



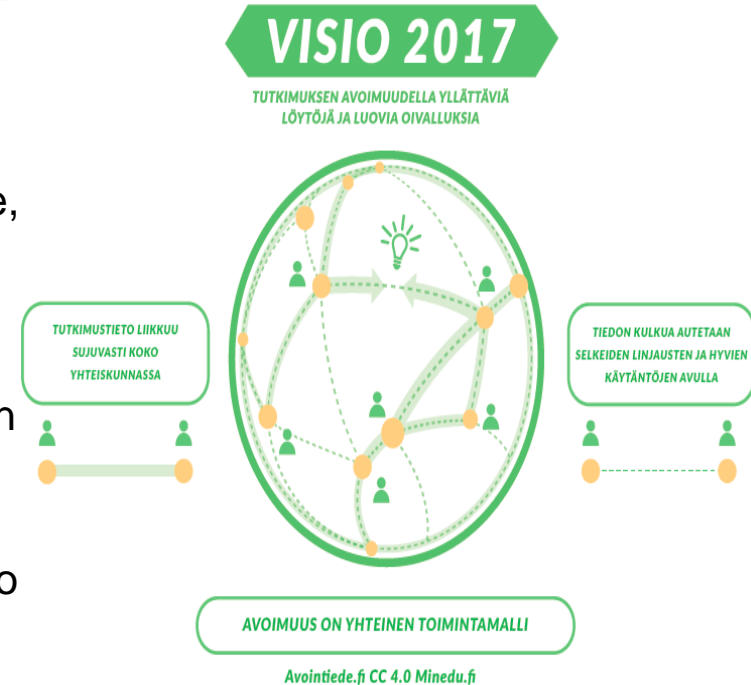
OPEN SCIENCE
AND RESEARCH

Miten avoin tiede ja tutkimus näkyy palveluiden kehityksessä?

Pirjo-Leena Forsström

Tieteen ja tutkimuksen avoimuudesta

- Avoin tiede ja tutkimus on **kansainvälisesti** merkittävä tapa edistää tiedettä ja tutkimuksen **vaikuttavuutta** yhteiskunnassa.
- Avoimuus on tieteen ja tutkimuksen keskeinen periaate, jonka kautta luodaan uusia mahdollisuuksia **osallistua** tieteen ja tutkimuksen tekemiseen tutkijoille, päätöksentekijöille ja kansalaisille.
- Edellytyksenä on tutkimuksen tuottamien ja tarvitsemien julkaisujen, datan, menetelmien, osaamisen ja tukipalveluiden laajamittainen **saatavuus**.
- Tutkimusprosessin digitalisoituminen ja avautuminen luo tutkijoille ja kaikille tutkimuksesta kiinnostuneille **uusia mahdollisuuksia** yhteistyöhön ja kommunikointiin. Luottamus tieteeseen kasvaa ja **yritystoiminta** vauhdittuu.



AVOIMUUDEN TUOMAT MAHDOLLISUUDET

- Toistettavuus
- Automatisointi
- Uudelleenkäyttö ennalta arvaamattomasti
- Yhteentoimivuus
- Mahdollisuus tavoittaa eri kohderyhmiä

