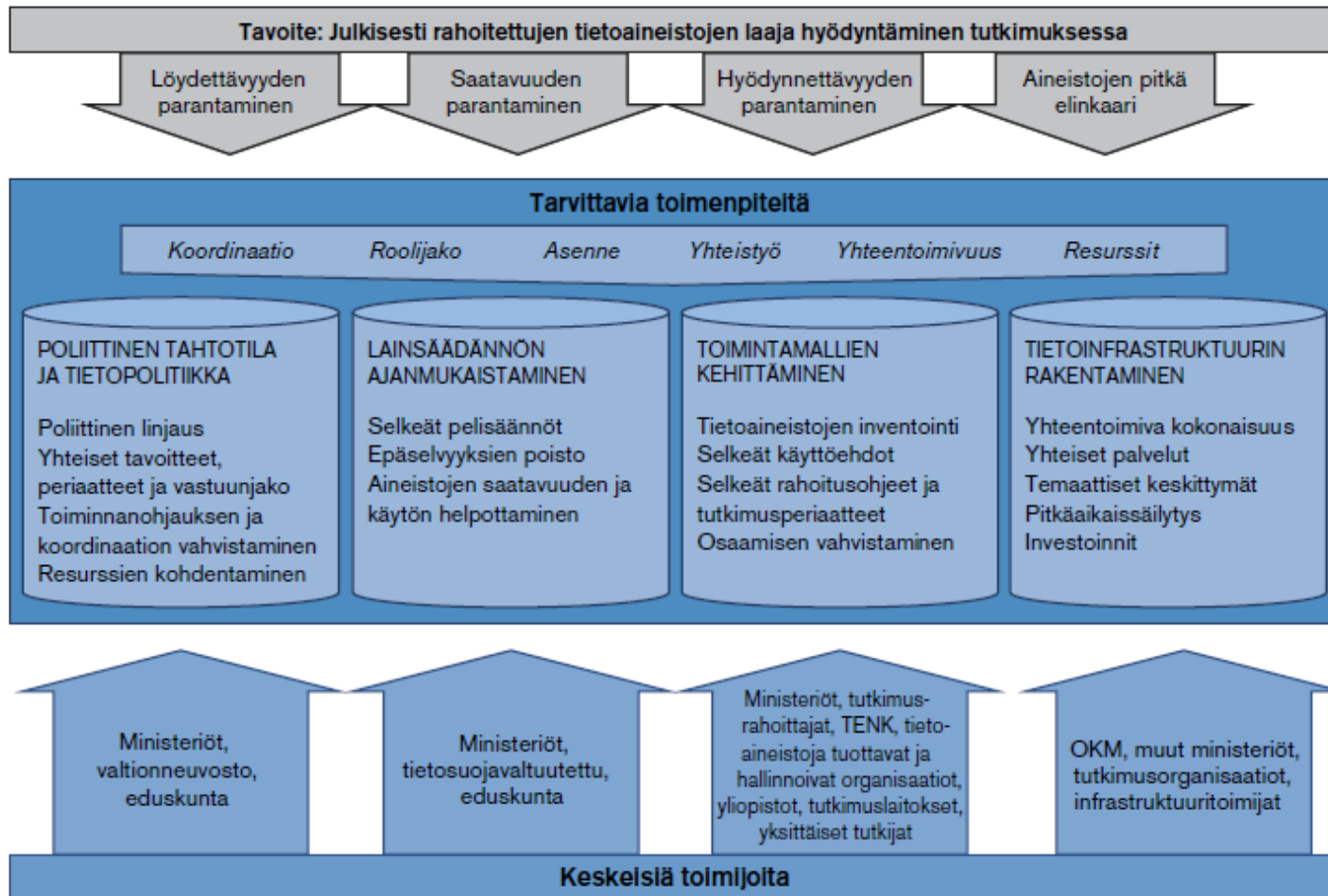




# TTA-hankkeen esittely

Pirjo-Leena Forsström  
TTA-hankkeen pääsihteeri

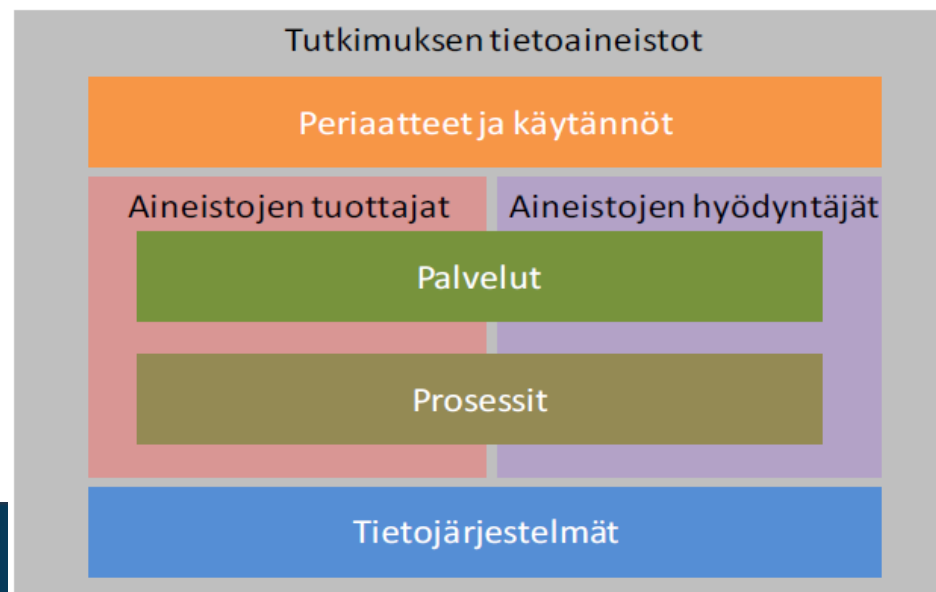
# Toimenpiteitä tietoaineistojen hyödyntämisen parantamiseksi



