

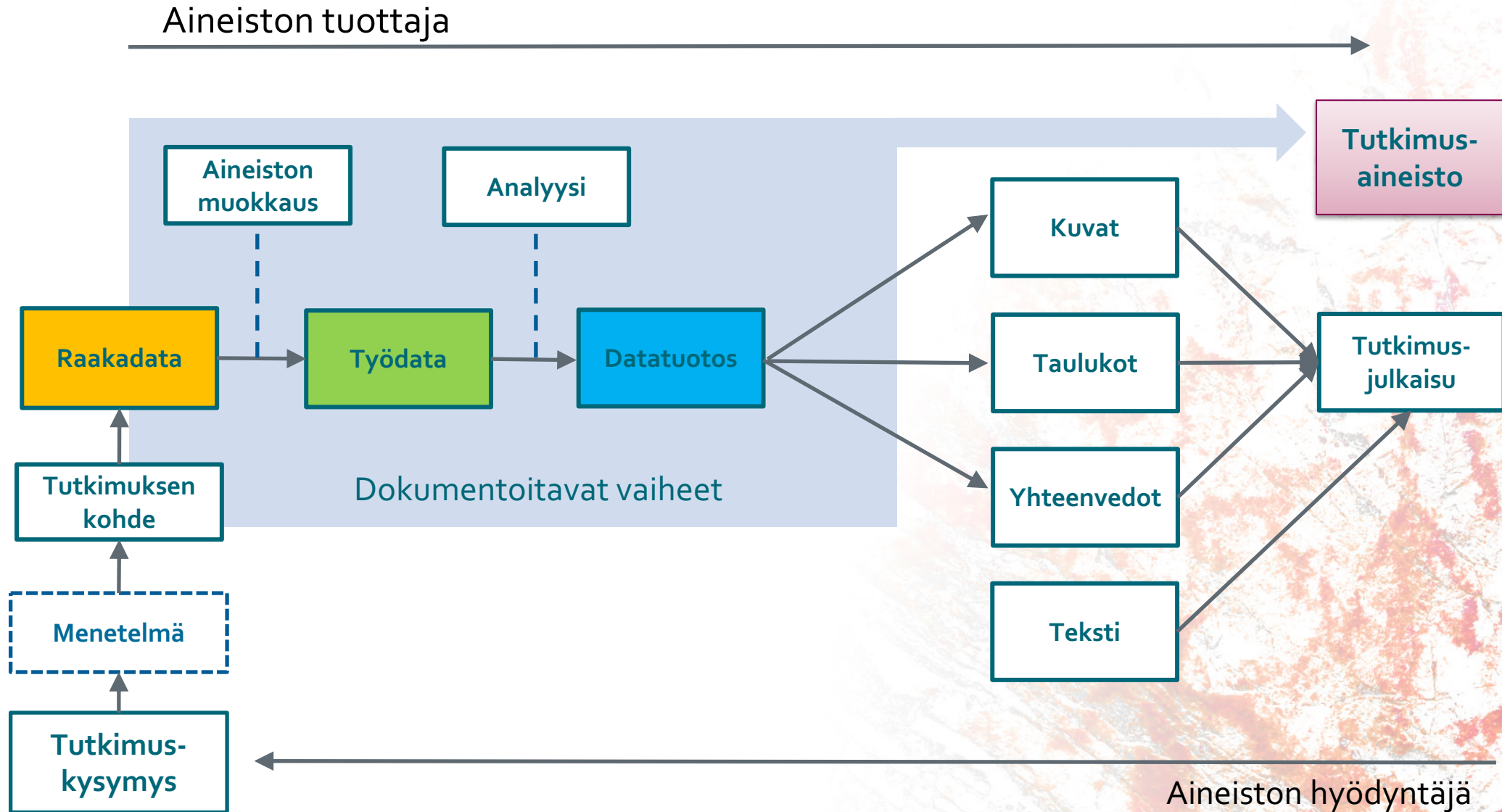


# Fairdata-palvelut ja PAS

PAS-seminaari 25.4.2023 / Juha Lehtonen



# Tutkimuksen tuotokset



# FAIR-periaatteet

**F****FINDABLE – LÖYDETTÄVÄ**

- Olennaiset tiedot kuvailtu riittävän tarkasti
- Kuvailusivu ja ainutkertainen pysyvä tunniste

