHA-annotaatioeditori

- Annotaatioiden tuottamiseen tarvitaan helppokäyttöisiä työvälineitä
- Annotaatioita tulee voida tuottaa hajautetusti
- SAHA on selaimella käytettävä annotaatioeditori
- Useat eri henkilöt voivat annotoida samaa dokumenttijoukkoa, annotaatiot tallennetaan keskitetysti palvelimelle



# SAHA - Annotaatioeditori

Ohjeistus Ohjeet

## Annotaatio

Linkkiosoitteita http://...  
Luokka: ...  
Annotaatio: ...

### Toiminnot:

- Sulje annotaatio...
- Päivitä annotaatio...
- Säilytä dokumentti...

Ominaisuus:	Arvo	Toiminnot
Annotointi:		
Tyylit:		
Käsitteet:		
Sähköpostiosoitte:		
Nimi:		

## Ominaisuuksien syöttö

Annotointi

Tyylit

Käsitteet

Sähköpostiosoitte

Nimi

Säilytä dokumentti

Annotoitava dokumentti

Skeema

Täytettävät ominaisuudet

ot:

tio...

atio...

ntti...

## Ominaisuuksien syöttö

### Ammatti

Hakusana Onkille

insi

Etsi Onkista

### Kollega

Valitse instanssi...

Luo uusi instanssi...

### Kotisivu

### Sähköpostiosoite





ot:

lio...

atio...

ntti...

## Ominaisuuksien syöttö

Onkista saatu käsite

### Ammatti

### Kollega

### Kotisivu

### Sähköpostiosoite

ot:

io...

tio...

tti...

