

Pysyvät tunnukset ja niiden hyödyntäminen

Arkistopalvelut uuteen nousuun 15.9.2015

Esa-Pekka Keskitalo, orcid.org/0000-0002-4411-8452

[URN:NBN:fi-fe2015091511591](https://nbn-resolving.org/urn:nbn:fi-fe2015091511591)



Sisältö

- Millaisista tunnuksista on puhe?
- Mitä hyötyä niistä on?
- URL, URN, URI
- Tunnistejärjestelmistä
- URN-tunnisteet

Pysyvä – ainutkertainen - toiminnallinen

Pysyvä

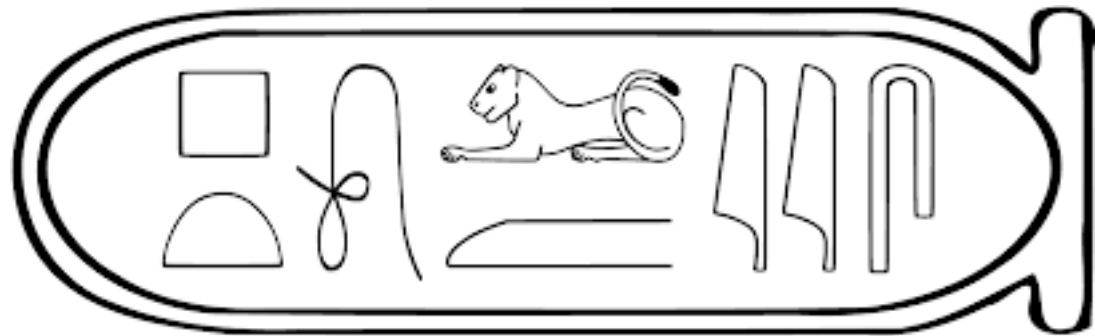
- Resurssilla on tunnus
- Resurssin tunnus ei muutu
- Resurssilla on ideaalisti yksi tunnus
- Tunnus annetaan resurssin elinkaaren alussa
- Resurssin kopioilla on sama tunnus



Pysyvä – ainutkertainen - toiminnallinen

Ainutkertainen

- Tunniste on globaalisti ainutkertainen
- Eri resursseilla on eri tunniste
- Resurssin muunnos saa eri tunnisteeseen
- Tunnuksia ei käytetä uudelleen