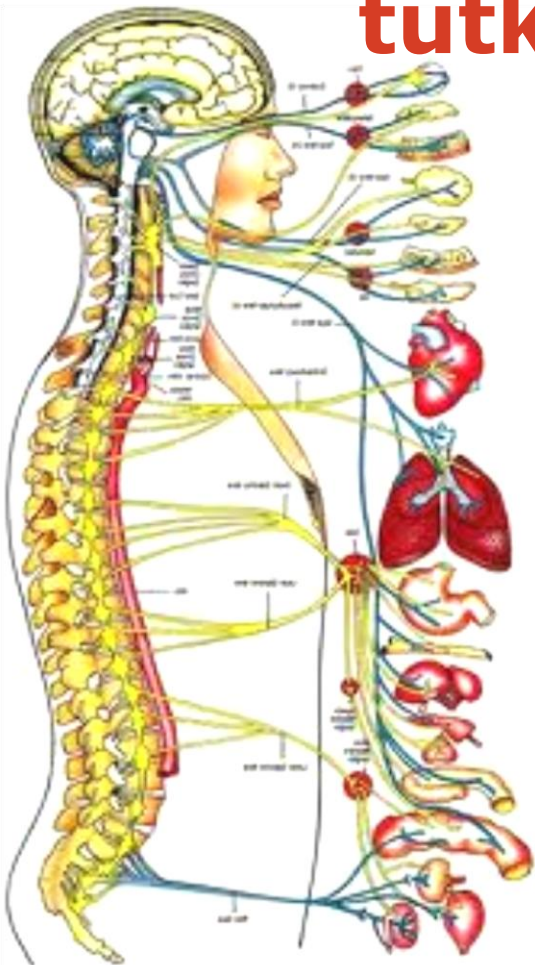
# TTA-VISIO



- Suomessa on selkeä tietopolitiikka, jota yhteiset sähköiset palvelut tukevat.
- Julkisen sektorin hallinnoimat tietoaineistot sekä julkisen tutkimusrahoituksen tuella syntyneet tietoaineistot ovat lainsäädännön ja yhtenäisten käyttöehtojen ohjaamina pääsääntöisesti maksutta yhteiskunnan hyödynnettävissä.
- Pitkäjänteinen tietoinfrastruktuurin kehittämisen ja ylläpidon rahoitusjärjestelmä takaa, että olemassa olevat ja uudet tietoaineistot kuvaillaan ja ne ovat helposti löydettävissä ja otettavissa käyttöön tietoverkossa olevien palveluiden avulla.
- Kannustava ja oikeudenmukainen meriittijärjestelmä varmistaa uusien, laadukkaiden tietoaineistojen liittämisen tietoinfrastruktuuriin.



# TTA, tutkimuksen tietoinfrastruktuuri



- "hermosto", tiedonvälitys- ja säätelyjärjestelmä
- TTA-palvelukokonaisuus edistää tutkimuksen tietoaineistojen kansallista, eurooppalaista ja kansainvälistä yhteentoimivuutta

# Määritelmiä

- Tutkimuksen tietoaaineistolla tarkoitetaan tässä työssä julkisin varoin tuotettuja sähköisiä tietoaaineistoja ja -varantoja. Tietoaaineistoihin kuuluvat sekä tutkimuksen tuottamat että tutkimuksen hyödynnettävissä olevat aineistot.
- Tutkimuksen tietoinfrastruktuurilla tarkoitetaan aineistojen sijaintiin, rakenteeseen, organisointiin, hakemistoihin ja luetteluihin, omistajuuteen, saatavuuteen, varmistukseen, tietoturvaan ja tietovarastoihin liittyvät keskeiset resurssit ja kyvykkyydet sekä niiden elinkaaren hallinta.
- TTA-palveluilla tarkoitetaan niitä tutkimuksen tietoinfrastruktuurin palveluita, jotka toteutetaan TTA-hankkeessa. Joulukuuhun 2012 mennessä tällaisiksi palveluiksi on sovittu tallennuspalvelu IDA, metatietopalvelu KATA sekä pitkäaikaissäilytys PAS. Jatkossa uusista palveluista sopiminen tapahtuu valitun hallintomallin puitteissa

## Käytännön tavoitteena on:

- tutkimuksen tietoinfrastrukturi palveluineen toimii saumattomasti yhteen muun kansallisen tietoinfrastruktuurin kanssa ja tarjoaa tutkimuksen tietoaineistojen säilyttämisen ja hyödyntämisen yhteiset palvelut.
- Tietoinfrastruktuurin rakentaminen, kehittäminen ja ylläpito on pitkäjänteistä ja takaa, että tietoaineistot kuvataan ja tuodaan tietoinfrastruktuuripalvelujen piiriin.
- Eri toimijoiden välinen roolijako on selkeä.
- Kaikista tutkimuksen kannalta merkittävistä tietoaineistoista on tuotettu tarvittavat metatiedot ja kuvaukset on koottu niin, että tietoaineistot on helposti löydettävissä.
- Tietoaineistot säilytetään pysyvästi ja niille on annettu pysyvä tunnus.
- Tietoaineistot ovat palveluiden avulla helposti löydettävissä, saatavissa ja käytettävissä.
- Tutkimusorganisaatiot ja yhteenliittymät vastaavat tietoaineistojen metatietotyöhön tarvittavista työkaluista ja järjestelmistä.

