

1. Yhteisen tahtotilan luominen ja tunnisteiden käytön edistäminen

WiP

1. Tietoisuuden lisääminen

- a. Verkkosivut (PID verkoston wiki)
- b. Yhteistyöverkostot ja niiden kanavat (esim. uutiset, seminaarit)
 - i. Digime
 - ii. Datatukiverkosto
 - iii. Kuntaliitto
 - iv. TiHa
 - v. Digitaalisen kulttuuriperinnön pyöreä pöytä
 - vi. Tutkimustietovaranto
 - vii. Yhteentoimiva Suomi
 - viii. [Tutkimusinfrastruktuuriyhteistyö -työryhmä](#)
 - ix. Avoimen tieteen koordinaatio
 - x. (Tiede)kustantajat, media
 - xi. DSpace-yhteisö?
 - xii. Tutkimusta tukevat organisaatiot (UNIFI, Arene, TulanetLa jne)
 - xiii. Datanhallinnan ammattilaisten verkostot
 - xiv. ...
- c. Oma tilaisuus (tiekartan julkaisutilaisuus)

- Wikidatasta olisi hyvä lisätä webinaarien kautta tietoutta, koska ovat linked data ja semantiikan osalta edelläkävijöitä
- Myös linked data-tietoutta olisi syytä lisätä esim. webinaarien tai uutisten kautta
- Koodistoista ja PID:eistä on hyvä viestiä lisää, esim. tilastokeskuksissa tehdään kansallisia luokituksia OPH-palvelussa ja yhteentoimivuus alustassa (DVV- Digi- ja väestötieto) annetaan omat URI:t samoille objekteille. Tästä olisi syytä sopia yhteisiä ja hyvin toimivia käytänteitä.
- URI:t - Mitä kielten koodistosta käytetään? Esimerkki: Nyt eri palvelut puhuvat kielistä eri koodistoilla. Mäppäyksiä koodistojen välillä olisi suotavaa. Eli koodilla tulisi olla PID:ejä myös jotta asia saataisiin ratkaistua. Tästä olisi myös hyvä lisätä tietoisuutta.
- Halutaanko pysytellä PID-keskusteluissa tutkimuksen piiriin? DOI:t ovat marginaalisia jos puhutaan laajemmasta kontekstistä kuin vain tutkimuksesta.
- Esimerkkinä myös CASRAI joka hajosi ja meni offline-tilaan, jolloin käsitelmäärityksiä ei ollut enää saatavilla ollenkaan. Eli jos noilla käsitteillä olisi ollut PIDit, niin olisi ehkä mahdollista uudelleen ohjata ne sinne missä määrittelyt nyt asuvat. Mutta jos ei oo niin homma hajoaa täysin. EuroCRIS ottaa nyt kopin katalogi elementeistä, mutta koska ei ole PID:ejä käsitteille annettu, niitä ei myöskään pystytä siirtämään. URN olisi ollut tässä pelastus. Linkki uutiseen: <https://eurocris.org/news/casrai-domain-handover>
- Tietokantojen pysyvyyden varmistaminen ja miten niihin viitataan kun näiden sisältö vaihtuu jatkuvasti?

2. Osaamisen lisääminen

1. Koostetaan tähän dokumenttiin lähdeluetteloa, m.m. standardit
 - a. URN, DOI, Handle, OID, RaID, GRID, ORCID (Handle), ISNI (organisaatiot että toimijat, CrossRef Funder Registry, ARK, ISBN, Y-tunnus, Q-tunniste (Wikidata))
 - b. Kattava lista PID:eistä: <https://koodistot.suomi.fi/codescheme;registryCode=research;schemeCode=PID>
 2. Oma tilaisuus (PID verkosto)
 3. Koulutusmateriaalit
 4. RDA, FAIR-ryhmät, EOSC (PID TF), Knowledge Exchange, FDO forum, PID Alliance yms. seuraaminen ja raportointi
 5. Standardien ja teknologioiden kehityksen seuraaminen ja soveltamisesta keskustelu
- Linked data-kyvykyys olennaista kehittää laajemmin.
 - Metadata-työkalu olisi hyvä kehittää, jotta metadatojen hallinta olisi helpompaa
 - Metadata scheme registry = paikka johon voi rekisteröidä tietomalleja, työkalu joka tullaan rakentamaan EOSC-CORE-projektissa. Tähän rakentavat mäppäysväline, jolla voi kaksi skeemaa määrittää toisiinsa. Data muunnin olisi ideaalinen tulevaisuuden kehityskohde, joka auttaisi mappauksia menevään yksi yhteen. Data Type Registry:ssä pitäisi pysytellä muuttamaan datakoodistot yhteismitallisiksi.