## Ominaisuuksien syöttö

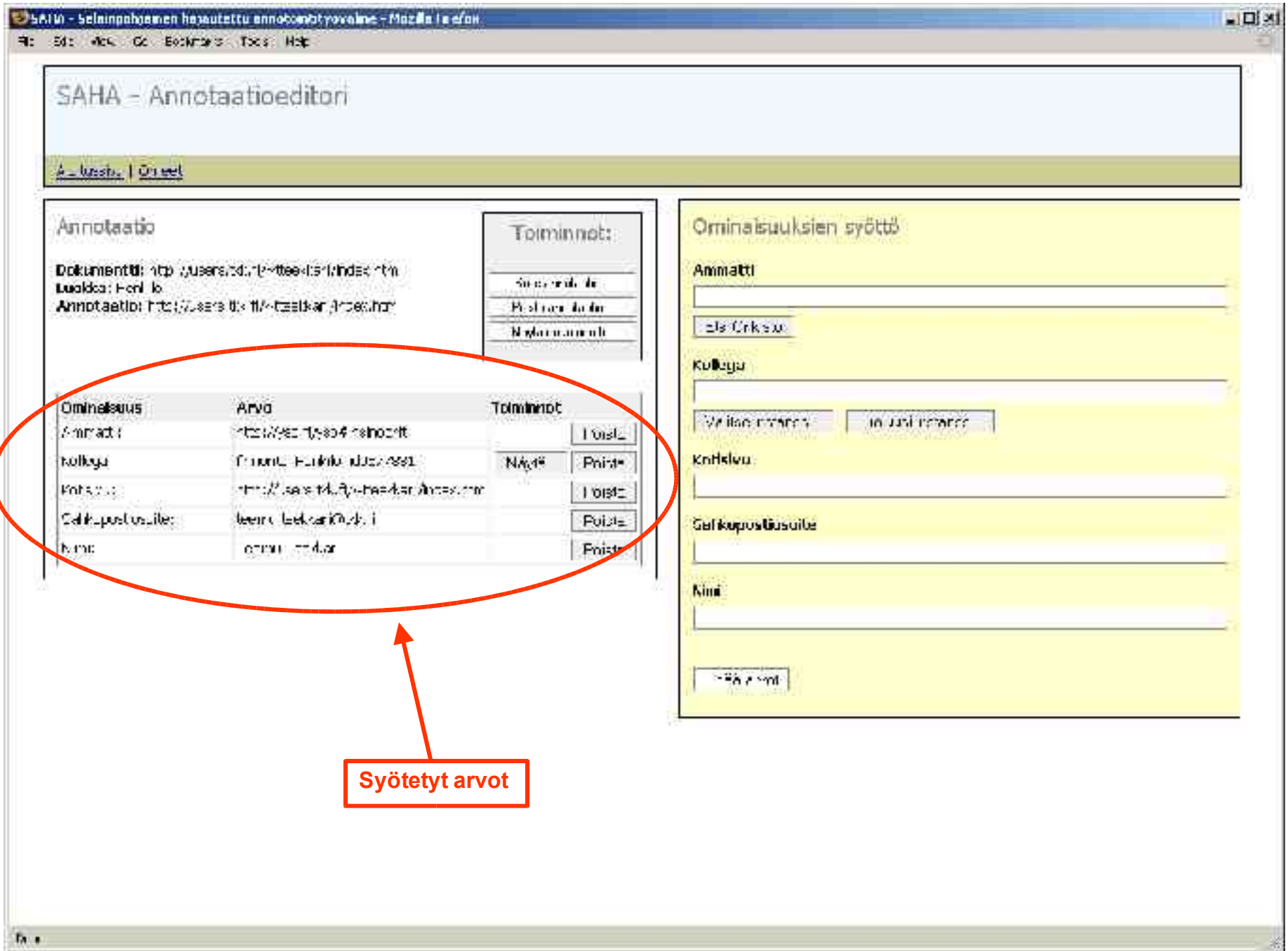
### Ammatti

### Kollega

### Kotisivu

### Sähköpostiosoite

### Nimi



# SAHA - Annotaatioeditori

[A - tulos](#) | [Oikeet](#)

## Annotaatio

Dokumentti: http://users.tkk.fi/~teekeis/index.html  
Luokka: Foni 10  
Annotaatio: http://users.tkk.fi/~teekeis/index.html

## Toiminnot:

- Sisällysluettelo
- Historia
- Näytä

Ominaisuus	Arvo	Toiminnot
Annamatti	http://www.tkk.fi/~teekeis/	<input type="button" value="Tulosta"/>
kollage	Frontend: Foni10, Järjestelmä: Foni10	<input type="button" value="Näytä"/> <input type="button" value="Print"/>
Kohde	http://www.tkk.fi/~teekeis/index.html	<input type="button" value="Tulosta"/>
Gefikoputusuute	teem: teekari@tkk.fi	<input type="button" value="Print"/>
nimi	teekeis	<input type="button" value="Print"/>