# Tietoaineistojen käyttötapoja



# TTA-hankkeen hyödyt

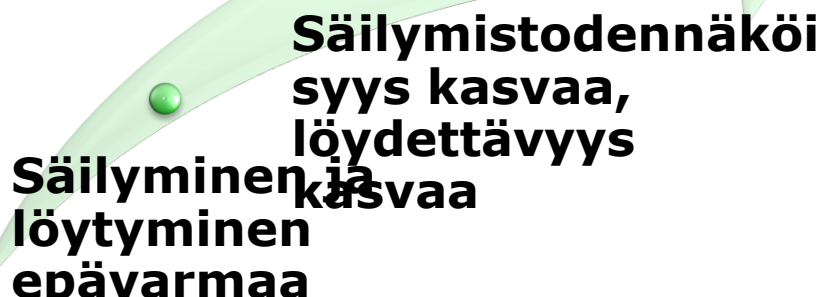
- Keskitettyä tietoa tutkimuksen tietoaineistoista: helpompi löytää, helpompi käyttää
- Yhtenäisempiä käytäntöjä aineistojen hallintaan
- Yhteentoimivuuden lisääntyminen: metatietomalli, rajapinnat
- Monipuolisen palvelukokonaisuuden kehitys
- Tietoaineistojen säilymisen turvaaminen

**Tier 1** – International data services

**Tier 2** – National data services

**Tier 3** – Institutions (Universities & Institutes)

**Tier 4** – “Small science” researchers & research groups



Säilyminen ja löytäminen epävarmaa  
Säilymistodennäköisyys kasvaa, löydettävyys kasvaa



# Tutkimuksen tietoaineistot TTA

TTA-hanke edistää tietoaineistojen **kuvausten yhtenäistämistä, säilytystä ja käyttöä.**

TTA-hankkeessa tuotetaan tätä tukemaan:

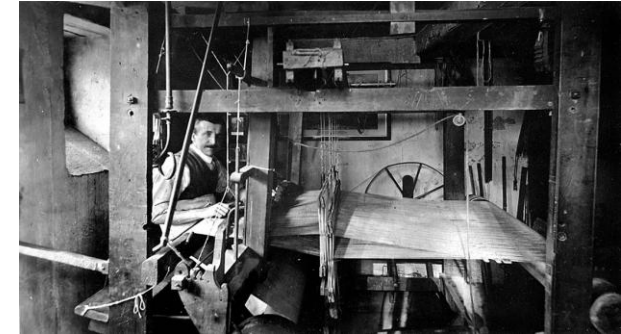
- *tutkimuksen tietoaineistojen prosessikartta*
- *tietoinfrastruktuurin palvelukokonaisuuden suunnittelu*
- *metatietomalli tutkimustiedon hallinnalle*
- *yhteinen metatietokatalogi/hakupalvelu tutkimusaineistoille*
- *tallennuspalvelu tutkimuksen tietoaineistoille*
- *yhteisen pitkäaikaissäilytysratkaisun valmistelu*