**A****ACCESSIBLE – SAAVUTETTAVA**

- Voidaan hakea internetistä
- Versiointi ja elinkaaren dokumentointi
- Muistosivu jos data on poistettu

**I****INTEROPERABLE – YHTEENTOIMIVA**

- Käytössä yleiset, dokumentoidut ja avoimet tiedostomuodot
- Yhteisten koodistojen käyttö & linkitykset muihin tuotoksiin

**R****RE-USABLE – UUELLEENKÄYTETTÄVÄ**

- Data on hyvin dokumentoitu ja ymmärrettävää
- Käyttöoikeudet/lisenssit selkeästi esillä



# Aineiston avaaminen

Mitä käyttöä aineistollani voisi olla?

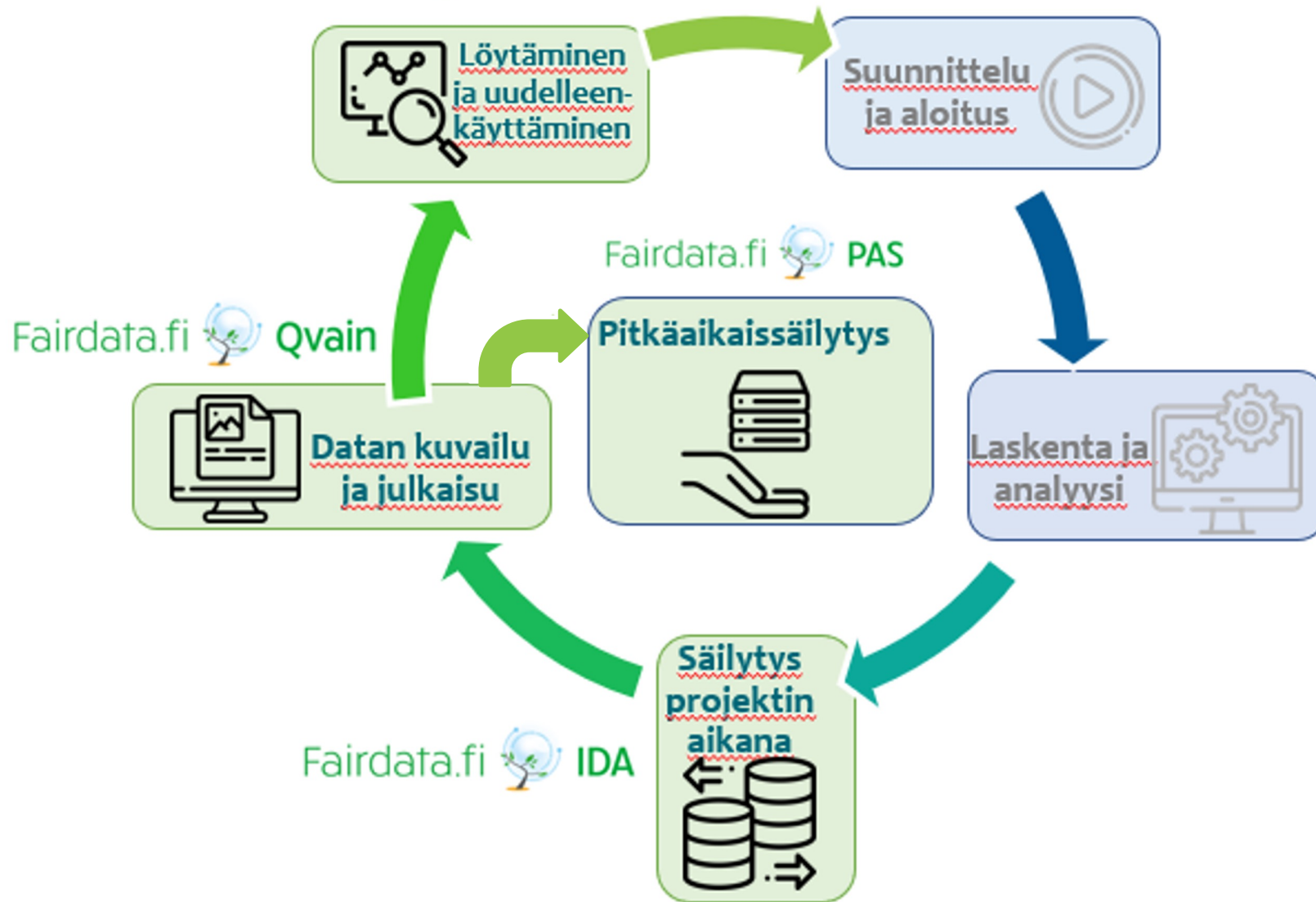
1. Tutkimuksen verifiointi
2. Jatkoanalyysi
3. Meritoituminen
4. Tutkimusalan yhteinen resurssi / tavoite
5. Tutkimusjulkaisut (data paper)
6. Opetuskäyttö
7. Oma käyttö

