Tietoarkisto on FAIR

Avoimen tieteen kevätpäivät 2019

Tampere

14.5.2019

Hannele Keckman-Koivuniemi

tietopalvelupäällikkö

TIETOARKISTO

Esityksen sisältö

FAIR periaatteessa

FAIR käytännössä. Tietoarkisto on vastuullisesti FAIR

Tietoarkisto on jopa FAIR+

Parannettavaakin on

FAIR periaatteessa

Koneluettavuus

F

FINDABLE – LÖYDETTÄVISSÄ

F1. (Meta)datalla on yleisesti käytetty yksilöivä pysyvä tunniste

F2. Data kuvaillaan monipuolisesti

F3. Metadata sisältää kuvailemansa datan pysyvän tunnisteen

F4. Metatiedot on tallennettu paikkaan, josta ne ovat haettavissa

A

ACCESSIBLE –SAAVUTETTAVISSA

A1. (Meta)data on tunnisteidensa avulla haettavissa standardoitujen protokollien avulla

A1.1. Protokollat ovat avoimia ja yhteiskäyttöisiä

A1.2. Protokolla sallii tunnistautumisen ja valtuutuksen tarpeen mukaan

A2. Metatiedot ovat saatavissa, vaikkei data enää olisi

I

INTEROPERABLE – YHTEENTOIMIVIA

I1. Kuvailussa käytetään vakiintuneita, helposti saatavissa olevia, jaettuja ja laajasti sovellettavia tietomalleja

I2. Kuvailussa hyödynnetään FAIR-periaatteiden mukaisia yhteisiä, hyvin dokumentoituja sanastoja

I3. (Meta)data sisältää asiamukaiset viittaukset muihin lähteisiin

R

REUSABLE – UUDELLEENKÄYTETTÄVIÄ

R1. Kuvailu on kattava ja riittävän täsmällinen

R1.1. Kuvailussa on määritelty selkeästi datan omistajuus, käyttöehdot ja lisenssit

R1.2. (Meta)datan alkuperä, konteksti, prosessoinnit, versiomuutokset kuvataan tarkasti

R1.3. Kuvailut tehdään tieteenalan hyvien käytäntöjen ja standardien mukaisesti

FAIR käytännössä. Tietoarkisto on vastuullisesti FAIR

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

Tietoarkisto

hoivaa ja avaa vastuullisesti suomalaisen yhteiskunnan, ihmisten ja kulttuuristen ilmiöiden tutkimiseen kerättyjä sähköisiä aineistoja

Findable
Accessible
Interoperable
Reusable

- n. 3 000 rekisteröitynyttä asiakasta
- n. 1 450 aineistoa
- 1 954 käyttölupahakemusta (2018)
- n. 3 750 aineistoihin liittyvää julkaisuviitettä
- käyttöaste 88 %

- Toiminut vuodesta 1999, Tampereen yliopiston erillisyksikkö
- CTS-sertifioitu (CoreTrustSeal) sähköinen data-arkisto: pitkäaikaissaatavuus = säilytys+hoiva+jakelu
- Valtakunnallinen tutkimusinfrastruktuuri
- Pääasialliset asiakasryhmät: opiskelijat, tutkijat, opettajat yliopistoissa ja ammattikorkeakouluissa
- Palvelut maksuttomia
- CESSDA ERICin kansallinen palveluntuottaja

CORE TRUST SEAL

TIETOARKISTO

Tietoarkiston palveluita

- **Aineistoportaali Aila** – etsi, lataa ja tallenna tutkimusaineistoja
- **Tietopalvelu**
- **Aineistonhallinnan käsikirja**
- **Menetelmäopetuksen tietovaranto** (MOTV)
- **Penna** – Kirjoitusaineistojen keruutyökalu
- **Pohtiva** – Poliittisten ohjelmien tietovaranto
- **Tietoarkisto-lehti** ja **Tietoarkistoblogi**

Aila

haka Login

Tukea ja meriittiä tutkijoille.
Näkyvyyttä aineistoille.
Vaikuttavuutta tutkimukselle.

Koneluettavuus

F

FINDABLE – LÖYDETTÄVYYS

Annamme aineistoille aina uniikin pysyvän tunnisteen (URN).

Tietoarkistoon arkistoidut aineistot on kuvailtu yksityiskohtaisesti suomeksi ja englanniksi .

Aineistot ovat löydettävissä Tietoarkiston Ailasta ja esimerkiksi CESSDAn aineistoluettelosta sekä Finnasta ja Etsimestä.

A

ACCESSIBLE – SAAVUTETTAVUUS

Metadata on saatavilla Ailasta sekä Tietoarkiston OAI-PMH-rajapinnasta (Kuha1 ja Kuha2) (DDI 2.0, OAI Dublin Core, EAD).

(Rekisteröityneet) käyttäjät voivat ladata dataa Ailasta käyttöehtojen mukaisesti.

Aila hyödyntää HAKA-käyttäjätunnistusjärjestelmää.

I

INTEROPERABLE – YHTEENTOIMIVUUS

Aineistot kuvaillaan kansainvälisellä DDI Codebook -kuvailuformaatilla. Sisällön lisäksi myös muuttujatasolla.