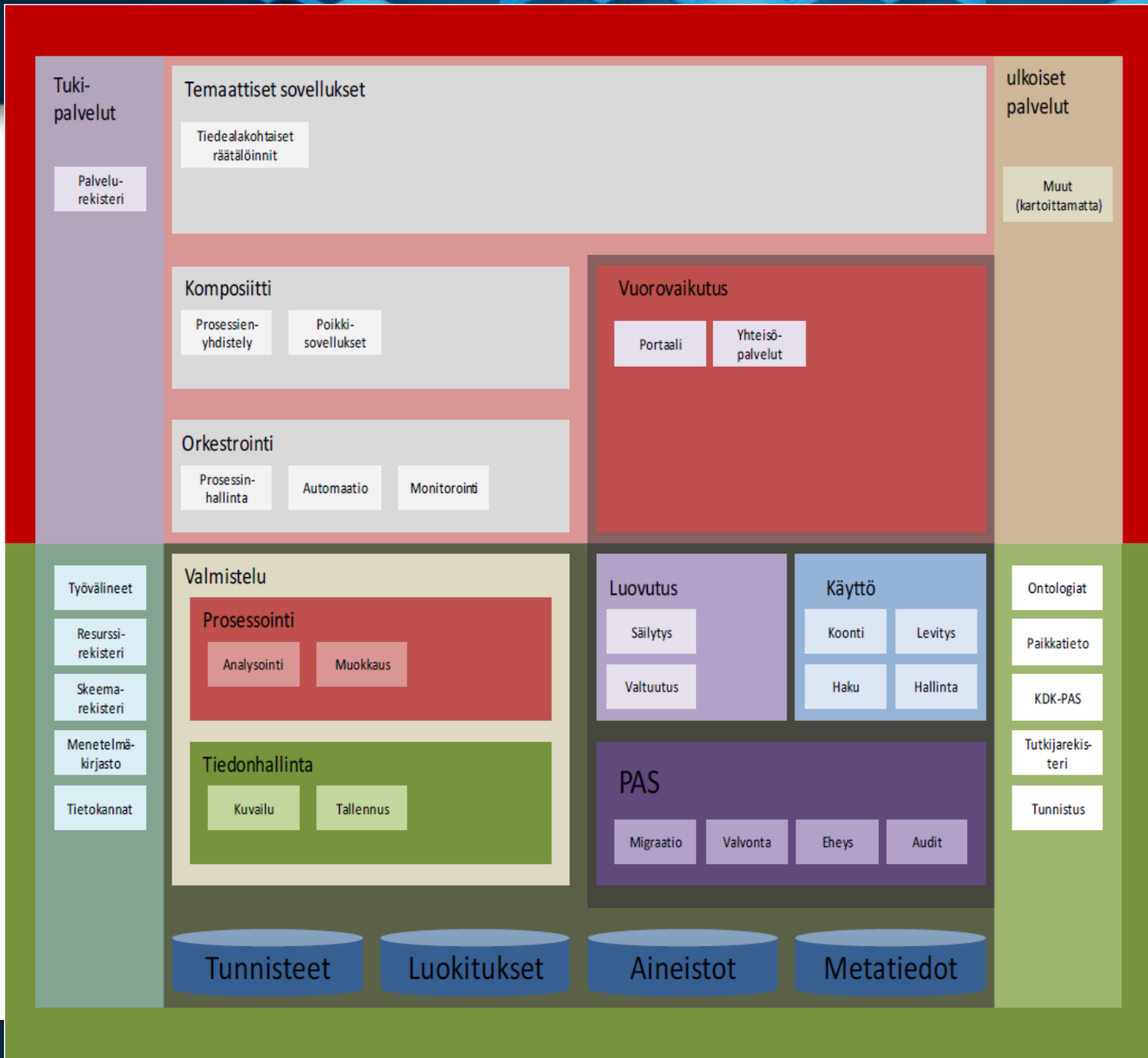
TTA-palvelukokonaisuus edistää tutkimuksen tietoaineistojen kansallista, eurooppalaista ja kansainvälistä **yhteentoimivuutta.**

# Tutkimustiedon infrastruktuurin

täytyy tukea:

- koko tutkimusprosessia
- tietoaineistojen siirtoa tieteenalojen välillä
- avoimen datan tuottoa
- työvuopohjaista toimintaa
- julkaisujen ja tietoaineistojen linkitystä
- tiedepolitiikan kehikkoa





# TTA palvelu-arkkitehtuuri

Koordinaatioryhmä

Ohjausryhmä

Metatieto  
-  
työryhmä

Tietoinfra-  
strukturi-  
työryhmä

TURE-  
pienryhmä

PAS

Tiedon  
saatavuus-  
työryhmä

TUTKIMUS-IDA

DATAKATALOGI

DATANHALLINNAN  
PILOTTI

AVAA-ALUSTA

TTA-hankkeen palveluiden periaatteet koostuvat

- 1) yleisistä arkkitehtuuriperiaatteista
- 2) yleisistä palveluperiaatteista
- 3) palvelukohtaisesta käyttö- ja palveluperiaatteista

# TTA-hankkeen arkkitehtuuriperiaatteet



- Tutkimuksen tietoinfrastrukturi noudattaa ohjaavia periaatteita
- Tutkimuksen tietoinfrastruktuuria johdetaan strategialähtöisenä kokonaisuutena
- Tutkimuksen tietoinfrastruktuurissa on vältettävä tarpeetonta monimutkaisuutta
- Tutkimuksen tietoinfrastrukturi on kerroksellinen, toiminnalliset ja tekniset tasot ymmärretään erillään toisistaan mutta toisiaan tukevia ja täydentäviä.
- Tutkimuksen tietoinfrastrukturi pyrkii olemassa olevien komponenttien uudelleenkäyttöön ja yhdistelyyn tavoitteita tukevaksi kokonaisuudeksi
- Tutkimuksen tietoinfrastrukturi palvelee yhteiskunnan kokonaisetua
- Tutkimuksen tietoinfrastrukturi mukautuu eri tieteenalojen ja toimintojen tarpeisiin
- Tutkimuksen tietoinfrastrukturi edistää tietoaineistojen uudelleen käyttöä
- Tutkimuksen tietoinfrastrukturi kokoaa yhteen paikallisia ja keskitettyjä palveluita
- Tutkimuksen tietoinfrastrukturi on teknologialtaan hallittu kokonaisuus

# Yleiset palveluperiaatteet



- Palveluiden elinkaari ja palvelulupaus on kestävä ja hallittu (täsmentyy hallintomallityössä).
- Palveluiden ja aineistojen avoimuus, julkisuus ja tietoturva ovat tasapainoisesti hallittu kokonaisuus (täsmentyy lakiselvityksessä).
- Suomalaisessa tutkimusjärjestelmässä toimiville tutkijoille TTA-palvelujen käyttö on periaatteessa maksutonta ja hakemuksesta myönnettävää. Edellytyksenä maksuttomuuteen on, että tutkimus on julkista.
- TTA-hankkeen työryhmien ja palvelukehityksen tuloksille käytetään lisenssiä, jotta opetus- ja kulttuuriministeriö voi edistää syntyneen aineiston tehokasta käyttöä tekijänoikeuksien estämättä.

