

Ontologiakäytön askeleet

Timo Westkämper
Mysema Oy



Tarve

Tarve formaalille sanasto-käytölle

Sanaston rajaaminen vaikeata ilman formaalia viitekehystä

Uudelleenkäyttö ja jakaaminen myös vaikeata

Linkitys mahdotonta ilman yhteisten standardien käyttöä

Ontologia ja sanasto-standardit avuksi



Ontologia-standardit

RDF Schema (1998 julkaistu, 2004 W3C standardiksi)

OWL (Web Ontology Language)

1.0 (2004)

Light, DL (Description Logic) ja Full dialekteinä

2.0 (2009)

EL, QL ja RL profiilit

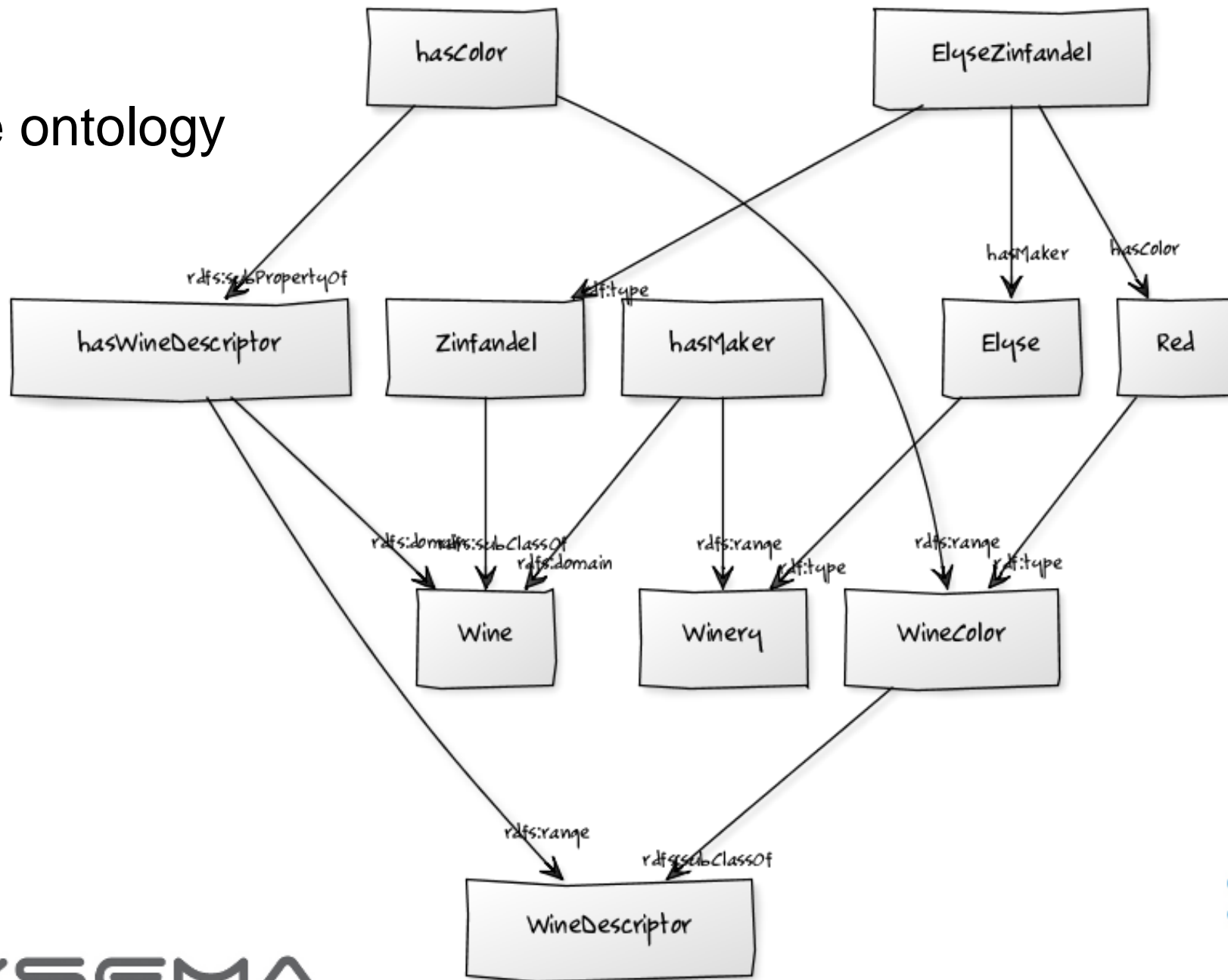
SKOS (Simple Knowledge Organization System)

