



# TTA-minimimetatietomalli

Mari Kleemola ja Panu Kalliokoski

TTA-seminaari: Metatietomallit ja kokonaisarkkitehtuuri 19.3.2013

# Tässä esityksessä...

- Työryhmä
- TTA-minimimetatietomalli
- Minimimetatiedot käytännössä
  - aineistojen kuvailu Ida-palvelussa
  - tunnisteiden löytäminen
  - minimi vs. DDI-Codebook
- Tulossa

# Työryhmä

- Mari Kleemola, FSD, puheenjohtaja
- Panu Kalliokoski, CSC, sihteeri
- Pekka Järveläinen, CSC
- Kimmo Koskinen, Helsingin yliopiston kirjasto
- Juha Hakala, Kansalliskirjasto
- Osma Suominen, Aalto-yliopisto
- Jaana Kilkki, Kansallisarkisto
- Saija Ylönen, Tilastokeskus
- Pekka Henttonen, Tampereen yliopisto
- Hanna Koivula, HY Luonnontieteellinen keskusmuseo
- Eero Mikkola, Metla

# Työryhmän tehtävät

- Laatii ehdotuksia/suosituksia:
  - standardeista tutkimusaineistojen metatietojen keräämiseen ja vaihtamiseen: metatietoformaatit, rajapinnat ja protokollat
  - kansallisen metatietoinfrastruktuurin vastuunjaosta, aineistokatalogien välisistä rajapinnoista ja tiedon kulkeutumisesta
  - aineistojen kuvailun metatietomallista ja kuvailun vähimmäisvaatimuksista aineistotyypeittäin
- Luo konkreettisia ohjeita tutkimusaineistojen metadatakksi
- Yhteentoimivuus tärkeää