**Tuotosten
saatavuus**

**Koneluettavuus,
rajapinnat, tunnisteet,
linkitykset, viittaukset...**

**Lisenssit ja
oikeudet**

**Semanttinen
yhteentoimivuus:
tietomallit, ontologiat etc.**

**Hyvä
saatavuus ja
käytettävyys**

Tutkimusaineisto, tutkimuksen tietoaaineisto



- **Avoimen tieteen käsikirja:** Tutkimusaineistoilla tarkoitetaan niitä resursseja, joita tutkija tuottaa tai joita hän käyttää tutkimusprosessin aikana. *Tutkimuksen tietoaaineistolla tarkoitetaan digitaalisessa, analogisessa ja fyysisessä muodossa olevia tieteellisen tutkimuksen perusaineistoja ja niistä jalostettuja aineistoja, joihin tutkimuksen löydökset ja tulokset perustuvat.* Tutkimusaineistot voivat olla aineellisia tai aineettomia. Jotta aineisto olisi tutkimukseen kelpaavaa, siihen täytyy liittyä tietoja ainakin sen alkuperästä. Tietoaaineistoihin yleensä kuuluvat sekä tutkimuksen tuottamat että tutkimuksen hyödynnettävissä olevat aineistot. Esimerkiksi datan liitteeksi tarvitaan kuvailevaa ja teknistä tietoa siitä mitä informaatiota se sisältää. Siksi tutkimusdataan liittyy paljon tietoa siitä, miten se on rakenteistettu ja koodattu, miten se on syntynyt ja miten sitä on käsitelty. Nämä tiedot on aina syytä tallentaa esimerkiksi metatietoihin, koodikirjoihin ja/tai muuhun dokumentaatioon. Yhdessä datan kanssa tästä kokonaisuudesta muodostuu tutkimusaineisto. On siis huomattava, että kaikki tutkimukselle hyödylliset tietoaaineistot eivät välttämättä ole alun perin kerätty tutkimuskäyttöön vaan esimerkiksi hallinnollista seurantaa, valvontaa, raportointia ja tilastointia varten.

Tietäminen ja tieto

- Mielekkään merkityksen, ajattelun ja kommunikaation horisontti, jossa:
 - Tiedon käytännöt: menetelmät, dokumentaatio, huolellisuus, tarkkuus, kriittisyys, objektiivisuus, eettisyys, aineistot ja dokumentaatio
 - Tietämisen tapahtumat: aiemman tutkimuksen tuntemus, harhojen tunnistaminen ja hallinta, johtopäätöksiä tekeminen
 - Tulevien tietämisen mahdollisuuksien jäsentely, toistettavuus ja todennettavuus

Viittaaminen,
ontologiat,
tunnisteet

Laadukas tutkimus kantaa ns. "episteemistä taakkaa" joka mahdollistaa merkityksen ja rajaa ja ehdollistaa mahdollisuuksia toimia. Hyvä tutkimusdatainfrastruktuuri helpottaa tätä taakkaa.

Uudistuva tutkimustieto

Tapahtumat

Käytännöt

Mahdollisuuksien jäsentely

Hyvä tieteellinen käytäntö, toistettavuus

Tietomallit, saatavuus, pysyvyys

Laadukasta tutkimusta tukee esimerkiksi:



- Tutkimusdatat, tutkimusjulkaisut ja tutkimusmenetelmät ovat avoimesti ja helposti löydettävissä ja saatavilla. Vastuut näiden elinkaaren eri vaiheisiin liittyvistä tehtävistä on sovittu ja kaikkien tiedossa, ja tehtävät on resursoitu riittävästi.
- Avoin tietoarkkitehtuuri mahdollistaa yhteentoimivan tutkimustietoinfrastruktuurin kehittämisen. Pitkäaikaissaatavuuteen liittyvä standardointityö on edistänyt tutkimusinfrastruktuurien ja korkeakoulujen parempaa tutkimustiedon hallintaa. Arkistot, kirjastot ja museot ovat luonteva osa tätä kokonaisuutta.
- Perusta on kunnossa suuriin haasteisiin vastaamiseksi, ja tutkimustieto ja tutkimusmenetelmät ovat aktiivisessa käytössä. Palvelut käytön helpottamiseksi ovat helposti löydettävissä ja niitä kehitetään toimintalähtöisesti.
- Toiminnan yhteentoimivuus kansallisesti ja kansainvälisesti on hallittua, selkeästi kuvattua ja helppoa.
- Hyvin toteutettu pitkäaikaissaatavuuspalvelu on tutkijoille helppo ja kannustava tapa toimia, ja aineistojen elinkaaresta huolehtiminen on luonteva osa tutkimuksen käytäntöjä.