## Tähän käytetään lisenssiä, jossa käyttäjälle annetaan oikeus

- Jakaa eli kopioida, levittää, näyttää ja esittää teosta sekä
- Valmistaa muutettuja teoksia, ehdoilla että:
- Nimeä eli teoksen tekijä on ilmoitettava siten kuin lisenssin antaja on sen määrännyt (mutta ei siten että ilmoitus viittaisi lisenssinantajan tukevan lisenssinsaajaa tai teoksen käyttötapaa)
- Epäkaupallinen eli käyttäjä ei saa käyttää teosta kaupallisesti
- Jaa samoin eli jos käyttäjä tekee muutoksia tai käyttää teosta omien teostensa pohjana, tulee johdannaisteos jakaa samalla tai samankaltaisella lisenssillä.

Lisenssi on: Creative Commons Nimeä-Epäkaupallinen-JaaSamoin 3.0  
Muokkaamaton.

Lisenssi päivitetään vastaamaan CC 4.0 versioita sen ilmestyttyä.

# Tutkimuksen tietoaaineistopyramidi





# IDA tallennuspalvelu

- Yhteinen tallennuspalvelu tutkimuksen digitaalisten tietoaaineistojen säilyttämiseksi ja uudelleenkäytön lisäämiseksi
- Tietoaaineistojen ja metatiedon turvallinen säilytys
- Datan säilyminen eheänä ja muuttumattomana turvataan kopioiden ja niiden eheyden hallinnalla

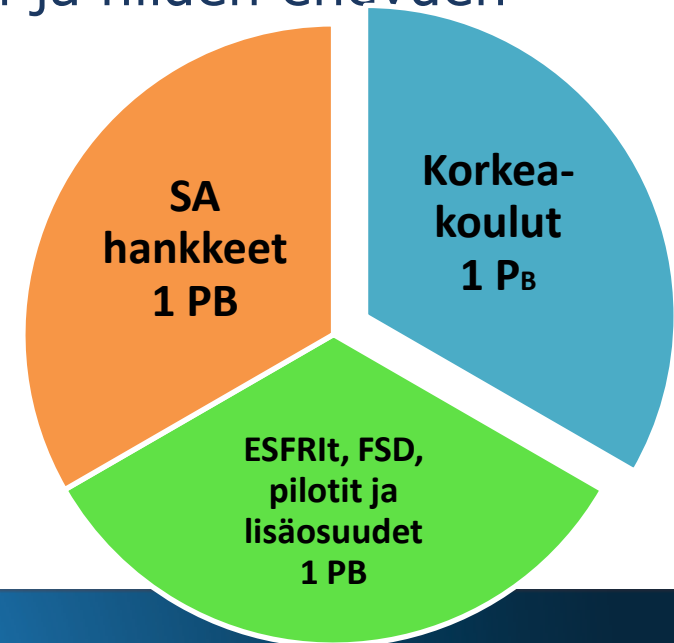


# IDA palvelulupaus

- Palvelu takaa aineistojen säilyvyyden vähintään vuoden 2017 loppuun asti. Tähän mennessä selvitetään, jatketaanko tätä säilytystapaa vai hoidetaanko säilyttäminen uudella ratkaisulla. Palvelun käyttäjän ei tarvitse tässä vaiheessa erikseen huolehtia tietojen siirrosta.
- Palvelu takaa vähintään yhteensä 3 petatavun kapasiteetin palvelussa.
- Aineistoja palveluun siirrettäessä niihin liitetään automaattisesti minimimetatietoja.
- Käyttöosuuksien mukaisesta käytöstä ei tänä aikana koidu käyttäjille kustannuksia.
- Aineistojen omistajat päättävät itse niiden avoimuudesta ja käyttöpolitiikasta. Aineistojen omistajuus- ja tekijänoikeusasioiden selvittämistä ennen aineistojen siirtoa palveluun suositellaan voimakkaasti.
- Vuoden 2017 jälkeen aineistojen säilytys vaatii minimimetatietoja laajempia kuvailutietoja.

**IDA** : tallennuspalvelu tutkimuksen digitaalisten tietoaaineistojen (jatkossa data) säilyttämiseksi ja uudelleenkäytön lisäämiseksi. Ida mahdollistaa tutkimuksen tuottamien tietoaaineistojen ja niihin liittyvän metatiedon turvallisen säilytyksen.