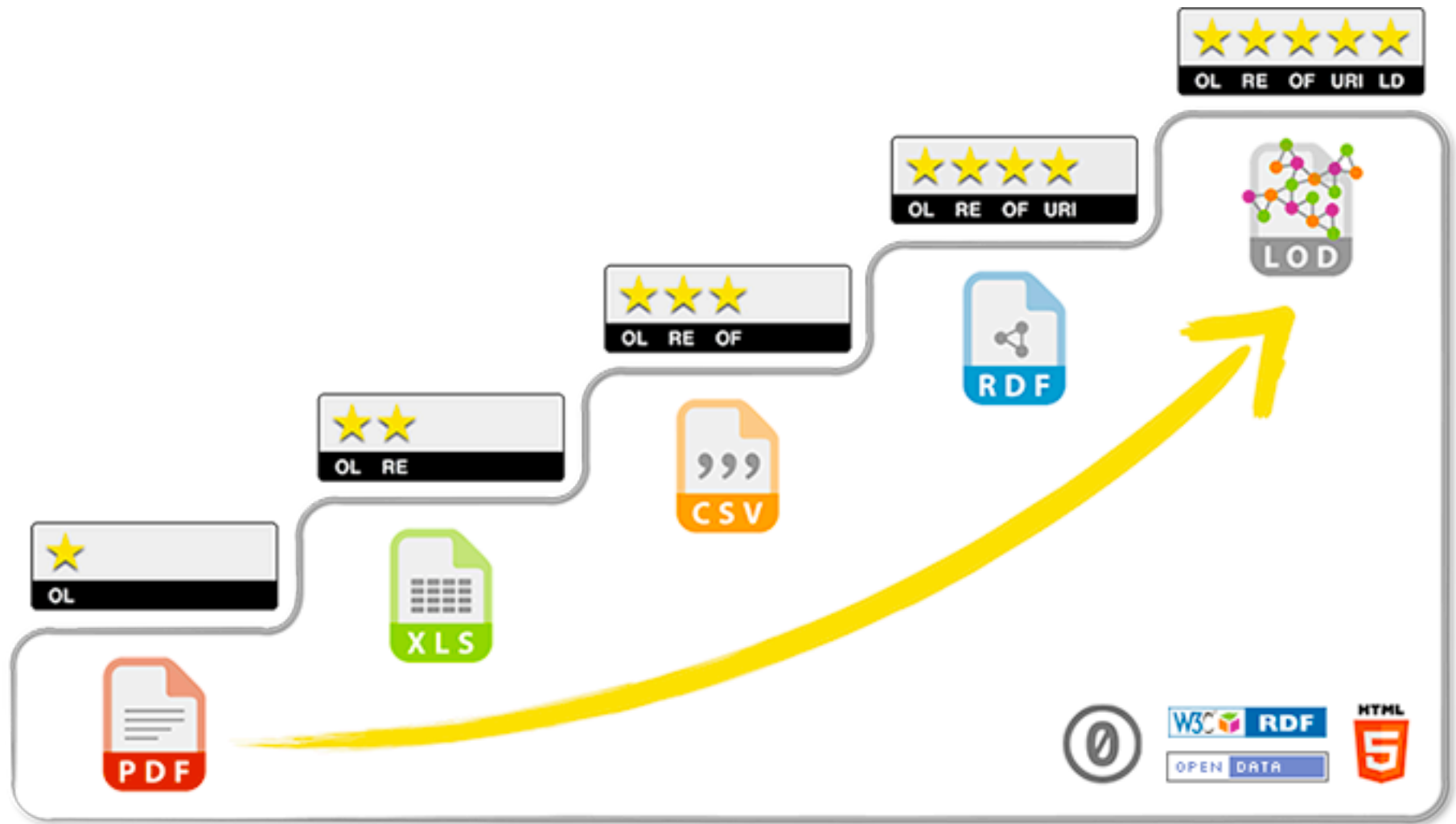


Pysyvä – ainutkertainen - toiminnallinen

Toiminnallinen

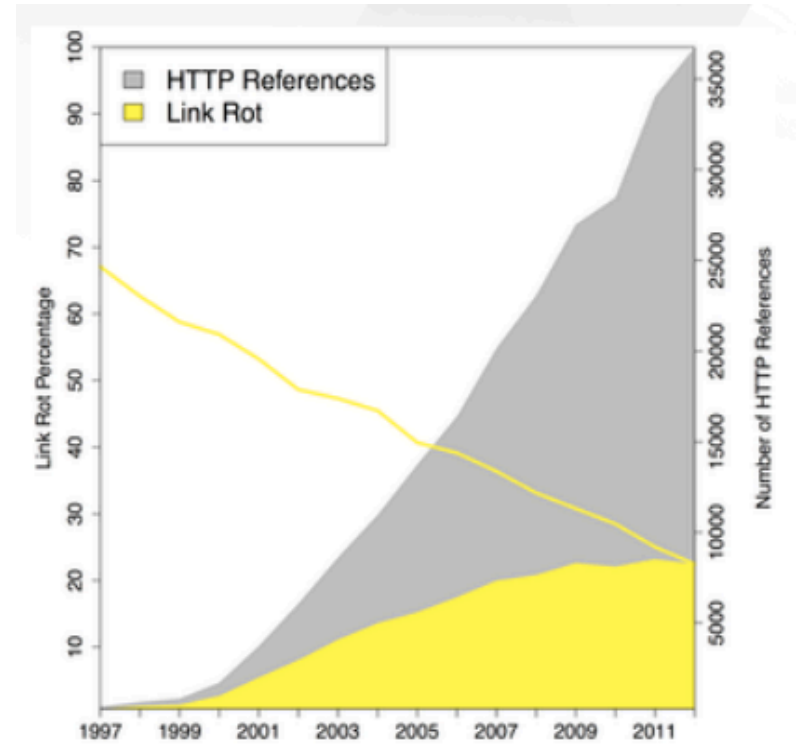
- Tunniste on linkki, joka johtaa resurssiin
- Tunniste on koneluettava
- Linkki voi vastata eri tavoilla
- Linkki toimii yli ympäristön muutosten

