

Mikä ihmeen pysyvä tunniste?

Näkökulmat Julkaistu 13.07.2016 klo 16:31 Kirjoittaja Jessica Parland-von Essen

Jokainen meistä on varmaan törmännyt ns. kuolleisiin linkkeihin. Tasaisin väliajoin näytölle pamahtaa http-virheilmoitus “404 Not Found”. Näin käy, kun sisältöä on siirretty palvelimella vaikka vain kansioista toiseen tai se on poistettu. Syitä ja virheitä, jotka johtavat linkkien hajoamiseen voi olla monenlaisia, mutta ne eivät ole tutkijalle erityisen relevantteja. Silti kyse on tieteelle ja tutkimukselle merkittävästä asiasta.



Rikkinäiset linkit ovat ongelma. Kuva: Kristian Niemi / Flickr CC BY-NC-ND 2.0

Tutkijalle on tärkeää, että julkaisu tai aineisto löytyy ja on jatkuvasti saatavilla. Verkossa tämä on kuitenkin pitkän päälle surullisen vaikeaa. Tekniikka päivittyy jatkuvasti ja digitaalinen tieto on luonteeltaan muutenkin epävakaata. Miten voimme varmistaa aineiston löytymisen?

Ennen kaikkea on tärkeää huolehtia siitä, että aineistoa säilyttää luotettava taho, esimerkiksi tietotai julkaisuarkisto. Ratkaisuksi tarjotaan lisäksi “PID:ejä” eli pysyviä tunnisteita, jotka toimivat (http://-etuliitteen kera) myös linkkeinä. Niiden avulla voidaan esimerkiksi tiedostoja paitsi luotettavasti yksilöidä, myös huolehtia siitä, että linkit eivät turhaan katkeile. Ajatuksena on, että pysyvän tunnisteiden avulla löytää aina vähintään aineiston perustiedot. Jos aineisto ei ole avoin,

verkosta pitäisi löytyä tieto siitä kuka siihen pääsee käsiksi ja miten. Mikäli aineistoa ei enää ole, sekin pitää kertoa, sekä miksi ja milloin se on hävitetty. Joissain tapauksissa linkki johtaa suoraan yksilöityyn sisältöön, mutta se ei välttämättä ole suositeltava malli tutkimustuotosten kohdalla.

Pysyvien tunnisteiden idea on, että uudelleenohjauksesta on huolehdittu suunnitelmallisesti siten, että sisällön siirtyminen tai esimerkiksi palvelun tai organisaation nimen muuttuminen eivät riko linkkejä. Riittäväällä sopimisella, hallinnoinnilla ja suunnitelmallisella toiminnalla huolehditaan myös siitä, että tunnisteet todella ovat yksilöivä, eikä samaa tunnistetta vahingossakaan käytetä uudestaan.

Pysyviä tunnisteita on monenlaisia, joista tutuimpia ovat ehkä Kansalliskirjaston myöntämä [URN](#) ja kansainvälinen [DOI](#). Tunnisteita voidaan antaa paitsi perinteiselle verkkosisällölle myös esimerkiksi henkilölle (tutkijoille [ORCID](#)) tai käsitteelle (esim. [Finto](#)). Näin tietoja voi paremmin hallinnoida ja linkittää. Kun pysyviä tunnisteita käytetään ahkerasti, tietojärjestelmät osaavat aina linkittää oikeat tekijät julkaisuihin, aineistot tekijöihin ja niin edelleen.

Toivottavasti virheilmoituksetkin jäävät hiljalleen vähemmälle. Se on meidän kaikkien vastuulla.

Lue lisää

Juha Hakala: Yleiskatsaus pysyviin tunnisteisiin. Tietolinja 2/2010 <http://urn.fi/URN:NBN:fi-fe201012143109>