Sanastot, Thesaurukset etc



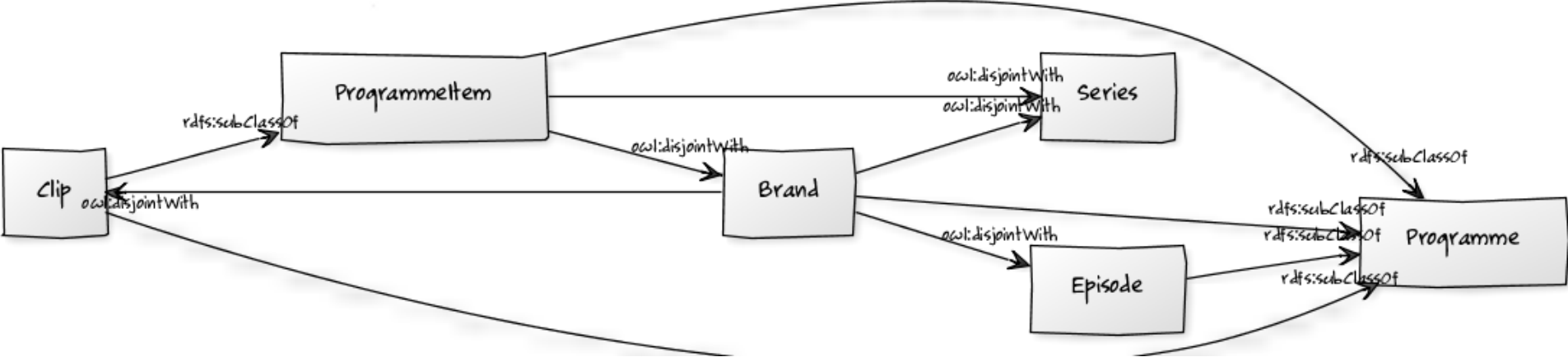
RDF Schema esimerkki

Wine ontology



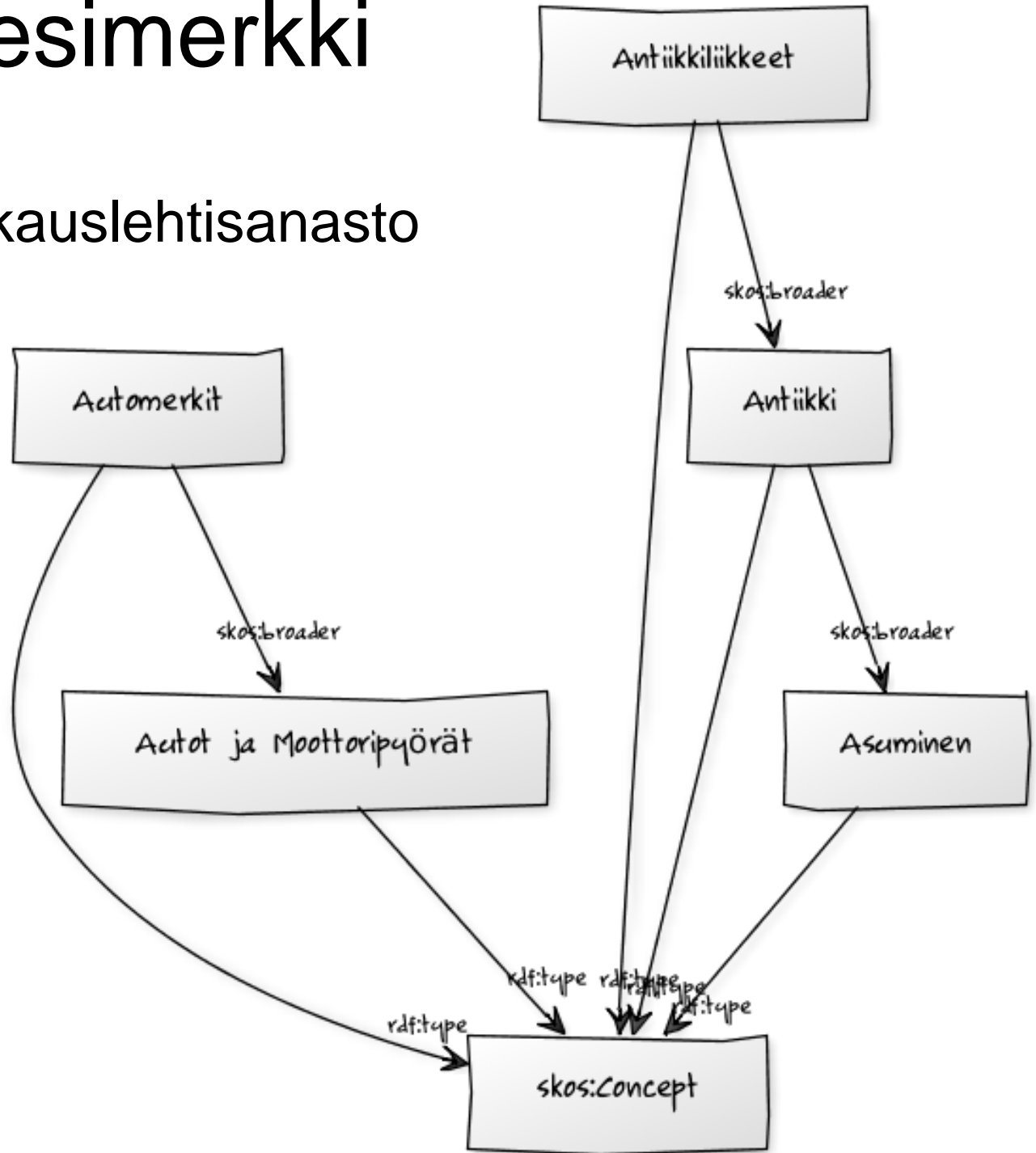
OWL esimerkki

BBC Programmes



SKOS esimerkki

ONKI: Aikakauslehtisanasto



Oman ontologian luonti

Käsittemallinnus

Päämallinnusdimensioiden määrittely
aliluokitus, alikäsitteet, ryhmittely

Formaatin valinta (RDF/S, OWL tai SKOS)

Toteutus (itse vai ulkoistettu, pelkkä ontologia vai palvelu)

Ylläpito, jatkokehitys, jakelu



Omien ontologioiden suhde ulkoisiin

Käyttö

esim. predikaattien arvoavaruus

Laajennus

aliluokat, alipredikaatit, suppeammat ja laajemmat käsitteet

Viittaukset

liittyy, sama kuin jne.



Olemassa olevan ontologian käyttö

ONKI

<http://onki.fi>

semanticweb.org

<http://semanticweb.org/wiki/Ontology>

Swoogle

<http://swoogle.umbc.edu/>



ONKI

72 ontologiaa

47 OWL muodossa

2 RDF Schema muodossa

22 SKOS muodossa

Tuki

email

ilmoitukset Twitterin kautta



ONKI pohjaisen ontologian käyttö

ONKI palvelun HTML pohjainen käyttö (ONKI Browser)

Ohjelmalliset rajapinnat (SOAP/XML, REST/JSON beta)