julkaisutietokannat

**Tutkimusprojekti**  
*Hallinnollinen yksikkö, jolla on ihmiset ja rahoitus, tuottaa aineistoja ja julkaisuja*

**Julkaisu**  
*monografia, artikkeli aikakauslehdessä tai muu vastaava*



TTA-minimikuvailu

**Aineisto**  
*Sellaisenaan hyödynnettävissä oleva tietokokonaisuus*

**Metatietotietue**  
*Yhtä aineistoa kuvaileva, samasta lähteestä tuleva metatietokokonaisuus*



**Tiedosto**  
*Jonkinkokoinen yksikkö raakadataa, teoriassa bittijono*

**PAS-paketti**  
*kokonaisuus, jossa PAS-järjestelmät vastaanottavat, luovuttavat tai siirtävät tietoja*

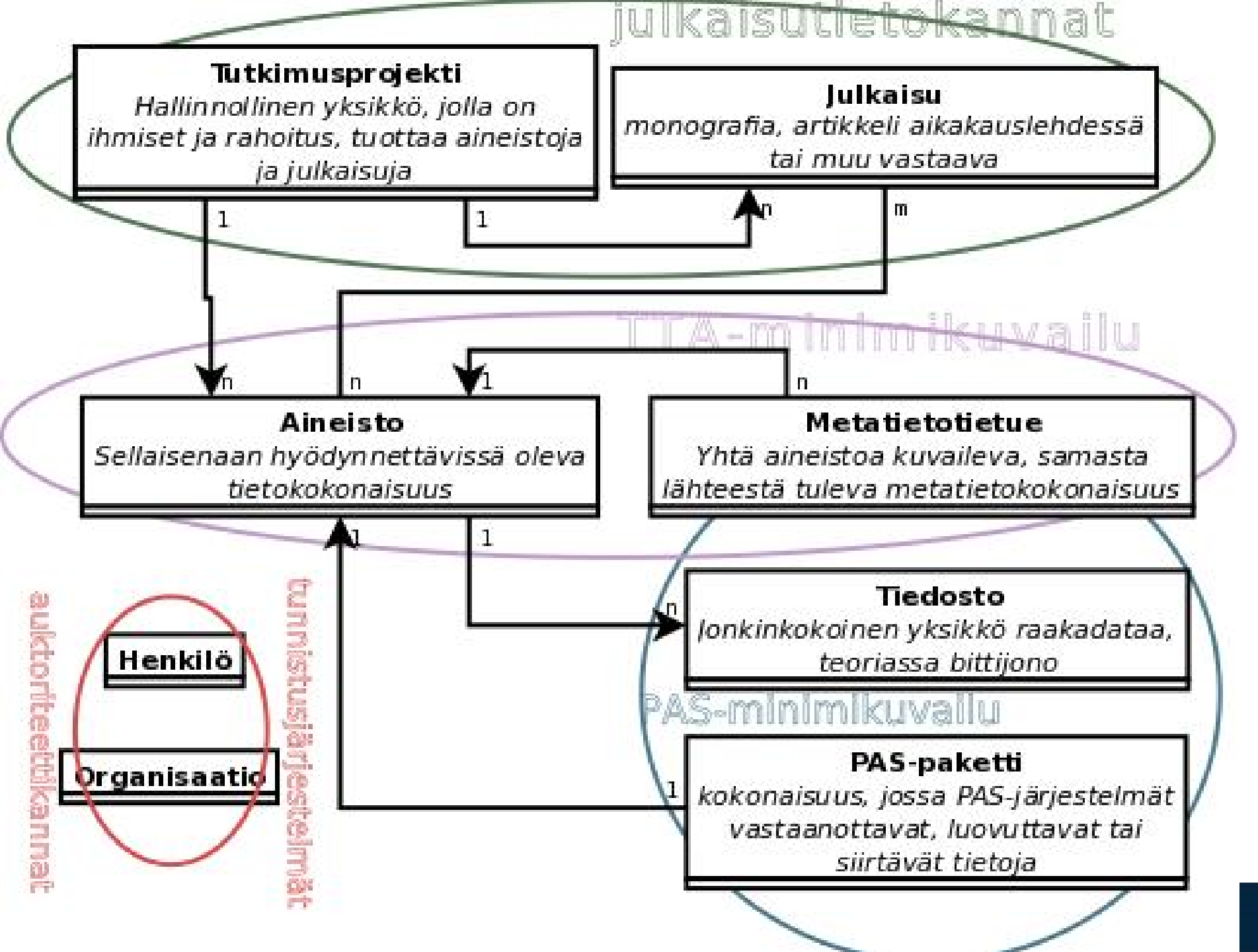
PAS-minimikuvailu

auktoriteettikannat

**Henkilö**

**Organisaatio**

tunnistusjärjestelmät



Julkaisutietokannat

**Tutkimusprojekti**  
*Hallinnollinen yksikkö, jolla on ihmiset ja rahoitus, tuottaa aineistoja ja julkaisuja*

**Julkaisu**  
*monografia, artikkeli aikakauslehdessä tai muu vastaava*

1

1

n

m

TTA-minimikuvailu

**Aineisto**  
*Sellaisenaan hyödynnettävissä oleva tietokokonaisuus*

**Metatietotietue**  
*Yhtä aineistoa kuvaileva, samasta lähteestä tuleva metatietokokonaisuus*

n

n

1

n

1

1

auktoriteettikannat



tunnistusjärjestelmät

**Tiedosto**  
*Jonkin kokoinen yksikkö raakadataa, teoriassa bittijono*

PAS-minimikuvailu

**PAS-paketti**  
*kokonaisuus, jossa PAS-järjestelmät vastaanottavat, luovuttavat tai siirtävät tietoja*

1

n

# TTA-minimimetatiedot

nro	elementin nimi	määrä	Tuottaminen
1.	Metatiedon tunniste	1	t/k
2.	Metatiedon muokkaamisaika	1	t/k
3.	Aineiston tunniste	1*	t/k+y
4.	Muokkaamisaika	1	t/y
5.	Jakelijan yhteystieto	1*	t/k+y
6.	Aineiston nimi	1	y(t)
7.	Käyttöehdot	1*	y
8.	Aihe	1*	y(t)
9.	Toimija	1*	t/k+y
10.	Aineiston kieli	0*	y

määrä:

