

Kokous ja sanastokoulutus 10.10.2014

TUHA-tietomallityöryhmän 2. kokous

Aika: Perjantai 10.10.2014 klo 12–13.30

Huom. Kokouksen jälleen TSK:n sanastokoulutus työryhmälle klo 14-16

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

Etäyhteys: connect.funet.fi/rakettitutki/

Kutsutut:

- Leena Huiku, Tampereen teknillinen yliopisto, puheenjohtaja
- Miika Alonen, CSC
- Jukka Haapamäki, opetus- ja kulttuuriministeriö
- Liisa Hallikainen, Lapin yliopisto
- Katja Kiviharju, Taideyliopisto
- Jenni Kokko, Jyväskylän yliopisto
- Anne Lehto, Tampereen yliopisto (etäyhteydellä)
- Hanna-Mari Puuska, CSC, sihteeri
- Tomi Rosti, Itä-Suomen yliopisto
- Pertti Tikkanen, Oulun yliopisto

Asiat

1. Kokouksen avaus ja työjärjestyksen hyväksyminen

Puheenjohtaja avasi kokouksen klo 12. Hyväksyttiin työjärjestys.

2. Edellisen kokouksen pöytäkirja

Hyväksyttiin edellisen kokouksen pöytäkirja.

3. Sanastotyön organisointi

Ehdotus organisointumiseksi:

Muodostetaan TUHA-tietomallityöryhmän sisältä kolme pienryhmää määrittelemään kuhunkin kokonaisuuteen liittyviä käsitteitä

- Tutkimustoiminnan tuotokset: Julkaisut, meriitit ja muut tuotokset
- Tutkimustoiminnan organisointi: Henkilöt, sisäiset organisaatiot (laitokset, tiedekunnat, tutkijakoulut, tohtorikoulut, tutkimusryhmät), tutkimusalueet
- Tutkimusprojektit: rahoitus, ohjaus, projektinhallinta ja dokumentit (hakemukset, suunnitelmat, projektikansio)

Näiden lisäksi työssä voidaan mahdollisuuksien mukaan hyödyntää muiden käsitteiden osalta (esim. taiteellinen toiminta, infrastruktuurit, mittarit, tutkimusdata) muita TUHA-ryhmiä sekä OKM:n asettamia työryhmiä.

Työryhmä keskusteli pienryhmien kokoonpanosta ja tehtävistä:

- Todettiin, että tehtävänä on tutkimuksen hallinnon sanaston luominen. Työssä voidaan hyödyntää mm. CASRAI-sanastoa ja linkittää sanastoa siihen.
- Todettiin, että tutkimuksen elinkaarijaottelu on tässä vaiheessa vaikeaa hahmottaa ja se kuuluu enemmän TUHA-viitearkkitehtuurityöryhmän tehtäviin. Siksi lähdetään liikkeelle nyt enemmän käsitteistä ja niiden muodostamista kokonaisuuksista.
- Kysyttiin, kuinka pitkiä käsitelisteja tulee. Todettiin, että osiot ovat hyvin erityyppisiä. Opintohallinnon OKSA-sanasto kattaa 700 käsitettä, tässä tuskin niin paljoa.
- Kysyttiin XDW:n roolista sanastotyössä. Todettiin, että olemassa olevia määrittelyjä voidaan hyödyntää, mutta ne eivät sido, vaan käsitteitä voidaan määrittellä myös eri tavoin kuin XDW:ssä. Sanastoa tulisi luoda siten, ettei se ole sidoksissa mihinkään tietojärjestelmään.
- Ehdotettiin, että kukin alaryhmä työskentelee omalla alustallaan, mutta samantyyppisellä pohjalla (esim. samat sarakkeet, joita voidaan kuitenkin tarpeen mukaan laajentaa kaikille alaryhmille, ks. Sanastotaulukko: https://docs.google.com/spreadsheets/d/1BTuJoGcj7PbOquwRaDGBb-_HK6IMsfPUalngGS9QWAO/edit#gid=1857719969). Jatkossa tavoitteena on, että hierarkkinen sanasto voitaisiin julkaista FINTOssa. Yläkäsite voi olla tällöin suoraan FINTOn käsite tai käsite omassa sanastossa.
- Täytyy siis esim. huomioida, kuinka julkaisu-käsite on määritelty YSON hierarkiassa. On hyvä määritellä käytettäväksi tarkoitetut käsitteet, mutta hierarkiaa ei ole syytä rikkoa. Erikoisontologiassa voidaan muodostaa oma käsitehierarkia. Esimerkiksi OKM:n tiedonkeruukäsikirjassa julkaisu toimii tutkimuksen tuotoksen synonyymina. Toisaalta esim. tieteellisten lehtien käyttämät termit eivät aina ole samoja kuin OKM:n luokitus.