Integraatio omaan Web-sovellukseen (ONKI Selector widget)

Ontologian lataus (RDF/XML)











ONKI Browser

Hae ontologiaa tai sanastoa:

72_{kpl}





Aiheen mukaan

- 8 Yleiskäsitteet 
- 12 Alakohtaiset 
- 4 Julkishallinto 
- 12 Kulttuuri 
- 6 Liike-elämä 
- 14 Luonto 
- 6 Paikkatieto 
- 10 Terveys 

Tyypin mukaan

- 33 Luokkaontologia 
- 17 Yksilöontologia 
- 17 Kehittynyt sanasto 
- 5 Yksinkertainen sanasto 

Julkaisutavan mukaan

- 40 Julkiset ONKI-ontologiat 
- 16 Yhdistelmäontologiat kehittäjille 
- 5 Julkiset OMA ONKI -ontologiat 
- 11 Asiakaskohtaiset ontologiat 

- ▶ Aikakauslehtisanasto 
- ▶ Allmän tesaurus på svenska - Allärs 
- ▶ Creative Commons Licenses 
- ▶ DCMI Type Vocabulary 
- ▶ Disaster ontology 
- ▶ Erityisryhmä-ontologia
- ▶ Eurooppalainen monikielinen terveyden edistämisen sanasto - HPMULTI 
- ▶ Iconclass
- ▶ Integrated Public Sector Vocabulary - IPSV 
- ▶ Julkishallinnon ontologia - JUHO 
- ▶ Kassu - kasvien suomenkieliset nimet 
- ▶ Kaunokki-ontologia - KAUNO 
- ▶ Kielitieteen asiasanasto
- ▶ Kielitieteen ontologia - KTO 
- ▶ Kielten nimien ontologia 