Kohti avointa linkitettyä dataa



Kohti toimivia viittauksia

- Vuonna 2014 tutkittiin vuonna 2012 julkaistuja tieteellisiä julkaisuja. 13% linkeistä arXiv-dokumenteissa ja 22% linkeistä Elsevierin lehdissä oli lakannut toimimasta.
- Useimmiten resurssi ei ole suorastaan kadonnut.
- Kumpi on pahempi: kadonnut sisältö vai muuttunut sisältö?
- Tunnuksilla voidaan parantaa linkkien pysyvyyttä



Scholarly Context Not Found: One in Five Articles Suffers from Reference Rot. M. Klein et al. PLOS ONE 2014.

Resoluutio

- Resurssi, jolla on tunnus ("nimi"), sijaitsee kullakin ajanhetkellä jossakin osoitteessa (URL)
- Osoitteiden täytyy voida muuttua.
- Osoitteita täytyy voida olla useita.
- Resoluutiopalvelu yhdistää nimen osoitteeseen
- Tietämyskantaa täytyy pitää yllä

URI – URL – Cool URI

- URI: <http://ww.example.com/example>
 - Uniform Resource Identifier
 - Merkkijono, joka johtaa resurssiin
 - Yllä ei ole määritelty, onko kyseessä HTML-sivu vai mitä
 - “Content negotiation” → mitä itse asiassa palautetaan
- URL on URIn erikoistapaus
- Cool URI = hyvin suunniteltu URI.
 - www.example.com/folder2/maiya/drafts/latestnews.cgi vs.
 - www.example.com/hauki-on-kala
- Hyvät URIt ovat tärkeitä mutta eivät korvaa tunnuksia.

Tunnusjärjestelmiä

- Handle, 1994
- DOI (Handlen erikoistapaus)
- Persistent URL (PURL), 1995
- Uniform Resource Name (URN), 1997
- Archival Resource Keys (ARK), 2001
- Extensible Resource Identifier (XRI), 2005

**KÄYTTÄKÄÄ *JOTAKIN*
PYSYVÄÄ TUNNUSTA!**

URN-tunnukset

- “A Uniform Resource Name (URN) is a URI that is assigned under the "urn" scheme and a particular URN namespace, with the intent that the URN will be ... a persistent location-independent resource identifier
- Muoto: URN:NID:NSS
- NID = nimiavaruuden tunnus
- NSS = nimiavaruudessa oleva merkkijono