3. Luottamuksen rakentaminen

- a. PID-politiikat organisaatioissa
- b. Templaattit ja esimerkkidokumentit, myös koulutusmateriaalit
- c. Sopimukset ja tekninen varmistus
 - i. ORCID-konsortio CSC:llä. Walter Rydman toimii meillä tässä (ORCID-konsortion sihteerinä)
 - ii. DataCite Finland Consortia
 - iii. ePIC-konsortion perustajajäsen (Pekka Järveläinen tekninen yhteyshenkilö)
 - iv. Mikropalvelu - Menee noin kuun vaihteessa tuotantoon
- d. Yhteistyön rakentaminen
 - i. Kuinka rakennetaan toimintatapa, jossa pystytään toimimaan itseorganisoiuvasti yhteisiä suuntaviivoja seuraten. Millaisia kaikkien käytössä olevia tietoteknisiä palveluita (=järjestelmiä) tarvitaan.
 - ii. Keskitetyt voisi edistää: "JulkiCT:ssä, jonka pitäisi vastata koko julkisen hallinnon tiedonhallinnasta. Viime aikoina ovat tosin keskittyneet enemmän tiedonhallintakartan lanseeraamiseen, mutta koska heidän pitäisi myös edistää kansallista kokonaisarkkitehtuuria, niin sinnehan tuon asian käsittely kuuluisi."- DVV/VM

4. Koordinaation järjestäminen

1. PID-verkosto
2. TiHan kylkeen?
3. Digivirasto?

4. PRH
 5. Konsortiot: DataCite, ORCID, CrossRef?
 6. Kansalliskirjasto, CSC
 7. MML, OM
- PID:ien koordinaatiota varten voitaisiin pystyttää esim. Suomen linked data-yhteisöä, koska yhteisen tahtotilan luominen on hyvin tärkeää. Aallossa on perustettu 'Linked Data Finland'

KAPPALE TIETOISUUDEN LISÄÄMISESTÄ

Ennen kuin voidaan keskustella tietoisuuden lisäämisestä, meidän tulisi miettiä miten rajoittaa keskustelua - halutaanko pysytellä PID-keskusteluissa tutkimuksen piiriin, vai laajennetaanko sitä kattaen myös tutkimuksen ulkopuolelle? DOI:t ovat marginaalisia jos puhutaan laajemmasta kontekstistä kuin vain tutkimuksesta. Yleisenä huomiona myös että jokaisen ei aina tarvitse tietää kaiken PID:eistä, koska prosesseja tulisi mahdollisimman pitkälle automatisoida.

Wikidatasta olisi hyvä lisätä webinaarien kautta tietoutta, koska ovat linked data ja semantiikan osalta edelläkävijöitä. Myös linked data-tietoutta sekä koodistojen PID-käytöstä olisi syytä lisätä esim. webinaarien tai uutisten kautta. Linked datasta koska xxx ja koodistoista koska esim. tilastokeskuksissa tehdään kansallisia luokituksia OPH-palvelussa ja yhteentoimivuus alustassa (DVV- Digi- ja väestötieto) annetaan omat URI:t samoille objekteille. Myös URI:en käyttö kielten koodistoissa on puutteita, esim. tällä hetkellä eri palvelut puhuvat kielistä eri koodistoilla. Mäppäyksiä koodistojen välillä olisi siis suotavaa ja eri koodistoilla tulisi olla PID:ejä myös jotta asia saataisiin ratkaistua. Näistä olisi syytä sopia yhteisiä ja hyvin toimivia käytänteitä. Esimerkkinä myös CASRAI joka hajosi ja meni offline-tilaan, jolloin käsitelmäryhtyksiä ei ollut enää saatavilla ollenkaan. Eli jos noilla käsitteillä olisi ollut PIDit, niin olisi ehkä mahdollista uudelleen ohjata ne sinne missä määrittelyt nyt asuvat. Mutta jos ei oo niin homma hajoaa täysin. EuroCRIS ottaa nyt kopin katalogi elementeistä, mutta koska ei ole PID:ejä käsitteille annettu, niitä ei myöskään pystytä siirtämään. URN olisi ollut tässä pelastus. Linkki uutiseen: <https://eurocris.org/news/casrai-domain-handover>

KAPPALE OSAAMISEN LISÄÄMISESTÄ

Eri PID-tyyppien listaaminen ja näiden rajoitus on haastavaa, koska niitä on olemassa hyvin monta. Lista kuitenkin lyhenee huomattavasti jos rajoitetaan siihen että ne resoluvoituu, jolloin niitä on vain neljä; (ARC, DOI, Handle ja URN). Mutta toisaalta, URN sisältää sekä DOI:n että Handle, jonka takia erottuu hieman muista PID:eistä. Jos ei tällaista rajoitusta tehdä, PID-tyyppejä on edellä mainittujen lisäksi esim. OID, RaiD, GRID, ORCID (Handle), ISNI (organisaatiot että toimijat), CrossRef Funder Registry, ARK, ISBN, Y-tunnus, Q-tunniste (Wikidata). Kattavamman listan PID:eistä löytyy täältä: <https://koodistot.fi/codescheme;registryCode=research;schemeCode=PID>. 'Elektronisten julkaisujen identifiointi'-opas kertoo miten PID:ien avulla voidaan tunnistaa sähköisiä julkaisuja. (Pysyvä tunnistus: URN:ISBN:978-951-51-7661-5).