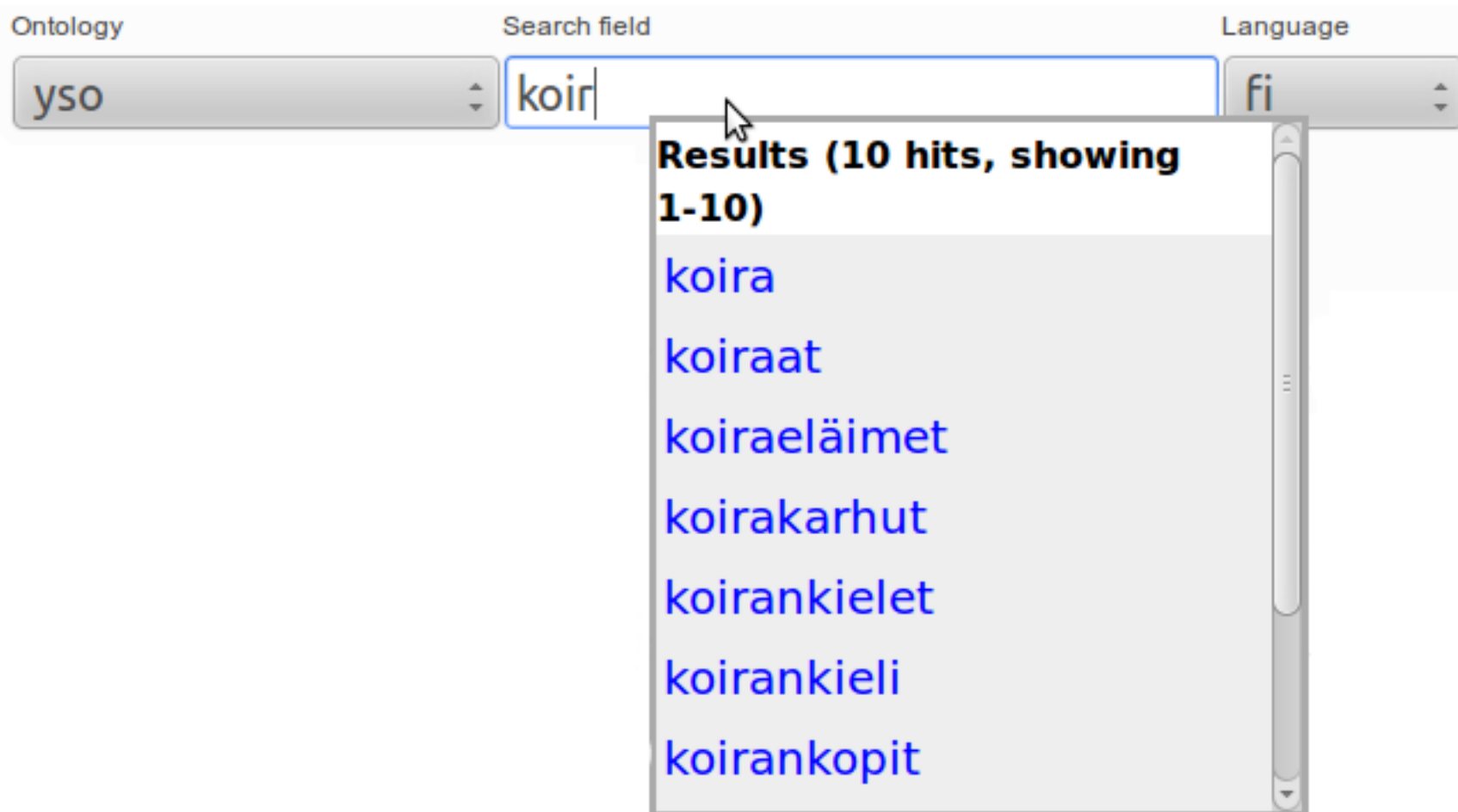
ONKI SOAP-rajapinta

getAvailableLanguages	languages used in concept names in an ontology.
getAvailableTypeUris	supported concept type URIs of an ontology.
search	searching for ontological concepts.
getLabel	fetching a label for the concept with a given URI.
expandQuery	expanding concept(s) with given URI(s) (e.g. to subconcepts) for queries.
getProperties	fetching properties for the concept with a given URI.
getConceptTree	fetching concept hierarchy for the concept with a given URI.



Ontologian käyttö ONKI selector widgetin kautta

Selector koodin käyttöönotto omassa ympäristössä



The screenshot displays the ONKI selector widget interface. It features three main components: an 'Ontology' dropdown menu set to 'yso', a 'Search field' containing the text 'koir', and a 'Language' dropdown menu set to 'fi'. Below the search field, a results panel is open, displaying the text 'Results (10 hits, showing 1-10)' followed by a list of search results: 'koira', 'koiraat', 'koiraeläimet', 'koirakarhut', 'koirankielet', 'koirankieli', and 'koirankopit'. A mouse cursor is positioned over the search field.



ONKI ontologian käyttö omalla palvelimella

Ontologian ja instanssidatan tallennus

SQL (MySQL, PostgreSQL, Oracle jne)

RDF-kanta (4Store, Virtuoso)

NoSQL (Mongodb, CouchDB, Cassandra)

Indeksointi

Päivitykset

Kuinka usein?

Vanhan rinnalle vai sijalle?

Lokalisointi



Kysymyksiä?



Kiitos!

www.mysema.com