- Tutkimusdata on saatavilla 1) tutkimuskäyttöön, 2) yhteiskunnallisen päätöksenteon tueksi ja 3) kaikille kiinnostuneille
 - 1) Tutkijoille: tutkimuksen resurssit (paikalliset, kansalliset, kansainväliset), tieteenalakohtaiset säilytyspaikat ja jakeluportaalit, vertaisverkosto
 - Jatkokäyttö, vertaisarviointi
 - 2) Yhteiskunnallinen päätöksenteko: koottuja datasettejä, koskien esim. Suomea tai suomalaisia
 - Esim. ilmastotutkimuksen havaintoaineistot
 - 3) Kiinnostuneet: avointa dataa hyvin kuvailtuna
 - joukkoistetut datan kuvailut ja keruut
 - Osaamisen kasvattaminen

Tutkimusaineistojen hallinnan infrastruktuuri



- Tulee tukea toimijaa, ei viedä kontrollia
 - *Tekijänoikeus, lisenssit, luvat, hallittavuus, palveluun vaikuttaminen*
- Tulee auttaa kasvattamaan laatua
 - *tietomallit, standardit, metadata, ...*
- Tulee tukea toistettavuutta
 - *Pitkäaikaissaatavuus: sopimukset, käytettävyys*
- Tulee tukea saatavuutta ja yhteentoimivuutta
 - *Ei siilo, avoin yhteistyö ja kehitys*
- Tulee olla kustannuksiltaan ennakoitava

Datainfrastruktuurin "blueprint"

Tuota ja tutki

- Kerää, tallenna, tutki, keskustele
- Löydä muuta relevanttia dataa

Datan valmistelu ja hallinta

- Kuvaile, paketoi, ylläpidä ja huolehdi elinkaaresta

Datan jakelu

- Luo jakelumalli ja rajapinnat, jaa metatietoa

Koulutus, ohjeet ja tukipalvelut

Tallennus

Optimoidaan nopeutta

Optimoidaan säilyvyyttä

Optimoidaan saavutettavuutta

Analyysityökalut

Yhteistyöalustat

Hakutyökalut

Metatiedot tuki: tietomallien hallinta, tunnisteet, ontologiat

Pitkäaikaissaatavuus: paketointi, sopimus, seuranta

Avoimet rajapinnat

Käyttövaltuutus

Jakeluportaalit

Datapolitiikka /Aineistonhallintasuunnitelmat /Säilytysuunnitelmat

Suunnitteluperiaatteet

FAIR principles

To be Findable:

- F1. (meta)data are assigned a **globally unique and eternally persistent identifier**.
- F2. data are described with **rich metadata**.
- F3. (meta)data are **registered or indexed in a searchable resource**.
- F4. metadata **specify** the data identifier.

To be Accessible:

- A1 (meta)data are **retrievable by their identifier using a standardized communications protocol**.
 - A1.1 the **protocol** is open, free, and universally implementable.
 - A1.2 the **protocol** allows for an authentication and authorization procedure, where necessary.
- A2 **metadata are accessible**, even when the data are no longer available.

To be Interoperable:

- I1. (meta)data use a **formal, accessible, shared, and broadly applicable language** for knowledge representation.
- I2. (meta)data use **vocabularies that follow FAIR principles**.
- I3. (meta)data include **qualified references** to other (meta)data.

To be Re-usable:

- R1. meta(data) have a **plurality of accurate and relevant attributes**.
 - R1.1. (meta)data are released with a **clear and accessible data usage license**.
 - R1.2. (meta)data are associated with their **provenance**.
 - R1.3. (meta)data **meet domain-relevant community standards**.