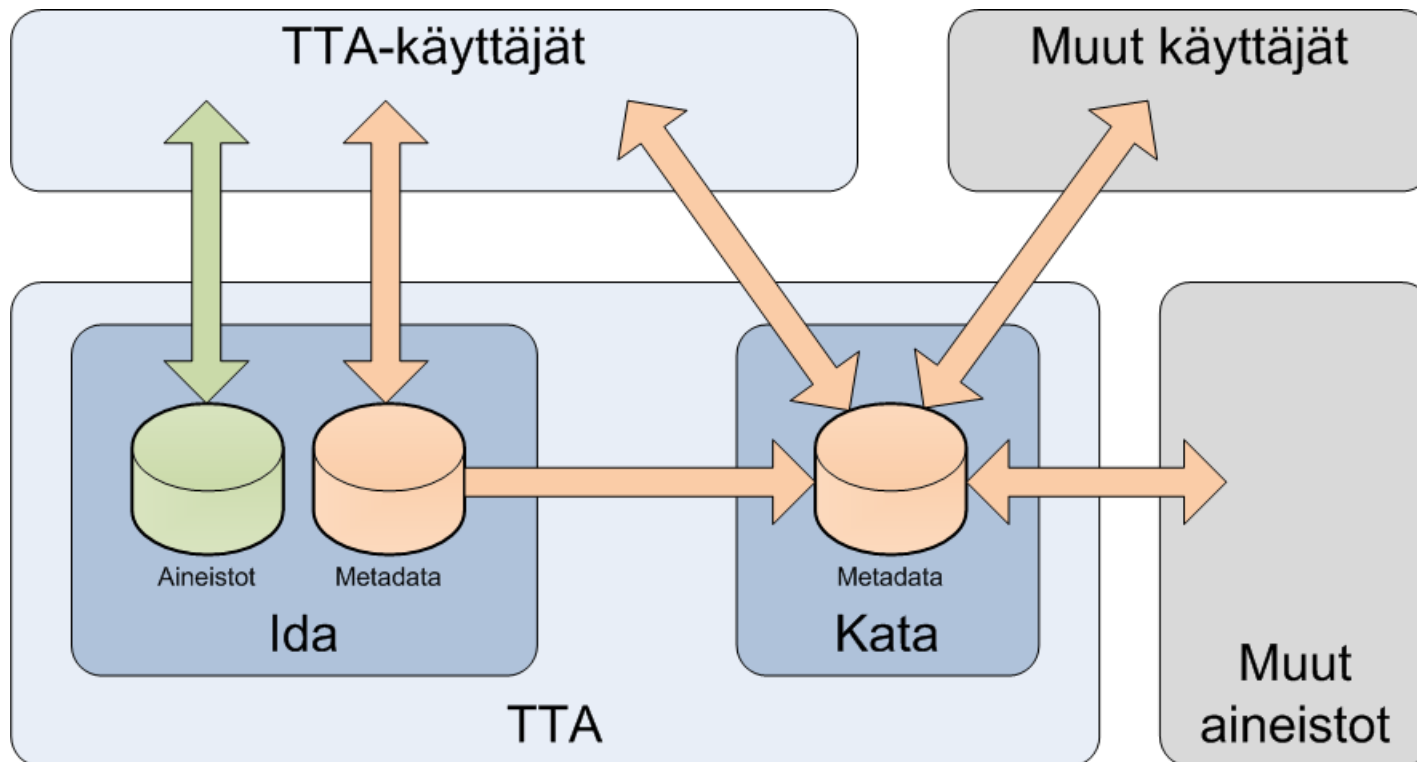
- Datat säilyminen eheänä ja muuttomattomana sekä asiakkaan helposti saatavilla on palvelun ensisijainen tavoite. Säilytettävän datan muuttumattomuus turvataan kopioiden ja niiden ehevden hallinnalla.



# KATA aineistokatalogi

- helpottaa tutkimuskäyttöön saatavilla olevien **aineistojen löytämistä**
- tuottaa tietoa aineistojen olemassaolosta **rahoittajille**
- mahdollistaa **yhtenäisen käyttöehto- ja käyttöoikeuskulttuurin** luomisen (omistajuus- ja hyödyntämisoikeustiedot aineistokatalogiin)
- mahdollistaa **meriitin syntymisen** tutkijalle
- auttaa tunnistamaan ja löytämään tietoaaineistoja **pitkäaikaissäilytykseen**

# IDA-KATA



# Digitaalinen pitkäaikaissäilyttäminen

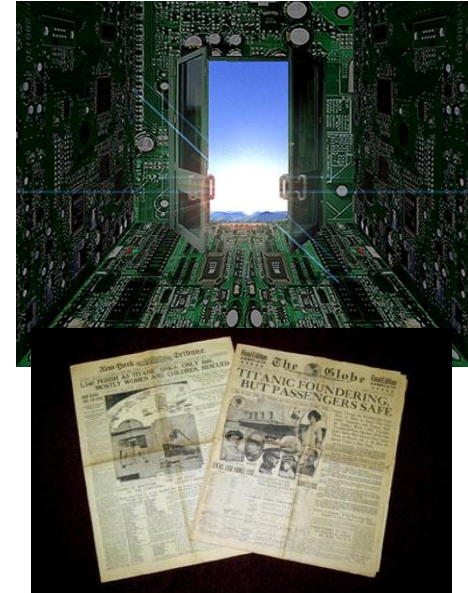


• prosessi, jossa digitaalinen kokonaisuus irroitetaan luontiympäristöstään, ja sen olemassaolo varmistetaan säilytysympäristössä autenttisuus ja eheys säilyttäen

• Kriittiset asiat:

- Autenttisuuden ja eheyden varmistaminen
- Teknologisten riskien hallitseminen
- Kustannusten hallitseminen

• Päämääränä infrastruktuuririippumattomuus, jotta voidaan käyttää mitä tahansa tallennusratkaisua



## HAASTEET:

Mitä pitää säilyttää?