Käytössä useita kansallisia ja kansainvälisiä sanastoja ja luokituksia (FINTOn YSA ja OKM-tieteenalaluokitus, ELSST-tesaurus, CESSDAn aihepiiriluokitus sekä DDI-allianssissa kehitetyt mm. otantamenetelmä- ja keruuvälinesanastot).

Metadata sisältää myös viittauksia muuhun metadataan, dataan ja julkaisuihin.

Yleisesti käytössä olevat dataformaatit.

R

REUSABLE – UUDELLEENKÄYTETTÄVYYS

Monipuoliset aineistokuvailut lisensoitu CC BY 4.0 –lisenssillä, suppeammat EAD- ja DC-aineistokuvailut CC0-lisenssillä.

Asiakkaat voivat ladata dataa arkistointisopimuksissa määriteltyjen ehtojen mukaisesti.

Käyttöehdot sisältyvät metadataan.

Tietoarkisto on jopa FAIR+

- FAIR-periaatteet keskittyvät (meta)datan saavutettavuuteen, eivät niinkään **datan pitkäaikaissäilytykseen**, mikä on edellytys sille, että toimiva data on saatavilla myös tulevaisuudessa. Data pitäisi tallentaa luotettavaan ja pysyvään sähköiseen data-arkistoon tai tallennuspalveluun, joka vastuullisesti hoivaa dataa (*Turning FAIR into reality 2018).*
- FAIR-periaatteet eivät huomioi kattavasti myöskään **tietosuoja ja -turvanäkökulma**a, joka Tietoarkiston toiminnassa on oleellista ja CTS-sertifionnissa huomioitu.
  - Mm. käynnissä oleva FAIRsFAIR –hanke (Fostering FAIR Data Practices In Europe, H2020) työstää parhaillaan myös näitä osioita.

Parannettavaakin on

- Tietoarkisto *kehittää Rajoja ylittävät työkalut ja palvelut* (C-BoTS) -hankkeessa (2019-2023) SA:n FIRI2018 -rahoituksella mm. palvelujen yhteentoimivuutta muiden kansallisten ja kansainvälisten tutkimusdatan palvelujen kanssa.
- *Aktiivinen Tietoarkisto CESSDAssa* (TA22) –hankkeessa (SA FIRI 2016: 2017-2022) kehitetään mm. aineistojen versiointia. Jatkossa jokaisen aineiston jokaiselle data- ja metadataversiolle tulee omat pysyvät tunnisteet – ja tarpeen mukaan hautakivet.
- Mahdollisuuksia: lisää käyttötilastoja, ORCID, CESSDAn ja FINTOn sanastojen tiiviimpi integroiminen Tietoarkiston palveluihin, metadatan rikastaminen VIRRAn ja METAXin avulla (julkaisut, tutkimushankkeet…), hakurajapinnan avaaminen, OpenAIRE, käyttöehtojen harmonisointi CESSDA-tasolla.

Käytännön FAIR ei ole automatiikkaa tai -magiikkaa

Tutkimusaineistojen koko elinkaaren hyvään hallintaan ja FAIR-periaatteiden noudattamiseen tarvitaan

- asiantuntijoita,
- läpinäkyviä prosesseja,
- yhteentoimivia teknologioita ja
- yhteistyötä.

FAIR-MYYTTEJÄ RIITTÄÄ...

ONNEKSI MYÖS VASTAUKSIA.

KATSO LISÄÄ:
Myths about FAIR
https://vidensportal.deic.dk/en/FAIR

www.vidensportal.deic.dk/FAIR

FAIR data are: **Findable, Accessible, Interoperable** and **Reusable.**
Making your data FAIR means maximizing the project's output, increasing your impact and enhancing your recognition as a researcher.

©2018 by DeiC /Deff /DTU /CBS /AAU /KU /Rigsarkivet /Det Kgl Bibliotek
Policies is made available under a Creative Commons 4.0 licence (international)

Lähteitä ja luettavaa

- FAIR for Beginners, https://vidensportal.deic.dk/en/FAIR
- FAIRsFAIR, https://www.fairsfair.eu/
- GO FAIR, https://www.go-fair.org/
- Kleemola, Mari (2017). Tietoarkisto on FAIR. https://tietoarkistoblogi.blogspot.com/2017/02/tietoarkisto-on-fair.html
- Turning FAIR into reality. Final report and action plan from the European Commission expert group on FAIR data. https://doi.org/10.2777/1524
- Wilkinson, Mark D. et al (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3, Article number 160018. http://dx.doi.org/10.1038/sdata.2016.18

Kiitos!

ota yhteyttä

asiakaspalvelu.fsd (a) uta.fi

www.fsd.uta.fi

Liity sähköpostilistalle
https://www.fsd.uta.fi/fi/ajankohtaista/fsdlistaohje/

f YouTube **tietoarkisto**

‘ ‘ ***Saat ensimmäisten joukossa tietoa muun muassa julkaistuista aineistoista, Tietoarkiston tapahtumista ja ajankohtaisista avoimen tieteen uutisista*** ’ ’

TIETOARKISTO