Pohdittiin, mitä lähteitä hyödynnetään organisaatio-kokonaisuuden käsitteiden määrittelyyn. Miten tulee huomioitua kaikkien organisaatorakenteet? Tiedekunta/laitos/laitoksen osa/osasto/tieteenalayksikkö/tutkimuskeskus/erillislaitos (yliopistojen yhteinen) voivat olla osin synonyymeja, eivätkä välttämättä hierarkkisia. Kommentoitiin, että tämä on nimenomaan sanastotyössä tehtävää työtä. Käsitteitä voidaan määritellä myös vaihtoehtoisiksi termeiksi. Esim. tiedekunnalla ja laitoksella voi olla yhteinen yläkäsite "organisaatioyksikkö". Jos niillä on eri käyttötarkoitus, ne pidetään erillään. Todettiin, että olisi syytä kartoittaa, mitä yksiköitä organisaatioissa on olemassa olevia dokumentaatioita hyödyntäen. Mainittiin, että voisi katsoa myös mitä koulutuksen puolella on tehty.

Sovittiin, että

- Puheenjohtajat kokoavat ryhmänsä ja sopivat yhteydenpidosta ja aikataulusta.
- Alaryhmien kokoonpanot: 1. [tutkimuksen tuotokset](#), 2. [tutkimuksen organisointi](#), 3. [tutkimusprojektit](#)
- Jokaiselle ryhmälle luodaan oma confluence-sivusto.
- Koko tietomalliryhmä kokoontuu seuraavan kerran seminaarityyppiseen tapaamiseen helmikuussa 2015, jolloin käsitellään alaryhmien siihenastiset tuotokset

4. Muut asiat

- EuroCRIS Strategic Members Meeting 11.-12.11.2014 Amsterdamissa: <http://www.eurocris.org/Uploads/Web%20pages/homepage/MM-programmeAmsterdam.pdf>

5. Seuraava kokous

Helmikuussa 2015

6. Kokouksen päätös

Puheenjohtaja päätti kokouksen klo 13.30.

Sanastokoulutus

[Sanastotyön perusteet -koulutusmateriaali](#) (Katri Seppälä, TSK)

Terminologian peruskäsitteet: tarkoite, termi ja käsite, käsitesuhde ja käsitejärjestelmä

Tarkoite:

- Tarkastelun kohteena oleva konkreettinen, abstrakti tai kuviteltu esine, olento tai ilmiö.

Termi:

- Käsitettä kuvaava termi. Yleiskielen sana saa täsmällisen merkityksen lauseyhteyden perusteella. Termi voi olla sana, yhdyssana, sanaliitto tai lyhenne.
- Hyvä termi on "läpikuultava", eli antaa alustavan mielikuvan käsitteestä. Lisäksi johdonmukainen ja neutraali sekä toisista termeistä erottuva.
- Lyhyin on yleensä toimivin termi. Yläkäsitettä kuvaava termi on hyvä olla produktiivinen, eli soveltuu johdoksien tekemiseen, esim. Julkaisu.

Käsite:

- Tiettyä tarkoitetta tai tarkoitejoukkoa vastaava mielikuva
- Käsitteitä on yleiskäsitteitä ja yksilökäsitteitä. Esimerkiksi. tietokone ja "maailman ensimmäinen tietokone".
- Käsitteistä muodostetaan hierarkia ja hierarkisuus pitää näkyä myös määritelmässä. Jokaisella alakäsitteellä on yksi tai useampi erottava tekijä.
- Tärkeää erottaa toisistaan koostumussuhteinen hierarkia ja alakäsite hierarkia. Koostumussuhteessa yläkäsite muodostuu alakäsitteistä, ts. koostumussuhteessa moottoriajoneuvo koostuu (pyörä, moottori) ja käsitehierarkiassa auto on moottoriajoneuvon alakäsite.

Käsitejärjestelmä:

- Kuvaa suhteita ja voidaan esittää käsitekaaviona. Yksi käsitekaavio kuvaa yleensä jotain osaa koko käsitejärjestelmästä.
- Käsitejärjestelmä auttaa tiedon jäsentelyssä ja selvittää käsitteiden välisiä suhteita. Piirtäminen helpottaa käsitteiden ja termien vertailua sekä auttaa määritelmien muotoilussa

Käsiteanalyysi:

- Käsitteen analysointi perustuu piirteiden tarkasteluun. Määritellään olennaiset ja epäolennaiset piirteet.
- Peruskysymyksenä kysytään mikä jokin on? Sitten mietitään mikä on määrittelyn kannalta olennaista.

