

PÖYTÄKIRJA(LUONNOS)

TUHA-tietomallityöryhmän 1. kokous

Aika: Torstai 28.8.2014 klo 10–12

Paikka: CSC, Keilaranta 14, Espoo (kh. Diskurssi, 1. Krs)

Etäyhteys: connect.funet.fi/rakettitutki/

Osallistujat:

Leena Huiku, Tampereen teknillinen yliopisto (puheenjohtaja)
Miika Alonen, CSC
Wilhelm Barner-Rasmussen, Svenska handelshögskolan
Jukka Haapamäki, opetus- ja kulttuuriministeriö
Markku Niemi, Helsingin yliopisto
Liisa Hallikainen, Lapin yliopisto
Katja Kiviharju, Taideyliopisto
Kati Mäki, Tampereen yliopisto (Anne Lehdon sijaisena, etäyhteydellä)
Hanna-Mari Puuska, CSC (sihteeri)
Pertti Tikkanen, Oulun yliopisto (etäyhteydellä)
Stina Westman, Aalto-yliopisto

Vierailija:

Suvi Remes, CSC

Asiat

1. Kokouksen avaus ja läsnäolijoiden esittäytyminen

Puheenjohtaja avasi kokouksen klo 10.03. Läsnäolijat esittäytyivät.

2. Kokouskäytännöistä ja työmuodoista sopiminen

2.1. Valitaan työryhmälle varapuheenjohtaja

Työryhmän varapuheenjohtajaksi valittiin Markku Niemi (HY).

2.2. Sovitaan kokouskäytännöistä (mm. varahenkilöt, kokouspaikat, etäosallistuminen)

Sovittiin, että jäsenet voivat lähettää kokoukseen varahenkilön, mikäli ovat itse estyneet osallistumaan. Kokoukset pidetään pääkaupunkiseudulla, ja niihin järjestetään aina etäyhteysmahdollisuus. Toivottiin, että kokouksia järjestettäisiin mahdollisuuksien mukaan samoille päiville kuin muut ryhmät, joissa samoja jäseniä (esim. ATT:n tietomallityöryhmä).

2.3. Sovitaan työmuodoista (mm. sähköpostilistat, työalustat, kokousmateriaalit)

Luodaan tietomallityöryhmälle sähköpostilista (tuha-tietomalli@postit.csc.fi). Kokousmateriaalit viedään työalustalle confluenceen, joka on lähtökohtaisesti julkinen ja kaikkien muokattavissa.

3. Kohti yhteentoimivuusmallia

Miika Alonen ja Suvi Remes (CSC) esittelivät korkeakoulujen tietomallin kehittämissuunnitelmia. (ks. esityskalvot). Jatkossa tavoitteena on ns. yhteentoimivuusmalli, joka sisältää kohdealueen keskeiset käsitteet ja määrittelee niiden luokat, ominaisuudet ja tietotyypit. Tavoitteena ovat yhteiset tietomäärittelyt, yhteensopivuus kansainvälisten standardien kanssa ja tietojärjestelmien yhteentoimivuus. Yhteentoimivuusmallin avulla voidaan kehittää yhteentoimivia tietomalleja eri käyttötarkoituksiin ja eri mallinustasoilla. Yhteentoimivuusmallin pohjana hyödynnetään sanastoja, joissa käsitteet on määritelty.

Kommentoitiin, että sanastotyöhön painottuminen vastaisi korkeakoulujen tarpeita. Esimerkiksi TUTKI-tietomallityöryhmässä merkittävien kehittämistarpe olivat yhteiset yleisesti käytössä olevat termit, ja haasteena oli niiden sovittaminen korkeakoulujen tietomalliin. Ryhmässä haluttiin käydä keskustelua käsitteistä ja niiden määrittely – kaaviot eivät niinkään olleet prioriteettiä. TUTKI-työryhmä tuotti paljon materiaalia, mutta niitä ei vielä systemaattisesti dokumentoitu. Materiaali tarjoaisi kuitenkin hyvän pohjan sanastotyölle.