<https://www.dcc.ac.uk/guidance/how-guides/five-steps-decide-what-data-keep>

**Webinaari: Aineistojen kuvailu ja julkaisu Fairdata-palveluissa**  
**Huomenna 26.4.2023 / [fairdata.fi](https://fairdata.fi)**

Uusia  
yhteistyömahdollisuuksia kun  
muut löytävät tutkimuksesi?

Muiden avaaman  
aineiston  
hyödyntäminen?



**Fairdata-palvelut** soveltuvat eri alojen digitaaliselle tutkimusmateriaalille ja niihin liittyville kuvailutiedoille

**Fairdata-palveluiden** käyttö on maksutonta suomalaisille korkeakouluille ja tutkimuslaitoksille

# Pitkäaikaissäilytys eli pitkäaikaissaatavuus

- Tarkoituksena säilyttää digitaaliset aineistot käyttökelpoisina ja autenttisina useiden vuosikymmenien tai satojen vuosien ajan
- Mahdollistaa digitaalisten aineistojen hyödyntäminen tuleville sukupolville
- Toteutuksen näkökulmasta korostaa jatkuvuussuunnittelua
  - Toiminnan elinkaari ylittää kaikkien toteutuksen osakokonaisuuksien käyttöajan
- Pitkäaikaissäilyttäminen on aina aktiivista toimintaa



# Pitkäaikaissäilytyksen OAIS-viitemalli

## ISO 14721:2012 (suomennos SFS 5972)

- Keskeisiä käsitteitä:

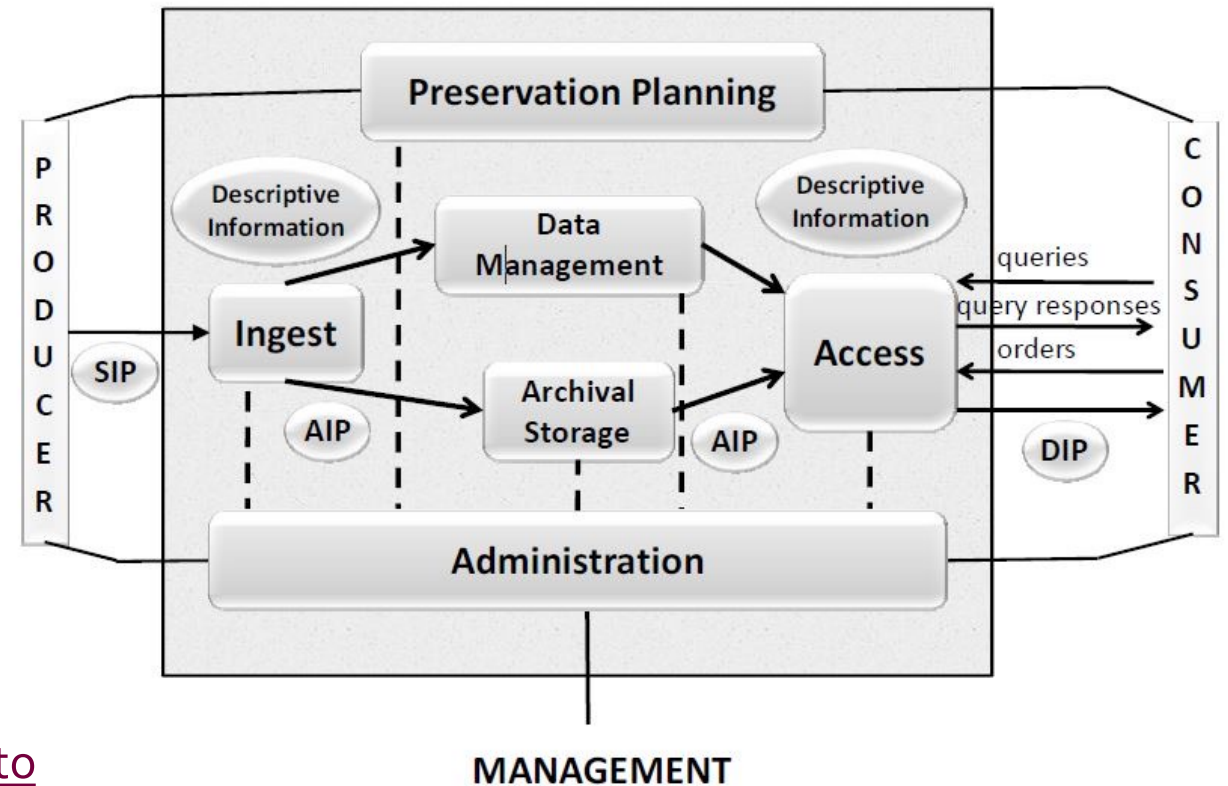
- Siirtopaketti (SIP)
- Säilytyspaketti (AIP)
- Jakelupaketti (DIP)

- Kohdeyhteisö  
(designated community)

”Tietty ryhmä potentiaalisia käyttäjiä, joiden pitäisi pystyä ymmärtämään tiettyjä aineistoja.”

- PAS-sanasto:

- <http://digitalpreservation.fi/specifications/sanasto>



# Soveltuuko aineistoni pitkäaikaissäilytykseen?

Tutkimusorganisaatio tekee kriteeristön ja valinnat

Kysymyksiä, joita esimerkiksi voi pohtia:

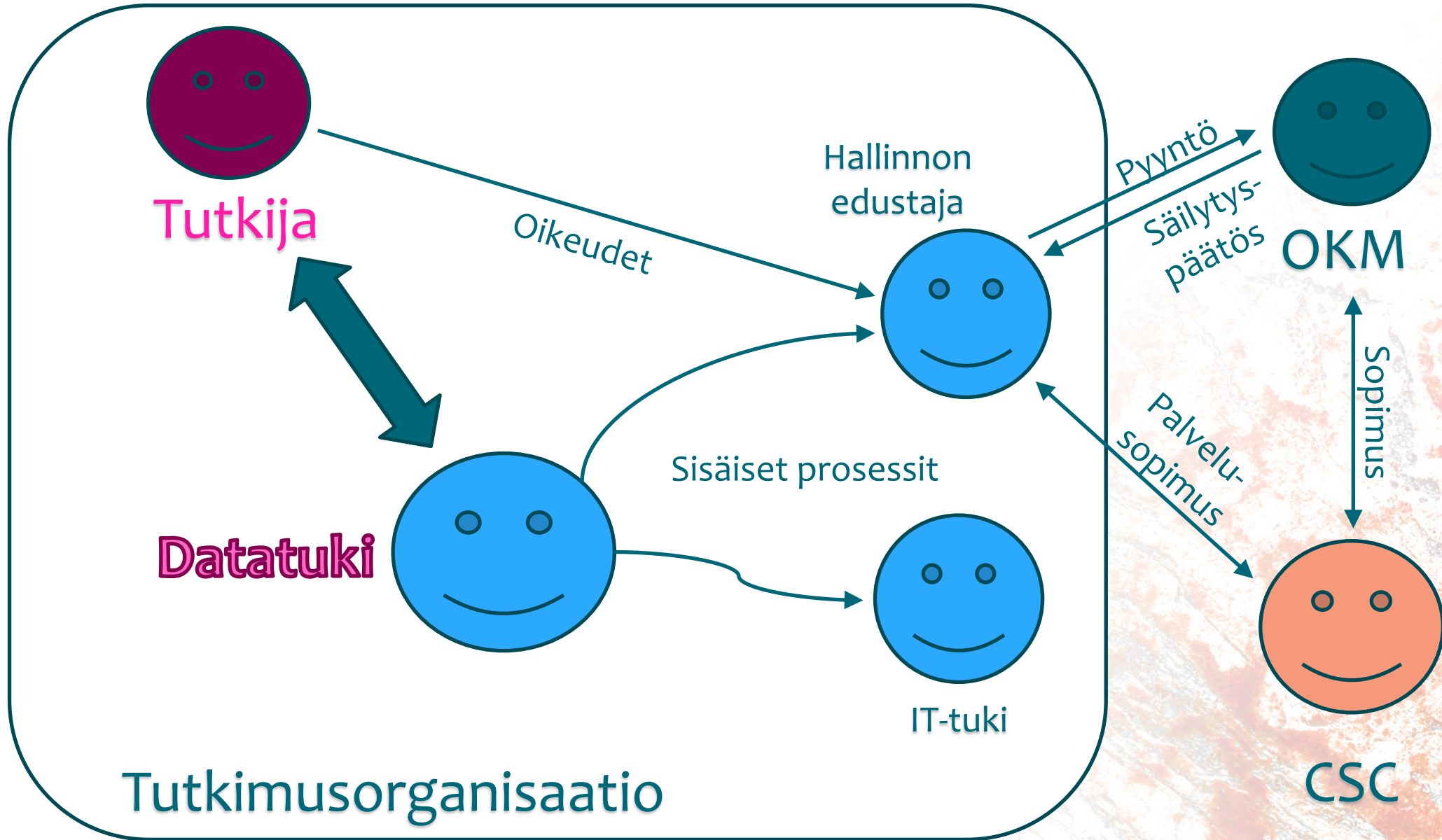
1. Edellytetäänkö pitkäaikaissäilytystä?
2. Onko aineistolla merkittävä tutkimuksellinen / kulttuurinen / historiallinen / taiteellinen arvo?
3. Tukeeko aineisto organisaation missiota?
4. Onko aineisto ainoa laatuaan?
5. Onko aineisto tuotettavissa uudelleen?
6. Säilyttämisen vs. kadottamisen hinta? Onko aineisto vaarassa kadota?
7. Sisältyykö aineistoon pitkäaikaissäilytykselle rajoituksia?
8. Onko aineisto kestävässä muodossa? (tiedostot, metatiedot)

- Mears: **Guidelines for selecting research data for retention and preservation**. The Open University, 2013.
- Tallman & Work: **Approaching Appraisal**. iPRES 2018, September 24-28, 2018, Boston, USA.
- Choy et al.: **The UNESCO/PERSIST guidelines for the selection of digital heritage for long-term preservation**. UNESCO/PERSIST Content Task Force, 2016.