Datainfrastruktuuripalvelut: vaihtoehtoja



- Hyödynnetään kv-palveluita:
 - Sopimukset (saatavuus, hinta, exit-mallit), tietomallit ja metatiedot
 - Yhteentoimivuus?
- Hyödynnetään kansallisia palveluita:
 - ATT, tutkimusinfrastruktuurit (FSD, Kielipankki etc), organisaatioiden kautta tarjottavat palvelut (Kansalliskirjasto etc)
 - Yhteentoimivuus?
- Tehdään itse:
 - Hyödynnetään tunnettuja avoimen lähdekoodin ratkaisuja
 - Työmäärä? Yhteentoimivuus?
- Hybridi:
 - Luodaan optimikombinaatio

Suunnittelun lähtökohdat



- Toiminta nojaa laadukkaaseen tutkimusaineistojen hallintaprosessiin
- **Hyvin määritellyt modulaariset palvelurakenteet ja muut kokonaisuudet**
- Tutkimusaineistojen suunnitelmallisen hallinnan tarjoaminen omistajilleen ja asiakkaalle keskiössä
- Kokonaisuus tukee ja ohjaa tutkijan ja organisaatioiden toimintaa pitkäaikaissaatavuutta tukevaan suuntaan
- Minkä voi automatisoida, se tehdään
- Minkä käyttäjä haluaa itse tehdä, siihen on mahdollisuus
- Työkaluja käytön ja prosessin laadun edistymisen seurantaan
- Järjestelmä ohjaa työn kulkua ja minimoi virheiden syntymisen
- Tuetaan niitä prosesseja, joissa tieto syntyy ja sitä käsitellään
- Yksinkertaistetaan prosesseja niin paljon kuin se on mahdollista

Tärkeimmät ATT-tietojärjestelmäpalveluiden suunnittelussa ja toteutuksessa noudatettavat SOA-ratkaisuperiaatteet ovat:

- Uudelleenkäytettävyys
- Systemaattisuus
- Palvelevuus toiminnan näkökulmasta
- Tekninen ja semanttinen yhteentoimivuus
- Joustavuus ja laajennettavuus