1 = vain kerran

1\* = useita kertoja

0\* = useita kertoja, voi puuttua

tuottaminen:

t = tallennuspalvelu automaattisesti

k = aineistokatalogi automaattisesti

y = aineiston julkaisija käsin

# Minimimetatietomallin mukaisten tietojen syöttö

- TTA-palveluihin voidaan viedä metatietoja monta kautta
  - Ida-palvelussa tiedoston metatiedot
  - Kata-aineistokatalogi
  - Kotiorganisaation aineistopalvelu
- Näistä on toteutettu Ida-palvelu
- Metatietoja voi muokata käsin tai viedä automaattisesti



# Ida kuvailuvaatimus

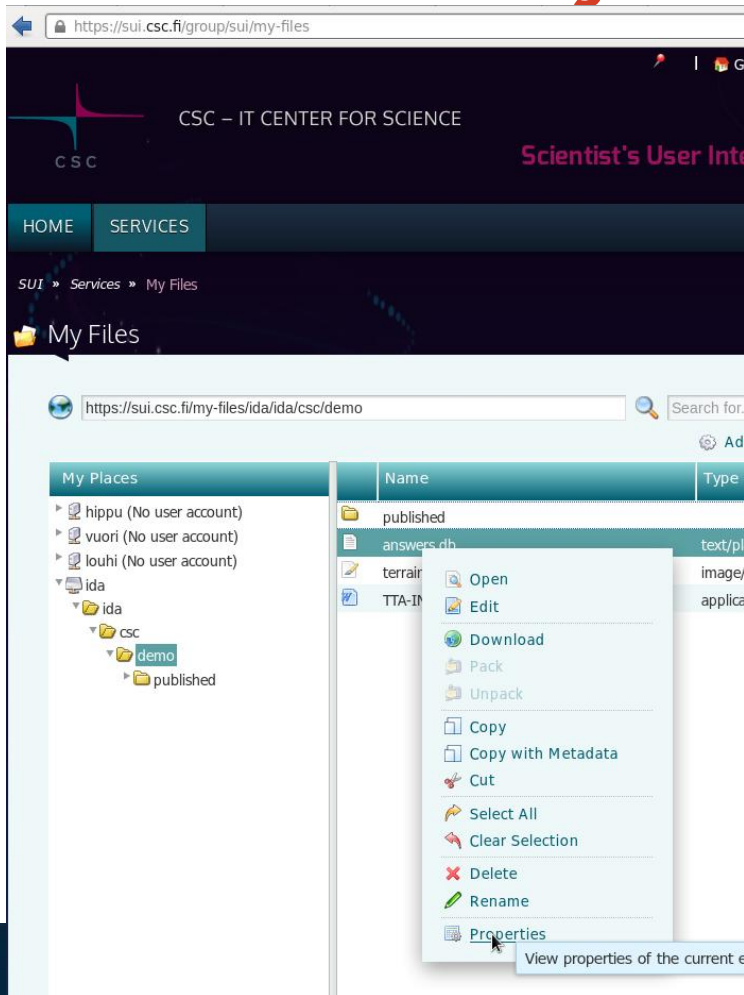
Nimi	Ida esitäyttää
Metadata.identifier	X
Metadata.modified	X
Identifier.version	X
Identifier.series	X
Modified	X
Contact.type	X
Contact.name	X
Contact.email	X
Contact.phone	X
Project.name	X
Project.funder	X
Project.funder_grant_number	X

Nimi	Ida esitäyttää
Title	
Title.lang	
Rights.category	
Rights.declaration	
Subject	
Creator	
Owner	
Language	