# Pitkäaikaissäilytyksen sopimukset





Tutkimusorganisaatio

CSC

# Voiko arkaluonteista aineistoa säilyttää PAS-palvelussa?

- Aineistoihin sisältyvien henkilötietojen suhteen hyödyntävä organisaatio vastaa rekisterinpitäjän velvollisuuksista
  - Hyödyntävän organisaation oma vaikutustenarviointi on oleellinen
  - CSC toimii henkilötietojen käsittelijänä
- DPIA-arviointipohja
  - DPIA: Tietosuojaa koskeva vaikutustenarviointi (data protection impact assessment)
  - Arviointipohjan tarkoituksena on helpottaa arvioinnin tekemistä
- PAS-sopimukseen liitetään käsittelytoimien kuvaus ja tarvittaessa turvallisuusliite
- PAS-palvelut sisältyvät CSC:n ISO/IEC 27001:2013 tietoturva-auditoinnissa tarkasteltuihin palveluihin
- Muut Fairdata-palvelut (esim. IDA) eivät sovellu arkaluonteiselle aineistolle

# Fairdata PAS-palvelun käyttötarkoituksia

Organisaatio	Käyttötarkoitus
Geologian Tutkimuskeskus	GTK:n tomografialaitteen tuottamat tietoaaineistot
Helsingin yliopisto	Helsingin yliopiston SMEAR-aineistojen valikoima meteorologia - ja ilmanlaatumittauksia
Helsingin yliopisto	M. cinxia and C. melitaeorum in the Åland metapopulation system
Helsingin yliopisto	FIRE (The Finnish Reflection Experiment)
Helsingin yliopisto	Luomuksen aineistot
Itä-Suomen yliopisto	SENSOTRA
Jyväskylän yliopiston kiihdytinlaboratorio	250-Nobeliumin hajoamisspektroskopia
Oulun yliopisto, Sodankylän geofysikaalinen observatorio	Havaintoaineistot
Tampereen yliopisto	Yhteiskuntatieteiden tiedekunnan Kansanperinteen arkiston A-K-kokoelma
Turun yliopisto	Historian, kulttuurin ja taiteiden tutkimuksen arkiston aineistot (HKT-arkisto)
Åbo Akademi	Samlingar vid Åbo Akademis bibliotek

# Mitä hyödyntäminen edellyttää organisaatiolta?

- Sopimusmalliin tutustumista
- PAS-palvelun hyödyntämisestä sopimista OKM:n kanssa
- Säilytettävän aineistokokonaisuuden tunnistamista
- Käyttöönottoprosessin läpikäymistä yhdessä CSC:n kanssa
- Aineistojen riittävää kuvailua
- Aineistojen paketointia PAS-määritysten mukaisesti
- Aineistojen siirtämistä PAS-palveluun

**Kaikki tämä edellyttää hyödyntävältä organisaatiolta henkilöresursseja palvelun käyttöönottoon ja käyttöön**

mutta...

**PAS-palvelu  
tukee  
organisaatioita  
kaikissa  
vaiheissa**

# PAS-palvelun hyödyntäminen on yhteistyötä



PAS ei ole “itsepalvelu”-palvelu

Ratkaistaan ongelmat yhdessä.  
Kysyvä ei tieltä eksy.



PAS-palvelu tukee organisaatioita  
kaikissa vaiheissa

Tutustumisesta käyttöönottoon ja  
käytön aikana.