- Materiaali joka pitää "pelastaa"
- Materiaali jota arvioimme tarvitsevamme pitkällä aikavälillä

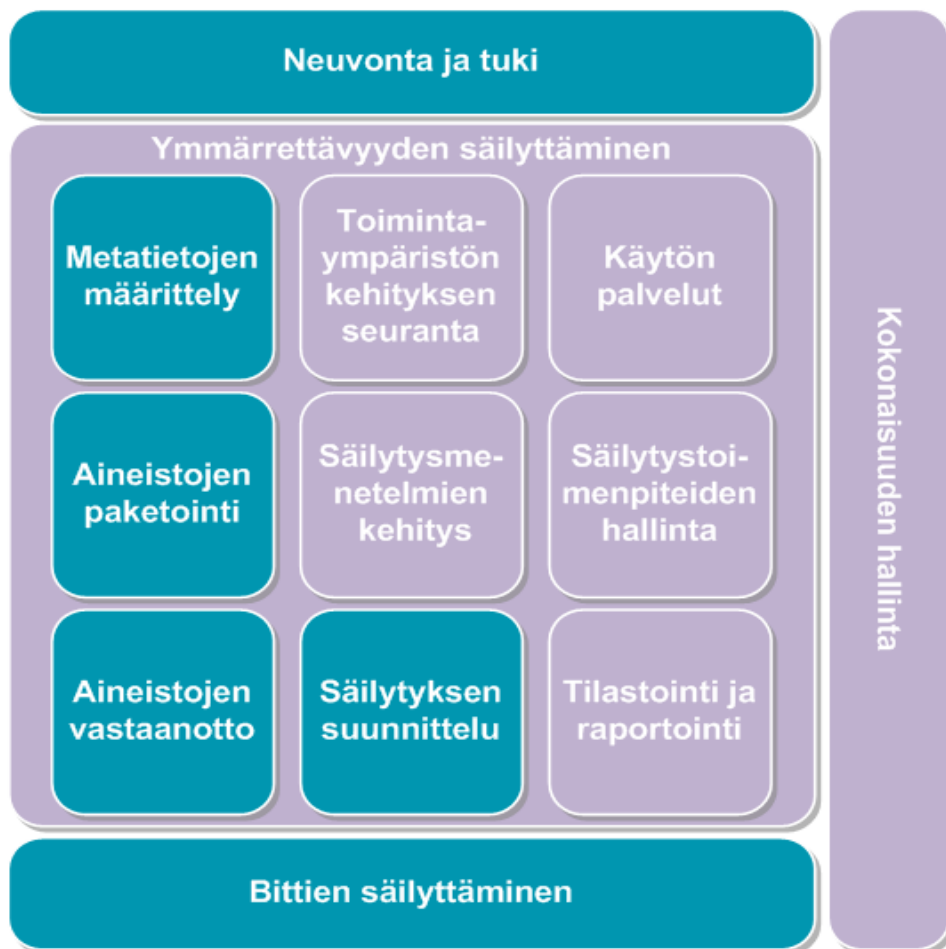
Kuinka se tulisi säilyttää?

- Formaatit
- Tallennusmediat
- Kuka on vastuussa, kuka tekee?

Kuka maksaa?

- Sisällön tuottajat?
- Laitokset ja instituutiot?
- Käyttäjät?

# PAS-ratkaisun valmistelu 2012 - 2013

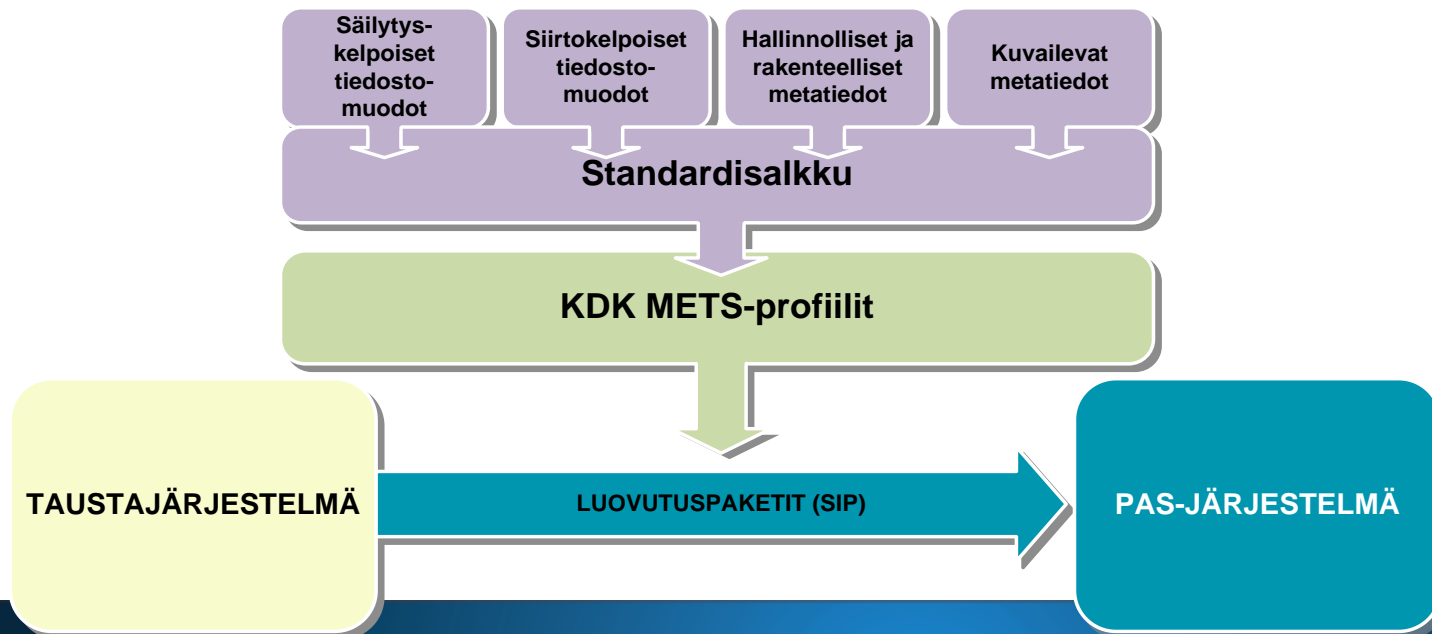


- Edetään toteuttamissuunnitelman massa ehdotetuissa vaiheissa
  - Vaihe 1: 2011-2013 PAS-ratkaisun valmistelu
  - Vaihe 2: 2014-2015 PAS-ratkaisun toteuttaminen