Osaamisen lisääminen tapahtuu parhaiten esimerkkien kautta opettaminen. Eli esimerkkitapausten esille tuominen on koulutusmateriaalina erityisen tärkeää. Osaaminen ja tietoisuus PID:eistä karttuu myös kun joutuu etsimään ja hyödyntämään avointa dataa eri lähteistä, niin samalla huomaa nopeasti mihin PID:ejä tarvitaan. Kansainvälisen kentän kehitystä tulee myös tiiviisti seurata, erityisesti standardien ja teknisten ratkaisujen osalta ja sopia yhteisistä käytänteistä yhdessä yli maan rajojen.

Linked data-kyvykyys olennaista kehittää laajemmin. Metadata-työkalut olisi hyvä kehittää, jotta metadataojen hallinta olisi helpompaa. Metadata scheme registry = paikka johon voi rekisteröidä tietomalleja, työkalu joka tullaan rakentamaan CORE4EOSC-projektissa. Tähän rakentavat mappausväline, jolla voi kaksi skeemaa mapata toisiinsa. Data-muunnin olisi ideaalinen tulevaisuuden kehityskohde, joka auttaisi mappauksia menevään yksi yhteen. Data Type Registry:ssä pitäisi pysyä muuttamaan datakoodistot yhteismitallisiksi.

KAPPALE LUOTTAMUKSEN RAKENTAMISESTA

Kansalliskirjasto valmistelee parhaillaan URN-sopimusta. TSV:n sopimus CrossRefin kanssa [Journal.fi](https://journal.fi) DOI-tunnisteen käytöstä. On keskusteltu myös että Kansalliskirjastosta tulisi DOI-taho, joka voisi käsitellä niitä DOI:ta joita ei sovellu [Journal.fi](https://journal.fi):hin taikka CSC:n piiriin. ORCID-konsortio CSC:llä. Walter Rydman toimii meillä tässä (ORCID-konsortion sihteerinä), DataCite Finland Consortia, ePIC-konsortion perustajajäsen (Pekka Järveläinen tekninen yhteyshenkilö), mikropalvelu - Menee noin kuun vaihteessa tuotantoon.

ORCID-konsortioista: ORCID-konsortio ei hallinnoi niinkään tunnisteita tai niiden käyttöä. Se huolehtii suomalaisten tutkimusorganisaatioiden ORCID-jäsenyyksistä, jonka avulla organisaatiot saavat oikeuden käyttää ORCIDin kattavinta kaksisuuntaista rajapintaa. Ts. tällä hetkellä pääasiassa "hallinnollinen toimija". ORCID-tunnisteen käyttöön jne. liittyvät asiat kuuluvat PID-verkostolle. eli PID verkoston tulisi luoda ja ylläpitää myös ORCID:ja koskeva politiikka/ohjeistus siitä mitä (meta)tietoja kunkin tunnisteen alla ylläpidetään. ORCID-tunniste on yksi PID muiden joukossa ja sopii luonnollisesti PID-verkoston piiriin. ORCID-tunnisteen tapauksessa se varmasti olisi lämmin suositus.

Yhteistyön rakentaminen

Kuinka rakennetaan toimintatapa, jossa pystytään toimimaan itseorganisoituvasti yhteisiä suuntaviivoja seuraten. Millaisia kaikkien käytössä olevia tietoteknisiä palveluita (=järjestelmiä) tarvitaan. Keskitetysti voisi edistää: "JulkiCT:ssä, jonka pitäisi vastata koko julkisen hallinnon tiedonhallinnasta. Viime aikoina ovat tosin keskittyneet enemmän tiedonhallintakartan lanseeraamiseen, mutta koska heidän pitäisi myös edistää kansallista kokonaisarkkitehtuuria, niin sinnehan tuon asian käsittely kuuluisi."- DVV/VM. PID:ien jakavien ja ylläpitävien organisaatioiden yhteinen kanava olisi hyvä olla.

KAPPALE KOORDINAATION JÄRJESTÄMISESTÄ

Ensinnäkin pitäisi määritellä mitä me järjestämisellä tarkoitetaan. Onko kyseessä informaatio-, vai budjettiohjaus? Onko itseohjautuva verkosto sopiva /riittävä? Kuuluuko verkostoon myös yksityisiä toimijoita, vai pelkästään julkishallinnon puolelta?

PID:ien koordinaatiota varten voitaisiin pystyttää esim. Suomen "Linked Data"-yhteisöä, koska yhteisen tahtotilan luominen tässä olisi hyvin tärkeää. Aallossa on perustettu 'Linked Data Finland', mutta sitä ei olla lähdeyt skaalaamaan kansalliselle tasolle soveltuvaksi. Tosin voidaan ajatella että PID-verkosto voisi ajaa tätä asiaa ja "Linked Data"-aiheeseen liittyen voisi hyvin sopia verkoston avoin työryhmä.