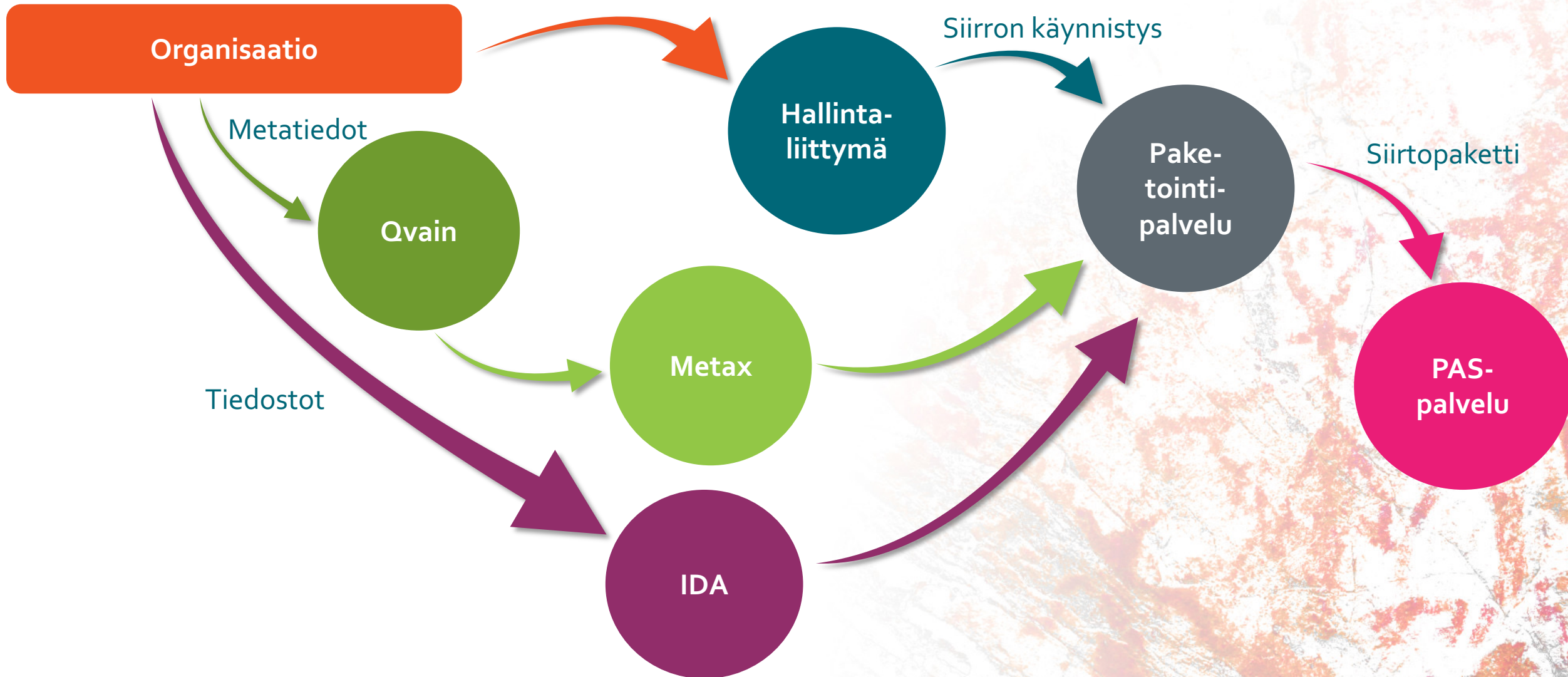
Keskustellaan parhaat tavat  
alusta asti yhdessä

Ja jatketaan yhteydenpitoa.