Määritelmä:

- Hyvä määritelmä kuvaa käsitteen sisällön olennaiset piirteet ja määrittää suhteet muihin käsitteisiin.
- Lisäksi yksilöi käsitteen niin että se eroaa muista käsitteistä.
- Luo ohjeet ja normit käsitteen käytölle tietyllä alalla.
- Sitoo yhteen käsitteen ja termin jotta käsitteeseen voidaan viitata yhteisesti sovitulla vakiintuneella termillä ja termin merkitys selkiytyy
- Auttaa selkiyttämään viestintää erikoisalan sisällä, sekä asiantuntijoiden ja maallikoiden välillä
- Kerrotaan mikä on lähin yläkäsite
- Olennaiset ja erottavat käsitepiirteet
- Ei toissijaisia tai (redundanteja) perittyjä käsitepiirteitä

Määritelmän perusrunko:

- termi (on) jotain joka on jonkun alakäsite/vieruskäsite ja jolla on (erottavia piirteitä)

Määritelmälle ei tule pistettä, ja se aloitetaan yleensä pienellä. Huomautus alkaa isolla ja huomautuksessa käytetään pistettä. HUOM: Meillä kuvataan määritelmä ja huomautus eri sarakkeisiin. Määrittelyssä voidaan noudattaa annettua ohjetta.

Huomautus:

- Sisältää tärkeitä käsitteen piirteitä jotka eivät kuitenkaan ole epäolennaisia, esimerkiksi viittaus säädökseen. Huomautus määritellään ajattelun käyttötarpeen mukaan. Voi sisältää esimerkkejä.

Sanastotyö vs. Ontologia:

- Terminologinen sanastotyö kuvaa erikoisalan käsitteistöä. Määritellään systemaattisesti käsiteanalyysimenetelmällä.
- Ontologia määrittelee formaalin tavan muodostaa koneluettava sanasto. Jokaiselle käsitteelle annetaan koneymärteinen tunnus (URI).
- Sanastotyö toimii työkaluna ontologian muodostamiselle

Kysymyksiä:

Miten kahden eri todellisuuden väliset erot otetaan huomioon käsittemäärittelyssä? Esim. Yliopisto ja yliopisto yliopistolaissa?

- Esimerkiksi säädösten määritelmiä voidaan käyttää lähteenä, mutta yleensä eri lähteissä olevat määritelmät ottaa huomioon epäolennaisia käsitepiirteitä. Yhtenä ratkaisuna voidaan määritellä kaksi eri käsitettä, yliopisto ja sen alakäsite virallinen yliopisto tjms.

Miten erotellaan yksilökäsite ja yleiskäsite, esim. useampi ilmiö julkaisusta? Esim. monta jazz konserttia, yksi vai useampi julkaisu?

- Asia päätetään ilmiötasolla. Jokainen tapahtuma voi itsessään olla yksilökäsite mutta myös tyyppiä julkaisu.

Miten julkaisu suhteutetaan asiakirja käsitteeseen?

- (Tieteellisessä) julkaisussa laaditaan tuotos jolla pyritään tiedon levittämiseen.
- Asiakirja käsittelee ja dokumentoi jotain asiaa, lisäksi se voidaan myös julkaista, jolloin siitä tulee julkinen asiakirja.
- Tieteellisellä julkaisulla pitää olla myös auktorisoitu julkaisija
- Miten määritellään ns. itsejulkaisu tai tieteellinen itsejulkaisu. Lisäys: esim. Grigori Perelmanin matemaattinen todistus verkkosivuilla?

Mikä pitäisi olla suhde olemassaolevien sanastojen ja käytössä olevan termistön suhde uuden sanaston määrittelyyn?

- Olemassaolevat termit pitää ottaa huomioon, mutta kuitenkin niin että yleinen määritelmä ei rajoita yleistä käsitettä.
- Pitää ottaa huomioon tietyn alan vakiintunut sanasto, käsitteelle voidaan määritellä "scopenote".
- Määriteltävä pääkäsite ja alakäsitteet joissa on rajattu alakäsitteiden käyttötarkoitus, esimerkiksi. julkaisujen tilastointi.

Miten tärkeää on määritellä assosiaatiosuhteet (käsitekaavioon) ?

- Pääsääntönä funktiosuhteet näkyy määritelmässä, jos on määritelmässä, niin piirretään myös kaavioon. Joskus voi piirtää funktiosuhteen katkoviivalla, jos suhde ei näy määritelmässä.