Keskusteltiin, kuka omistaa esimerkiksi työryhmän työnä syntyvät sanastot. Kerrottiin, että FINTO on Kansalliskirjaston tarjoama julkaisupalvelu sanastoille, joita voivat julkaista organisaatiot (esim. OKM). Esimerkiksi opetuksen ja koulutuksen sanasto OKSA on OKM:n koordinoima (<https://confluence.csc.fi/pages/viewpage.action?pageId=8688189>). OKSAn piakkoin valmistuvan sanaston julkaisutapa on parhaillaan OKM:ssä pohdinnassa.

Kysyttiin, kuinka opetuksen ja tutkimuksen päällekkäiset käsitteet on tarkoitus huomioida (esim tohtori-koulutus). Lisäksi kysyttiin, mikä taho viime kädessä määrittelee käsitteen, jos sitä on määrittelemässä esimerkiksi useampi ryhmä. Kerrottiin, että FINTO-KOKO-sanaston on tarkoitus yhdistää sanastot yhteisiksi käsitteiksi. Niin ikään jo XDW:ssä oli tavoitteena, että käsitteiden päällekkäisyys eri sektoreilla otetaan huomioon. Todettiin myös, että vaikka sanastotyötä tehdään useissa eri ryhmissä, osittain samat henkilöt osallistuvat niihin. Tutkimushallinnon sanastotyötä ei tällä hetkellä kuitenkaan tietävästi tee mikään muu taho, ja tässä ryhmässä olisi paras asiantuntemus sen tekemiseen. Yhteentoimivuusmallin hallintamalli ja hyväksymisprosessi ovat vielä valmisteilla. Mainittiin, että valtionhallinnon kokonaisarkkitehtuurilinjauksen mukaan OKM:llä on vastuu kohdealueen sanastoista. OKM:n intressi on, että erityisesti tuotospuoli olisi yhteismitallinen.

Mainittiin, että tutkimusta tehdään muuallakin kuin yliopistoissa (tutkimuslaitokset, yliopistolliset sairaalat, ammattikorkeakoulut) ja todettiin, että ryhmään olisi hyvä pyytää edustusta myös näistä tahoista. Myös Suomen Akatemian edustus olisi hyvä olla mukana.

4. Työryhmän työn tavoitteet

Ehdotuksia työryhmän tehtäviksi:

- 1) Sanastotyö
 - Tunnistetaan kokonaisuudet, jotka kuuluvat tutkimushallinnon toimialaan
 - Mitä sanastoja, termejä, koodistoja, malleja löytyy? Mitkä kokonaisuudet halutaan määritellä yhdessä?
 - Työryhmä voisi toimia tutkimushallinnon sanastotyöryhmänä
 - Sanastojen julkaiseminen (esim. FINTOssa)
- 2) Yhteentoimivuusmalli
 - CERIF pohjana
 - Tarkistettava yhteentoimivuus muiden kohdealueiden (esim. OKSA-sanasto, JHS) käsitteiden kanssa
 - XDW-malli inputina
- 3) Sovellusprofiilit
 - Käytännön toteutukset
 - Mitä kokonaisuuksia löytyy, mitä halutaan siirtää tietojärjestelmien välillä
- 4) Loogiset tietomallit
 - Julkaisutiedonkeruun sanasto ja tietomalli

Työryhmä keskusteli ehdotuksista. Ehdotettiin, että työssä lähdettäisiin liikkeelle laajempien kokonaisuuksien hahmottamisesta. Lähtökohtana voisi olla tutkimusprosessin kuvaaminen sen elinkaaren kautta (esim. ideointi ja rahoitus, projekti, tuotokset, hyödyntäminen) – esim. mitä dokumentteja syntyy, mitä resursseja on käytössä, ja mitä rooleja tutkijalla on eri vaiheissa. Mainittiin, että työtä olisi hyvä linkittää viitearkkitehtuuryöryhmän työhön, jossa on myös tarkoitus hahmottaa tutkimukseen liittyviä prosesseja ja kokonaiskuvaa.

Kysyttiin millä tasolla määrittelyä halutaan tehdä (esim. eri tutkijanimikkeitä). Kommentoitiin, että esimerkiksi nimikkeet kuuluvat luokituksiin. Sen sijaan esimerkiksi "vastuullinen tutkija" olisi sentyypinen käsite, jota voitaisiin määritellä ja joka on tarpeellinen tietomallissa. Mainittiin, että silloinkin kun tehdään kansallisen tason mallia, kukin organisaatio miettii käyttötarkoituksiaan sisällään. Monia käsitteitä tarvitaan sekä sisällä että ulospäin. Sanastotyössä täytyy kuitenkin rajata, mitä käsitteitä sisällytetään. Todettiin, että lähtökohtana on käyttötarve – esim. jos tiedonkeruussa tarvitaan jotain luokitus, se voidaan ottaa mukaan.