Nämä vaativat käyttäjältä erillistä kuvailutyötä

<http://www.csc.fi/sivut/ida/ohjeet/kayttoohjeet/Kuvailuvaatimus>

# Metatietojen syöttö SUI:ssa



https://sui.csc.fi/group/sui/my-files

CSC – IT CENTER FOR SCIENCE  
Scientist's User Interface

HOME SERVICES

SUI » Services » My Files

My Files

https://sui.csc.fi/my-files/ida/ida/csc/demo

My Places

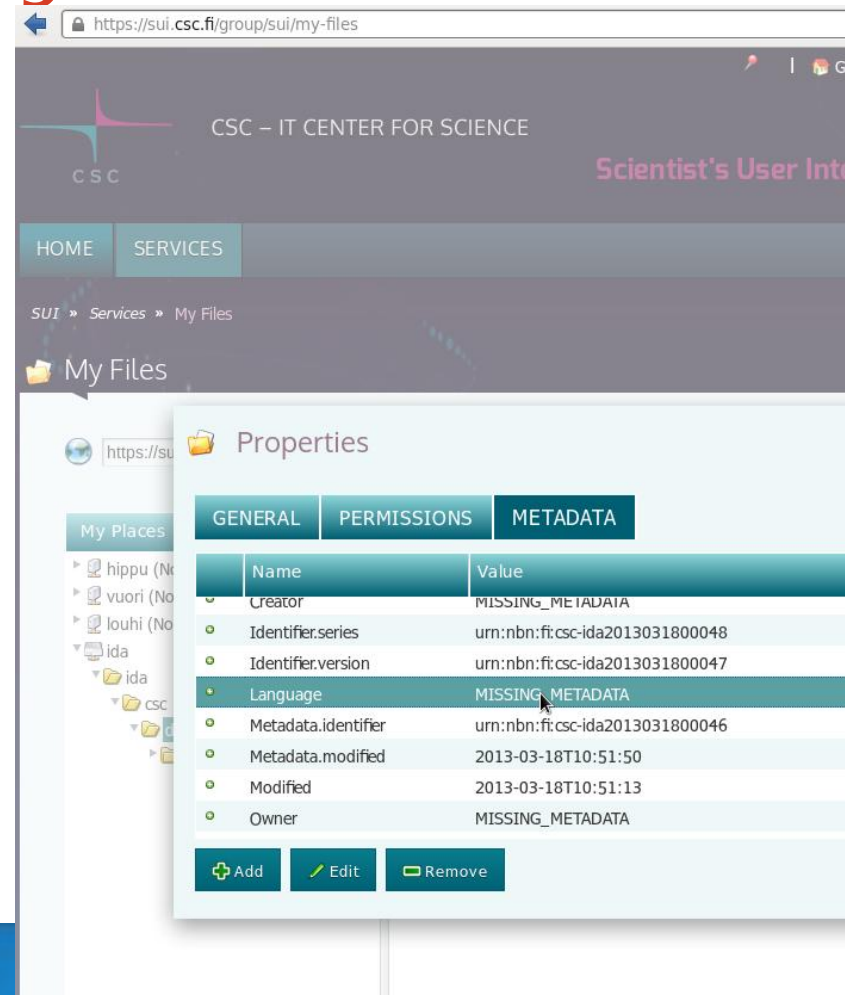
- hippu (No user account)
- vuori (No user account)
- louhi (No user account)
- ida
  - ida
    - csc
      - demo
        - published

Name	Type
published	Folder
answers.db	text/plain
terrain	image/png
TTA-IDA	application

Context menu options:

- Open
- Edit
- Download
- Pack
- Unpack
- Copy
- Copy with Metadata
- Cut
- Select All
- Clear Selection
- Delete
- Rename
- Properties

View properties of the current element



https://sui.csc.fi/group/sui/my-files

CSC – IT CENTER FOR SCIENCE  
Scientist's User Interface

HOME SERVICES

SUI » Services » My Files

My Files

Properties

My Places

- hippu (No user account)
- vuori (No user account)
- louhi (No user account)
- ida
  - ida
    - csc
      - demo
        - published

GENERAL		PERMISSIONS	METADATA
Name	Value		
Creator	MISSING_METADATA		
Identifier.series	urn:nbn:fi:csc-ida2013031800048		
Identifier.version	urn:nbn:fi:csc-ida2013031800047		
Language	MISSING_METADATA		
Metadata.identifier	urn:nbn:fi:csc-ida2013031800046		
Metadata.modified	2013-03-18T10:51:50		
Modified	2013-03-18T10:51:13		
Owner	MISSING_METADATA		

Buttons: Add, Edit, Remove

# Aineistojen käyttöpolitiikka

- Aineiston käyttöehdoista päättää aineiston omistaja
- Omistaja on lähtökohtaisesti tekijä, mutta voidaan sopia tekijänoikeuksien siirtymisestä
  - Esimerkiksi työsopimuksessa kannattaa määrittää aineistojen siirtyminen kotiorganisaatiolle
- Sopimukset voivat myös asettaa reunaehdot aineiston lisenssille
  - Esimerkiksi rahoitussopimuksessa voi edellyttää aineiston vapaata saatavuutta