# Esimerkkiväylä pitkäaikaissäilytykseen

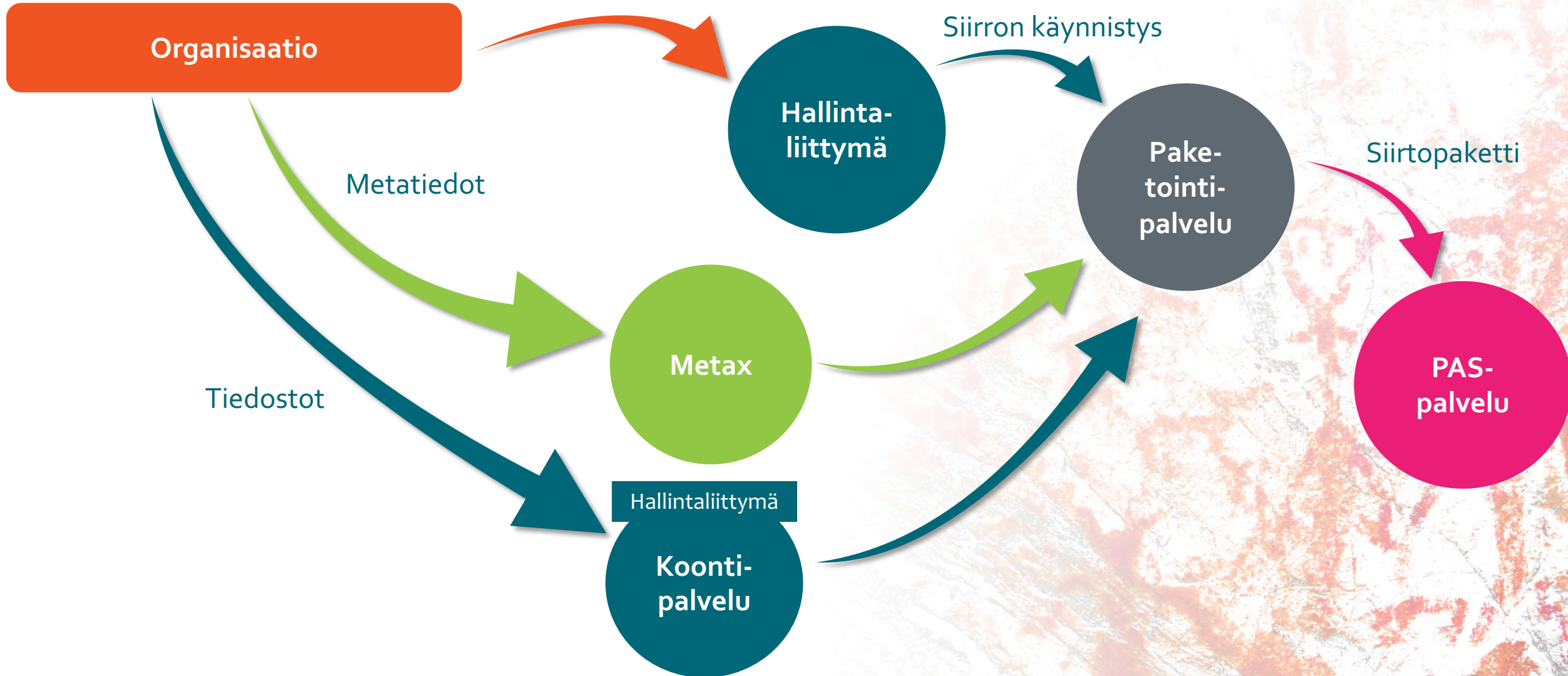


# Esimerkkiväylä: IDA & Qvain





# Esimerkkiväylä: Tiedostojen koontipalvelu



# Yhteistyötä organisaatioiden kanssa nyt ja tulevaisuudessa

(myös muiden kuin asiakkaiden)

- Fairdata-verkosto
  - <https://www.fairdata.fi/fairdata-verkosto/>
- PAS-palveluiden sähköpostilista
  - <https://www.digitalpreservation.fi/2021-liity-pas-palveluiden-sahkopostilistalle>
- Kuukausittaiset #PASKaffet
  - Joka kuukauden ensimmäinen perjantai klo 14:00-14:45
- #PASKlinikat
  - Säännöllisen epäsäännöllisesti
- Koulutustilaisuuksia enemmän tai vähemmän säännöllisesti
- [twitter.com/dpres\\_FI](https://twitter.com/dpres_FI)
- Virallinen tukiosoite: [pas-support@csc.fi](mailto:pas-support@csc.fi)
- <https://www.digitalpreservation.fi>
- <https://www.fairdata.fi>



**YOU'RE NOT  
ALONE...**

[pas-support@csc.fi](mailto:pas-support@csc.fi)