<http://urn.fi/URN:NBN:fi-fe2015081910975>

Esimerkkejä URNeista

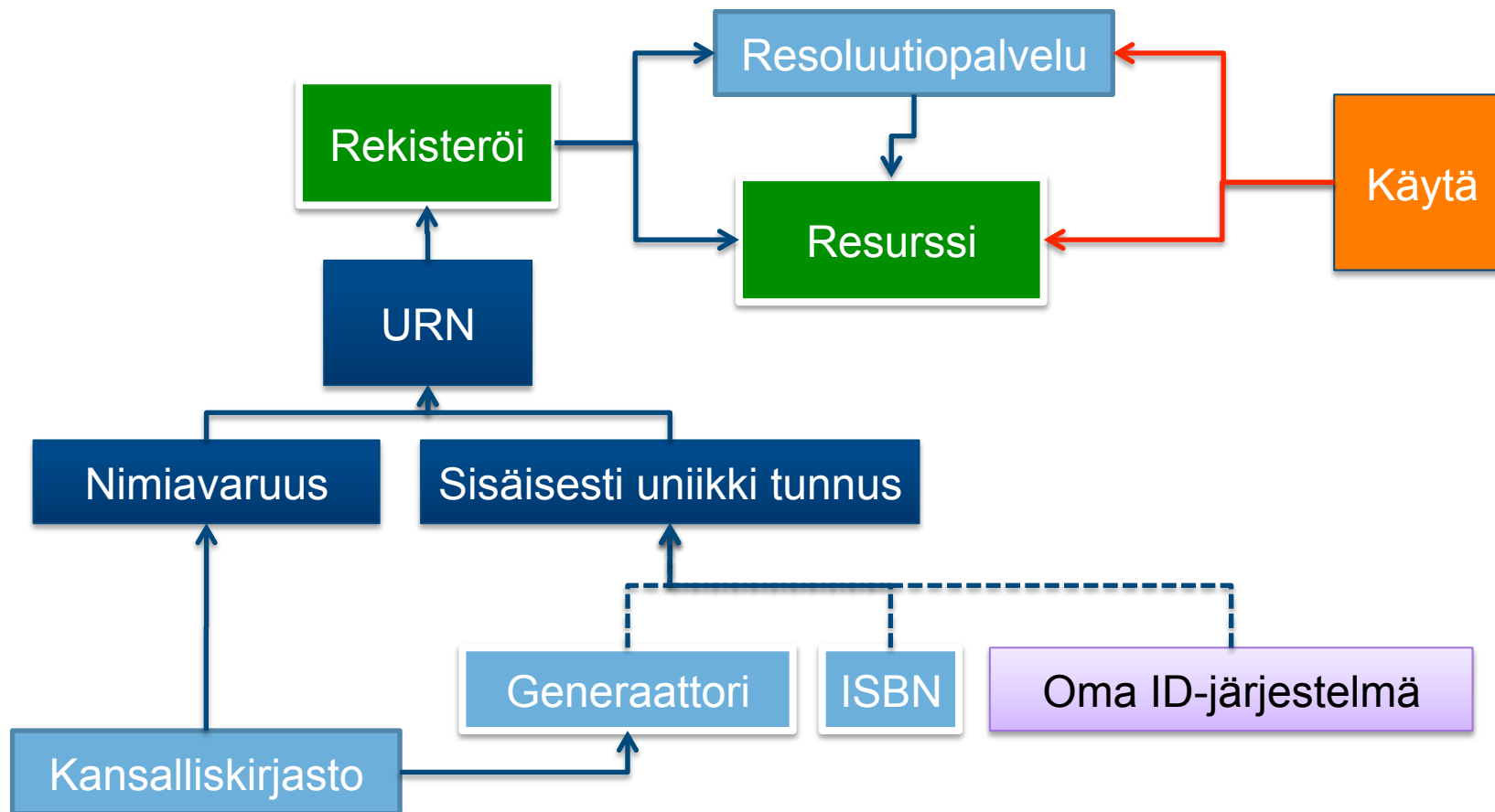
- <http://urn.fi/URN:NBN:fi:au:cn:60599A> (termi)
- <http://urn.fi/URN:ISBN:978-951-765-787-7> (monografia)
- <http://urn.fi/urn:nbn:fi:lb-201405275> (kohta korpuksessa)

Nimiavaruus ja alinimiavaruus

- Nimiavaruus: standardi tunnusjärjestelmä, joka on rekisteröity URN-nimiavaruudeksi.
- Alinimiavaruus: nimiavaruudesta lohkaistu osa, jota esim. organisaatio voi käyttää itsenäisesti.
- Alinimiavaruus voi olla ”melkein mitä tahansa”.

- **NBN** = National Bibliographic Number
- **NBN:fi** = Kansalliskirjaston hallinnoima Suomen aliavaruus
- **NBN:fi:ula** = Lapin yliopistolle annettu aliavaruus

URN-prosessi



Mitä URN vaatii käyttäjältä

- Hakee URNeja TAI hankkii URN-nimiavaruuden
- Antaa resursseille URNeja ja disseminoi niitä
- Rekisteröi URNit ja päivittää muutokset

Mitä organisaation pitäisi nyt tehdä

- Mitkä resurssit tarvitsevat tunnuksia?
 - Minkä tyyppiset tiedot; millä hienousasteella?
- Miten resurssit saavat tunnuksen
 - Prosessin suunnittelu
- Miten resurssit saavat standarditunnuksen
 - Tunnuksjärjestelmän valinta
- Miten varaudutaan tulevaisuuteen

Haasteita

- Tunnuksien antamisen automatisointi
- Tunnuksien oikea käyttö ja sen tuki
- Tunnukset termeille, toimijoille, paikkatiedolle
- Samuus – erilaisuus
 - Onko epub-versio sama vai eri kuin pdf-versio?
 - Ihmisen kannalta voi olla sama asia, koneen kannata ei.
- Resurssin verkossa säilymisen varmistaminen.
 - Esim. <http://hiberlink.org/>

Kiitos huomiosta!

esa-pekka.keskitalo@helsinki.fi
@keskitalo