Datainfrastruktuurin "blueprint" ja ATT-palvelut

Tuota ja tutki

- Kerää, tallenna, tutki, keskustele
- Löydä muuta relevanttia dataa

Datan valmistelu ja hallinta

- Kuvaile, paketoi, ylläpidä ja huolehdi elinkaaresta

Datan jakelu

- Luo jakelumalli ja rajapinnat, jaa metatietoa

Koulutus, ohjeet ja tukipalvelut

Tallennus

Optimoidaan nopeutta

Optimoidaan säilyvyyttä

Optimoidaan saavutettavuutta

IDA

Analyysityökalut

Metatiedot tuki: tietomallien hallinta, tunnisteet, ontologiat

Avoimet rajapinnat

Yhteistyöalustat

Qvain, METAX, Finto. IOW

Käyttövaltuutus

Hakutyökalut

Pitkäaikaissaatavuus: paketointi, sopimus, seuranta

REETTA

ETSIN

Tutkimuksen PAS-palvelu

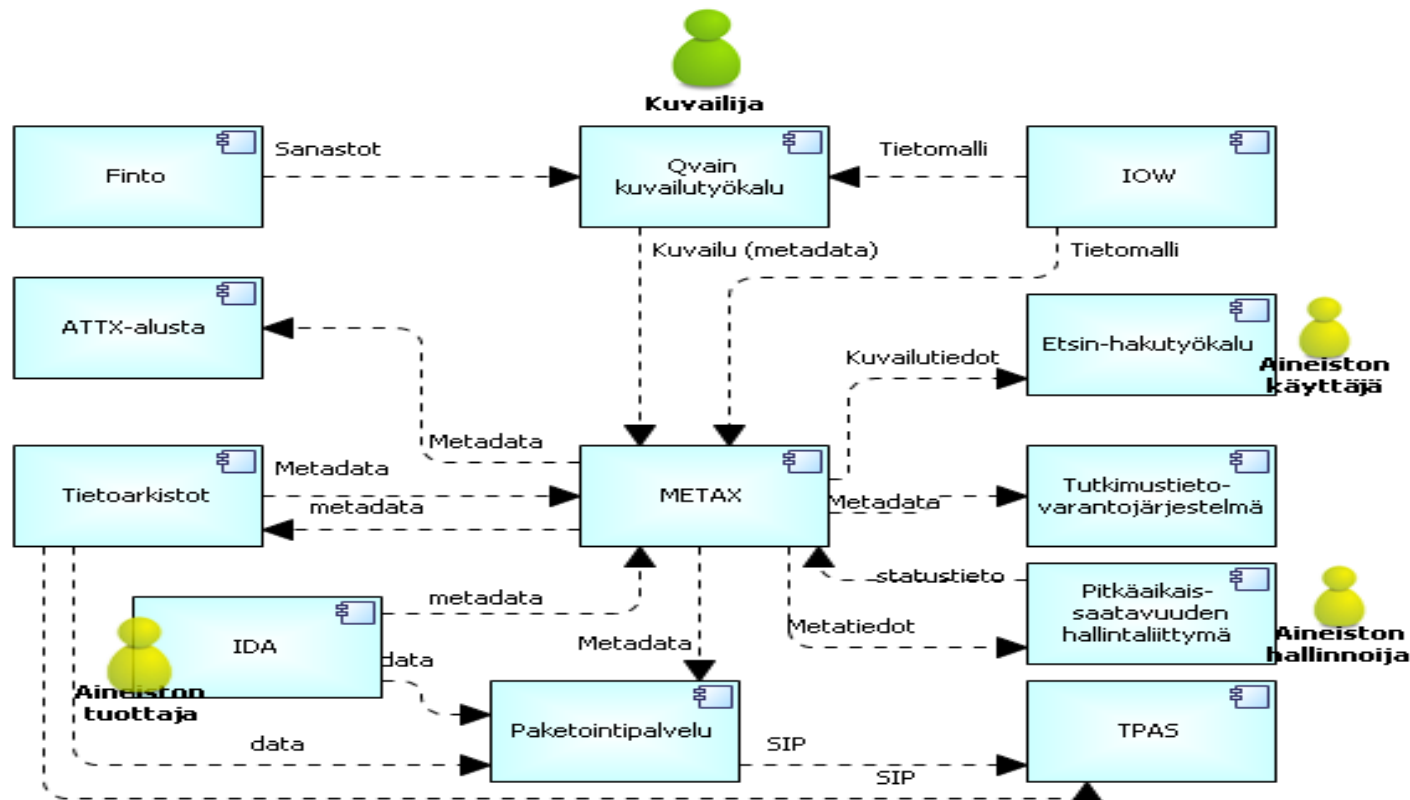
Jakeluportaalit

AVAA

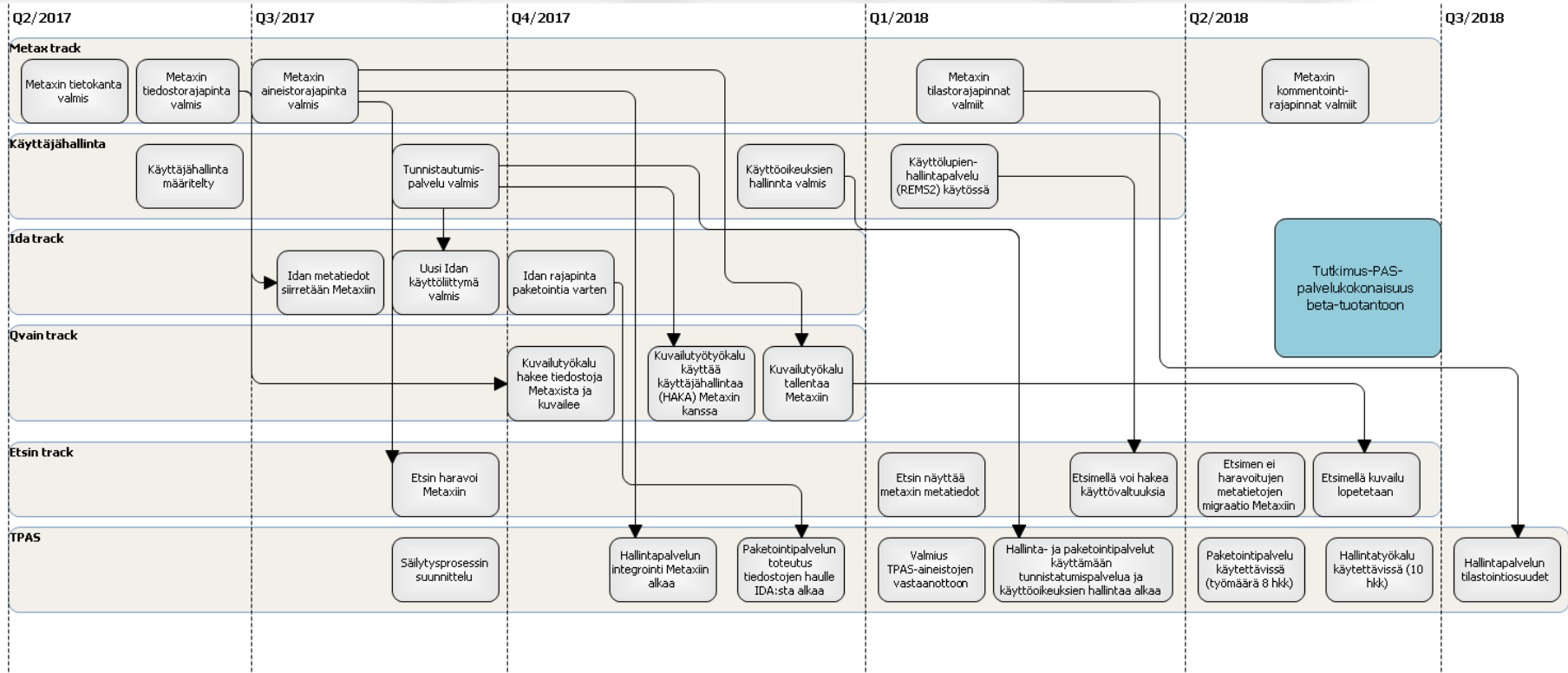
Datapolitiikka /Aineistonhallintasuunnitelmat /Säilytyssuunnitelmat

- Säilytys: IDA -> Keskipitkä säilytys -> Pitkäaikaissäilytys
- Kuvailu: tietomalli (IOW) -> Qvain tai haravointi -> METAX
- Hallinta: hyväksyntä -> sopimus -> siirto pitkäaikaissäilytykseen (paketointi) -> tilan hallinta
- Käyttö: metatietojen tarjoaminen rajapinnan kautta (muut palvelut) tai haku (Etsin)