## Ominaisuuksien syöttö

Annamatti

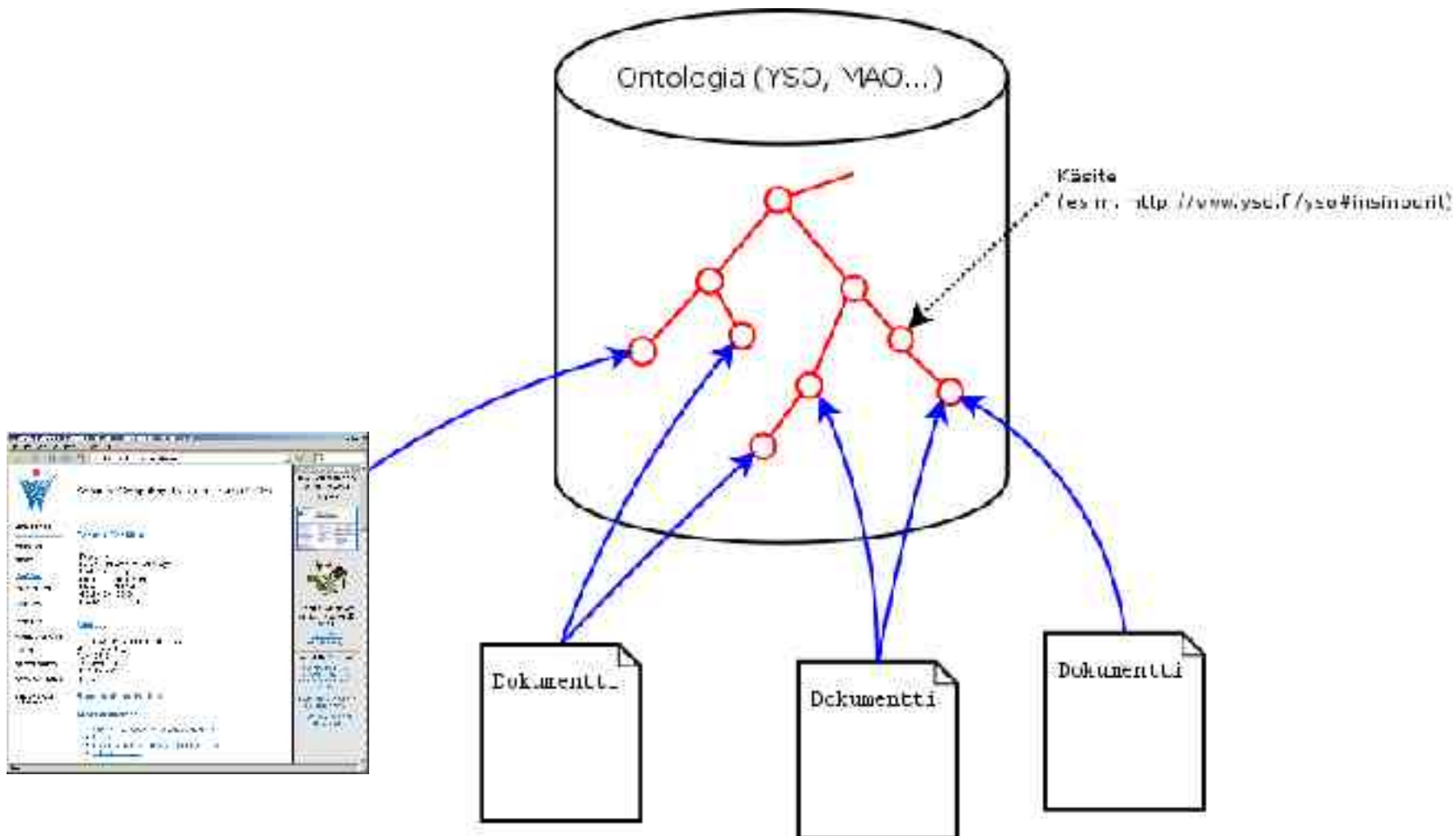
kollage

Kohde

Gefikoputusuute

nimi

**Syötetyt arvot**



# Puoliautomaattinen annotointi

- puoliautomaattisella annotoinnilla tarkoitetaan kieliteknologisilla apuvälineillä automatisoitua aineiston kuvailua
- automatisointi keskeistä laajojen aineistojen annotoinnissa
- tietokone voi avustaa käyttäjää tunnistamalla tekstistä
  - ontologian käsitteitä
  - henkilöitä sekä muita nimettyjä toimijoita
  - yksittäisistä käsitteistä muodostuvia tapahtumia
- annotointityökalun käyttäjä varmistaa laadun vahvistamalla halutut annotaatiot ja hylkäämällä väärät

SAHA - Annotaatioeditori

SAHA - Annotaatioeditori

Annotaatio

Dokumentti: http://www.tkk.fi/~teemu/teemu/index.shtml  
 Osoite: http://www.tkk.fi/~teemu/teemu/index.shtml  
 Annotaatio: teemu.teekkar@tkk.fi

Toinninnot

Suoritaannotaatio  
 Poistaannotaatio  
 Tyhjennäannotaatio

Ominaisuus	Arvo	Tunnus	toiminto
Arvo	http://www.tkk.fi/~teemu/teemu/index.shtml		Poista
URL	http://www.tkk.fi/~teemu/teemu/index.shtml		Poista
Käsitteet	http://www.tkk.fi/~teemu/teemu/index.shtml		Poista
Semanttinen	teemu.teekkar@tkk.fi		Poista
Nimi	teemu.teekkar		Poista

Ominaisuuksien syöttö

**Teemun** kotisivu  
 <person ID:1> <concept>

Tämä on **Teemu Teekkarin** kotisivu.  
 <person ID:1>

Olen tutkijana **Teknillisellä korkeakoululla**  
 <university>

**viestintätieteiden laboratoriossa.**  
 <department>

Tutkimuskohteina **minua** kiinnostavat  
 <person ID:1> <interest-in>

erityisesti **semanttinen laskenta** ja **tekoäly**.  
 <research: SW> <AI>

Sähköpostiosoitteeni on **teemu.teekkar(at)tkk.fi**.  
 <email ID:1>

Tunnistettuja  
 käsitteitä

tutkimuskohteina

henkilö: Teemu

verbi: kiinnostaa

viestintätekniiikan laboratoriossa.  
<department>

Annotoidaan tapahtuma:  
tekijä: Teemu Teekkari  
toiminta: on kiinnostunut  
kohde 1: semanttisesta laskennasta ja  
kohde 2: tekoälystä  
tarkennus: tutkimuskohteina.

Sähköpostiosoitteeni on **teemu.teekkari(at)tkk.fi**.  
<email ID:1>

tutkimusala: Semanttinen verkko

aihe: Tekoäly