

Datapolitiikasta käytäntöön: Tunnisteiden hallinnasta

Fakta Julkaistu 24.03.2016 klo 15:27 Kirjoittaja Jessica Parland-von Essen

Hyvien tutkijapalveluiden toteuttaminen ja tehokas palveluarkkitehtuuri vaativat organisaatioilta [tunnisteiden](#) hallintaa. Myös kansallisten arkkitehtuurien toimeenpano edellyttää organisaatioiden kiinnittävän huomiota semanttiseen yhteentoimivuuteen ja tiedon linkittämisen mahdollistamiseen. Tämä vaatii myös yksilöivien tunnisteiden ja niiden pysyvyyden pohdintaa.

Tarjolla on sekä organisaation sisäiseen käyttöön toimivia ratkaisuja että ulkopuolisia PID-palveluita. Esimerkiksi [paikkatiedolle](#), [julkaisuille](#), [tutkijoille](#) sekä [opetus- ja koulutuslalle](#) on olemassa eritasoista ohjeistusta ja palveluita yksilöivien tunnisteiden ja niiden käytön osalta. Asiasanoille löytyy puolestaan tunnisteita [Finto-palvelusta](#).

Mikäli käytössä on omia järjestelmiä, joissa on verkkoon näkyviä tunnisteita, on syytä huolehtia siitä, että ne ovat [EU-komission ohjeiden](#) sekä [W3C:n suositusten](#) mukaisia. Ulkoista palvelua valittaessa on puolestaan syytä tarkistaa, että järjestelmä on teknisesti luotettava, auktoritatiivinen, joustava metadatan esittämisen suhteen sekä yhteentoimiva omien ja kansallisten järjestelmien kanssa. Myös resolvoinnin tarvetta on syytä pohtia.

Erilaisia PID-järjestelmiä käytetään yhä useammin, sillä yksinkertaiset URI-tunnisteet eivät välttämättä ole riittävän pysyviä verkko-osoitteiden, -sivustojen tai organisaatioiden rakenteiden muuttuessa. URI-tunnisteet voivat kuitenkin olla pysyviä niin kauan kuin niitä hallinnoiva organisaatio omistaa kyseisen verkko-osoitteen. PID-järjestelmän tunnisteet ovat puolestaan pysyviä niin kauan kuin palvelu on olemassa, eikä esimerkiksi verkkosivun osoitteen muutos vaikuta niihin. PID-järjestelmän, kuten esim. Handlen, ylläpitäminen vaatii kuitenkin jatkuvaa teknistä ylläpitoa ja osaamista. Content negotiation -mekanismin avulla järjestelmä voi myös sopeuttaa vastauksensa riippuen kyselyn agentista siten, että esimerkiksi selaimelle vastataan html-dokumentilla ja RDF-lukijan saadessa RDF-tiedoston. Tämä mahdollistaa uusien formaattien lisäämisen järjestelmään myöhemmin.

[Ruotsin yhteentoimivuussihteeristö](#) on kehittänyt organisaatioita varten seuraavan tarkistuslistan, jonka avulla voi lähestyä tunnisteiden hallintaa:

- onko organisaatiollasi datapolitiikka?
- onko organisaatiollasi tunnistepolitiikka?
- ovatko organisaatiosi tunnisteet sisäisesti yksilöllisiä?
- ovatko organisaatiosi tunnisteet globaalisti yksilöllisiä?
- ovatko organisaatiosi tunnisteet sisäisesti pysyviä?
- ovatko organisaatiosi tunnisteet globaalisti pysyviä?
- Mikäli käytössänne on vain organisaation sisäisiä tunnisteita, ovatko ne sellaisia, että niitä voidaan käyttää suositusten mukaisesti osana ulkoista pysyvää yksilöivää tunnistetta?
- onko käytettyjen tunnisteiden luonti ja käyttö organisaatiossa hallittua?
- toimiiko linkkien päivitys organisaatiossanne?
- onko tarvetta ulkoisen resolverin käytölle?
- jos pysyvät tunnisteet ovat URI-tunnisteita, ovatko ne EU-komission suosituksen mukaisia?
[”http://{domain}/{tyyppi}/{käsite}/{viite}/”]

Mikäli käytössänne on PID-järjestelmä:

- käytetäänkö sitä vain sisäisesti vai myös ulkoisesti?
- onko järjestelmää integroitu muihin järjestelmiin?
- onko järjestelmälle olemassa hallintasuunnitelmaa?
- koulutetaanko järjestelmän käyttäjiä?
- onko järjestelmä luotettava koskien esim. tietoturvaa ja kopiointia?
- mikäli käytätte ulkopuolisen osapuolen tarjoamia tunnisteita tai resolvointia, onko niitä ylläpitävä taho auktoratiivinen ja luotettava?
- onko järjestelmä joustava ja tarjoaako se tietoa sen laatua heikentämättä?
- onko järjestelmä yhteentoimiva (kansallisten ja kansainvälisten tai avointen standardien mukainen)?

Lisää lukemista aiheesta

- [Digisam: Checklista: Beständiga identifierare](#)
- [EU-komission ohje](#)
- [Uniform Resource Identifier, Wikipedia](#)
- [URN:NBN \(National Bibliography Numbers\)](#)
- [Uniform Resource Name \(URN\)](#)
- [Persistent URL \(PURL\)](#)
- [Digital object identifier \(DOI\)](#)
- [ORCID](#) (suomalainen sivusto tulossa)
- [ePIC](#)
- [Archival Resource Key \(ARK\)](#)
- [Open URL](#)
- [Handle-järjestelmästä](#)
- [Best Current Practice RFC 7320 - URI Design and Ownership](#)

[”Implementing Persistent Identifiers: Overview of concepts, Guidelines and Recommendations”](#), Consortium of European Research Libraries European Commission on Preservation and Access, 2006

[”Persistent and unique identifiers”](#), CLARIN, 2008

Aineistoa W3C:ltä

- [Data Identification](#)
- [URI Design Principles](#)
- [Cool URISs don't change](#)
- [Cool URIs for the Semantic Web](#)