Tutkimus-PAS- kokonaisuus uudistuu- metatietovirrat



Tutkimus-PAS-palveluiden kehittämisen aikataulu



- OKM kysyi korkeakouluilta tiedossa olevista, merkittävistä tutkimusaineistoista sekä näiden
 - Säilytyspaikasta
 - Miten säilytyksestä sovittu
 - Miten saatavilla
 - Tarve saada pitkäaikaissäilytyspalveluun
- Mallista, joka olisi toivottava pitkäaikaissäilytykseen aineistoja otettaessa

- Kokonaisuuteen tarvittava prosessi rakentuu
 - Päätösprosessi
 - Millä ehdoilla ja miten aineistoja TPASiin?
 - Keskustelut korkeakoulujen kanssa tarpeista ja aikataulusta
 - Sopimusprosessi
 - Etenee keskusteluiden pohjalta
 - Saatavuusprosessi:
 - Määritetään tarpeen mukaan
 - Käyttäjille mahdollisimman helppo
 - Saumaton palveluketju

- Tavoite:
 - Niiden korkeakoulujen kanssa, jotka ovat ilmoittaneet voivansa ottaa palvelun käyttöön 2017 tai 2018, käydään neuvottelut etenemisestä elo-syyskuussa 2018. Tuloksena: tarkennettu aikataulu ja tarvetiedot.
 - Muiden korkeakoulujen kanssa käydään keskustelut loppuvuodesta 2017/alkuvuonna 2018.

- Elämme tietäen ja tiedämme eläen

KIITOS!