Vaihe I:  
PAS-ratkaisun valmistelu

Vaihe II:  
PAS-ratkaisun toteuttaminen

- Aineiston kuvailua ja rakennetta koskevilla määrityksillä varmistetaan että aineistoon on liitetty pitkäaikaisen säilyttämisen kannalta välttämättömät metatiedot
- digitaalisia objekteja ei voida siirtää PAS-järjestelmään sellaisenaan.
- Aineisto on siirrettävä luovutuspaketteina ja varustettava riittäväillä metatiedoilla.





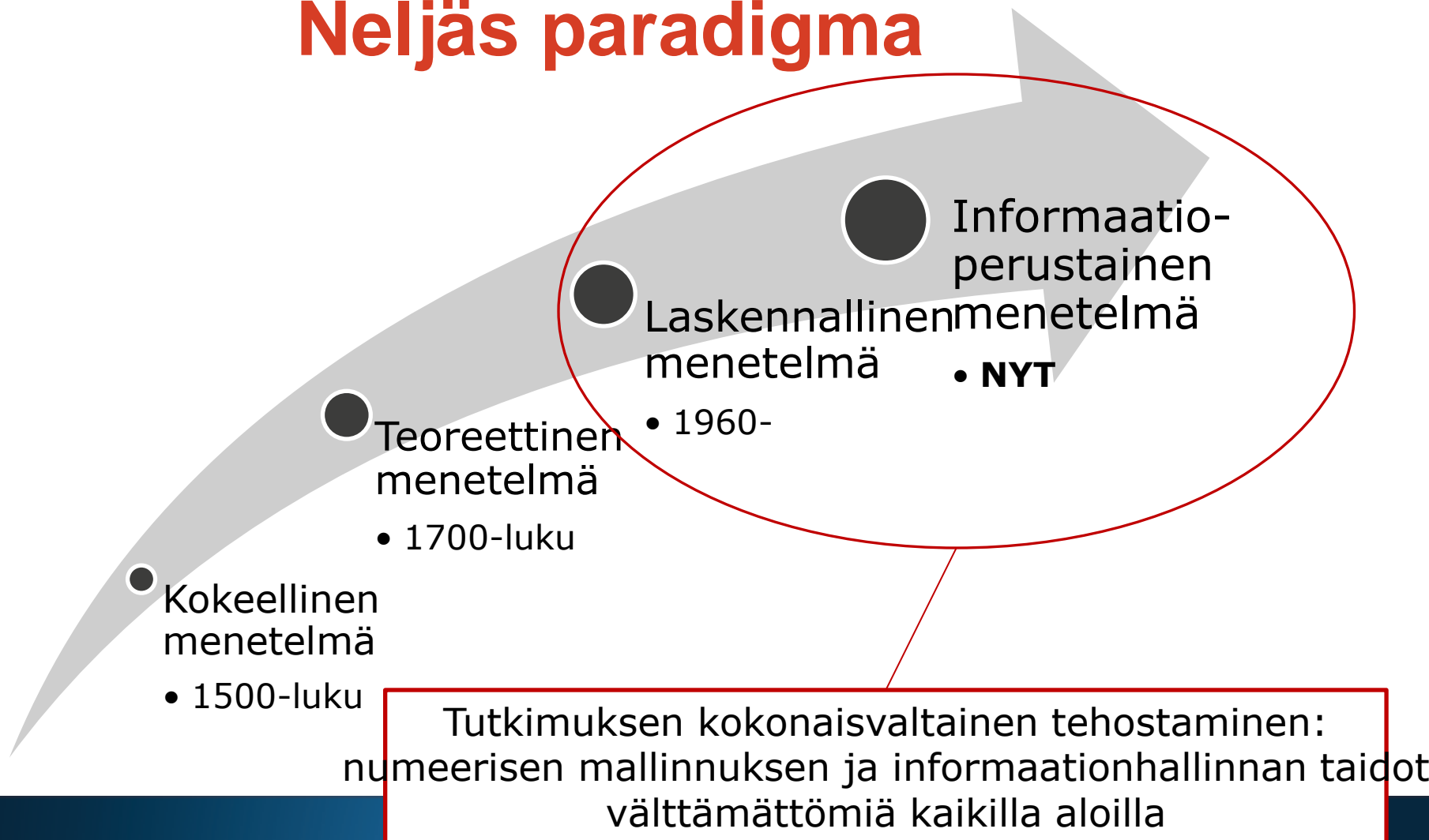


# Kiitos!





# Neljäs paradigma



# Tutkimuksen tietoaaineistopyramidi



# Tutkimustiedon infrastruktuurin

täytyy tukea:

- koko tutkimusprosessia
- tietoaineistojen siirtoa tieteenalojen välillä
- avoimen datan tuottoa
- työvuopohjaista toimintaa
- julkaisujen ja tietoaineistojen linkitystä
- tiedepolitiikan kehikkoa