## Toimijat: tekijä ja omistaja

- Henkilöihin viitataan ORCID- tai ISNI-tunnisteella
- Organisaatioihin Y-tunnuksella
- Tutkija pystyy perustamaan itselleen ORCID-tunnisteen osoitteessa <https://orcid.org/register>

# Käyttöehdot ja lisenssit

- Vaikuttavat olennaisesti aineiston uudelleenkäyttökynnykseen
- Avoin lähestymistapa helpottaa
  - Kaikille yhteinen, tunnettu lisenssi
  - Avoin lisenssi, esim. CC-BY tai CC-0
  - Avoin saatavuus, esim. www:ssä
  - Myös tietosuojakysymykset vaikuttavat

# Kaksi käyttöehtoesimerkkiä

- Avoin aineisto
  - Rights.category = LICENSED
  - Rights.declaration = <http://creativecommons.org/licenses/by/3.0/>
- Luovutussopimuksella oleva aineisto
  - Rights.category = CONTRACTUAL
  - Rights.declaration = <http://esimerkki.fi/luovutussopimus.pdf>

# Kieli

- Tarvitaan seuraaviin metatietoihin:
  - Aineiston kieli (Language)
  - Aineiston nimen kieli (Title.lang)
  - Vapaan aiheosan kieli (Subject)
- Ilmaistaan kolmikirjaimisella koodilla (ISO-639-3)
  - fin, eng, sve

# Aihesanoitus

- Auttaa toisiinsa liittyvien aineistojen tunnistamisessa
- Aiheita mm. tutkimuskohde, tutkitut ominaisuudet, menetelmät, aineiston sisältötyyppi, tutkimuspaikka ja -aika
- Aihesanat mieluiten yhteisistä sanastoista
  - Monikielisyys
  - Käsitteiden tunnistaminen samoiksi / eriksi



# TTA:ssa suositetut sanastot

Suosituimmuusjärjestyksessä:

- YSO- ja YSA-asiasanastot
  - <http://onki.fi/en/browser/overview/yso>
  - esim. sademäärä <http://www.yso.fi/onto/yso/p7001>
- Tieteenalakohtainen URL-tunnisteinen sanasto
  - esim. axonogenesis involved in innervation
  - [http://amigo.geneontology.org/cgi-bin/amigo/term\\_details?term=GO:0060385](http://amigo.geneontology.org/cgi-bin/amigo/term_details?term=GO:0060385)
- Muu tieteenalakohtainen tai erityissanasto
- Vapaa asiasana
  - esim. sve: bäckrödningar

# TTA-minimimeta vs. Tietoarkiston DDI -Codebook-kuvailut

no	TTA-minimimalli	Tietoarkiston DDI-Codebook -kuvailu
1	Metatiedon tunniste	1.1.1.5 IDNo (@agency)
2	Metatiedon muokkaamisaika	1.1.3.3 prodDate (@date) 1.1.6.1 version (@date)
3	Aineiston tunniste	2.1.1.5 IDNo (@agency)
4	Muokkaamisaika	2.1.6.1 version (@date)
5	Jakelijan yhteystieto	2.1.4.1 distrbtr (@URI)
6	Aineiston nimi	2.1.1.1 titl
7	Käyttöehdot	2.4.2.3 restrctn
8	Aihe	2.2.1.1 keyword (@vocab, @URI) 2.2.1.2 topcClas (@vocab, @URI)
9	Toimija	tekijä: 2.1.2.1 AuthEnty (@affiliation) muu tekijä: 2.1.2.2 othId (@affiliation) tuottaja: 2.1.3.1 producer (@affiliation, @role) jakelija/julkaisija: 2.1.4.1 distrbtr (@URI) omistaja: 2.1.3.1 producer (@affiliation, @role) rahoittaja: 2.1.3.1 producer (@affiliation, @role)
10	Aineiston kieli	-

# Metatietoryhmältä tulossa...

- Laajennettu malli
- Sanastosuosituksia
- Ohjeistusta



Kiitos!

Kysymyksiä?