Todettiin, että osa-alueiden työstäminen olisi hyvä tehdä pienemmissä ryhmissä kuitenkin siten, että koko ryhmä miettii ensin kokonaisuudet. Alustavasti esitettiin, että rahoitus ja talousasiat muodostaisivat yhden ja tuotokset ja tulosten hyödyntäminen toisen kokonaisuuden.

Todettiin, että ryhmällä on toistaiseksi vähän sanastotyökokemusta, joten yhteinen koulutus voisi olla tarpeen, jotta kaikilla olisi sama käsitys menetelmistä. Sovittiin alustavasti, että seuraavan kokouksen yhteydessä käydään aluksi noin tunnin pituinen johdatus terminologiseen sanastotyöhön. Sanastokoulu-

tus voisi auttaa hahmottamaan myös työmäärää (OKSA-sanastossa määritelty 700 käsitettä, tässä ehkä vähemmän).

Ensi vaiheessa käsitemäärittelyä olisi mahdollista tehdä taulukkopohjaisesti (käsitteet + niiden määritte-lyt + käsitteiden hierarkiat). Toinen vaihtoehto olisi ottaa halukkaiden käyttöön editorityökalu. Samalla alustalla tekeminen toisi synergiaetuja, kun kokonaisuuksia voisi linkittää toisiinsa. Kommentoitiin kuitenkin, että tässä vaiheessa työryhmän jäsenillä ei välttämättä ole aikaa perehtyä uusiin työkaluihin, ja alkuvaiheessa olisi parempi hyödyntää helppokäyttöistä menetelmää. Todettiin, että taulukoillakin päästään alkuun, mutta jossain vaiheessa on ehkä mahdollista tilata myös työkalukoulutusta (esim. Voc-Bench).

Kysyttiin, kuuluuko taideala mukaan sanastotyöhön. Mainittiin, että OKM on perustamassa työryhmää pohtimaan julkaisutiedonkeruun taidealan luokituksia ja määrittelyitä täydentämään JUREn taiteellisen työryhmän tekemää työtä. Kyseinen ryhmä voisi toimia taidetuotospuolen alaryhmänä tälle ryhmälle. Kommentteja muihinkin osa-alueisiin voidaan pyytää myös muilta TUHA-ryhmiltä ja kukin voi keskustella käsitteistä omissa organisaatioissaan. Esimerkiksi tutkimuksen määritelmään voidaan saada apua myös taloushallinnon puolelta.

Kysyttiin, onko joillakin ulkopuolisilla tahoja odotuksia ryhmän työlle. Todettiin, että mahdollisia tarpeita liittyy esimerkiksi julkaisutiedonkeruun uudistamiseen. Tässä vaiheessa ryhmä kuitenkin asettaa itse aikataulunsa.

Sovittiin, että seuraavaan kokoukseen CSC hahmottelee XDW:n pohjalta kokonaisuutta ja luonnostelee taulukkoa sanastotyön pohjaksi (mitä käsitteitä jo löytyy, mihin linkittyy, yhteys CERIF-käsitteisiin). Lisäksi tehdään yhteiselle työalustalle pohja kokonaisuuksista, jota ryhmän jäsenet voivat täydentää ja merkitä omat asiantuntemusalueensa. CSC selvittää koulutusmahdollisuutta seuraavaan kokoukseen.

5. Muut asiat

ATT on käynnistänyt tietomalliselvitysryhmän, joka työskentelee syksyn 2014. Ryhmän fokuksena ovat kuvailutietojen mallintamismenetelmät, eikä se tee sanastotyötä. Ryhmään kuuluu myös tämän ryhmän jäseniä. Todettiin, että tiedonvälitys puolin ja toisin on tärkeää.

6. Seuraava kokous

Seuraava kokous pidetään lokakuussa (ajankohta riippuu koulutuksen saatavuudesta).

7. Kokouksen päätös

Puheenjohtaja päätti kokouksen klo 12.