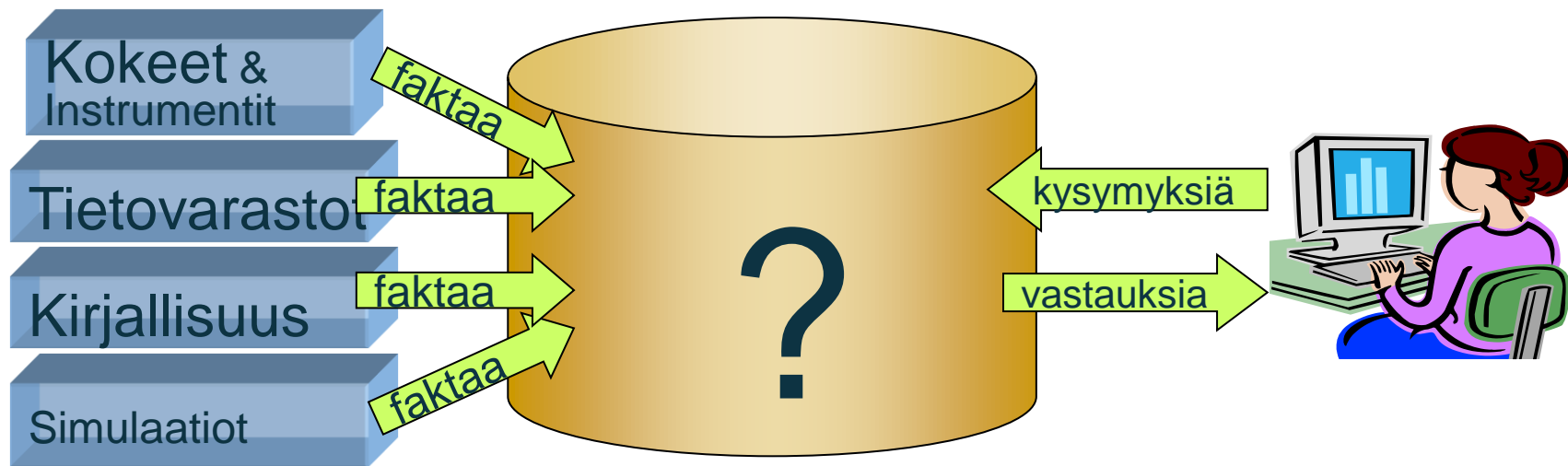




# TTA-hankkeen esittely

# eTutkijan ongelma



- Datan koostaminen
- Petatavumäärien hallinta
- Yhteiset skeemat
- Miten organisoida?
- Miten uudelleenorganisoida?
- Rinnakkaiselo ja yhteistyö muiden tutkijoiden kanssa?
- Kyselyjen ja visualisoinnin työkalut?
- Tuki ja koulutus
- Suorituskyky
  - Vastaukset kyselyihin riittävän nopeasti
  - Erittäin suurten massojen analysointi ja hallinta

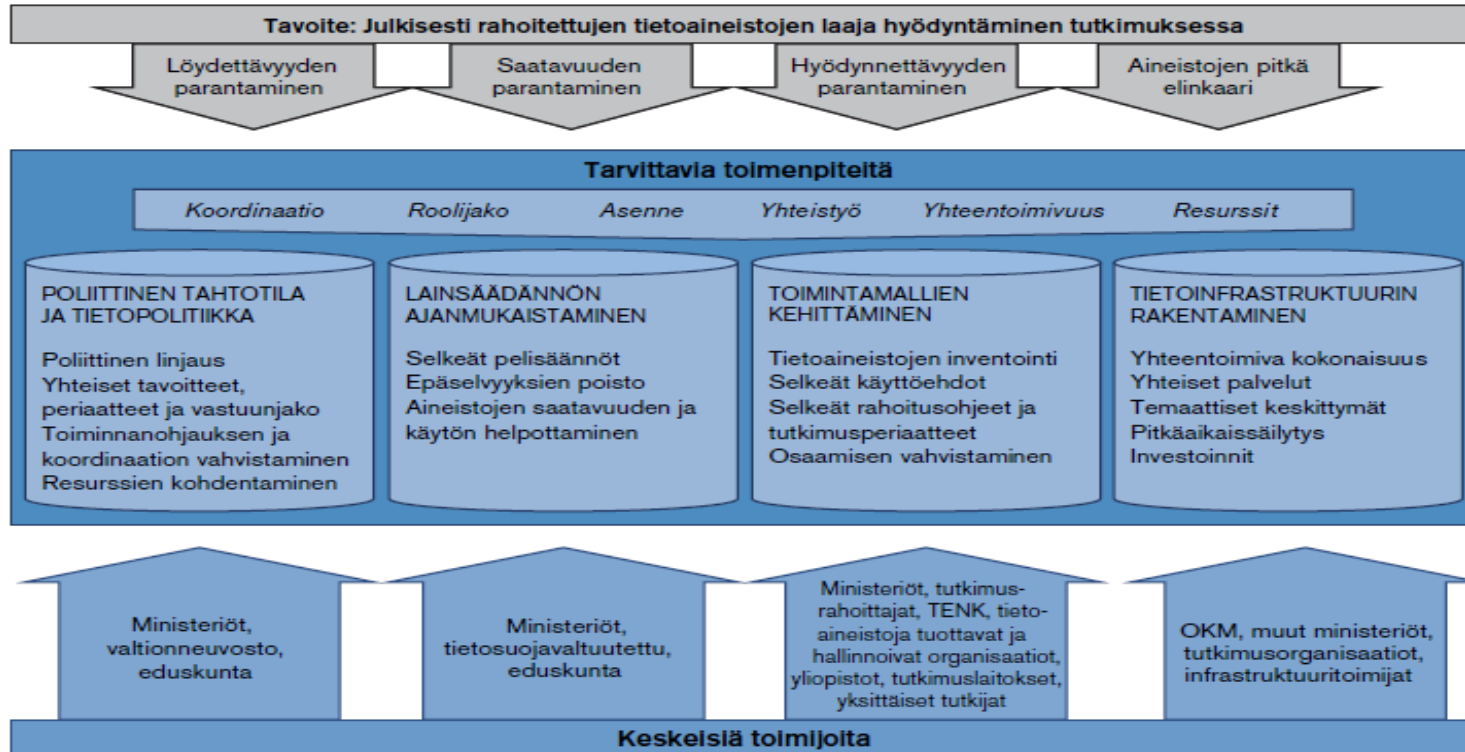
# Tutkimuksen tietoaaineistot TTA

- Opetus- ja kulttuuriministeriön käynnistämä TTA-hanke edistää tietoaaineistojen kuvausten yhtenäistämistä, säilytystä ja käyttöä.
- TTA-hankkeessa tuotetaan tätä tukemaan:
  - tutkimuksen tietoaaineistojen prosessikartta
  - tietoinfrastruktuurin palvelukokonaisuuden suunnittelu
  - metatietomalli tutkimustiedon hallinnalle
  - yhteinen metatietokatalogi/hakupalvelu tutkimusaineistoille
  - tallennuspalvelu tutkimuksen tietoaaineistoille
  - yhteisen pitkäaikaissäilytysratkaisun valmistelu
- TTA-palvelukokonaisuus edistää tutkimuksen tietoaaineistojen kansallista, eurooppalaista ja kansainvälistä yhteentoimivuutta.

# TTA-VISIO

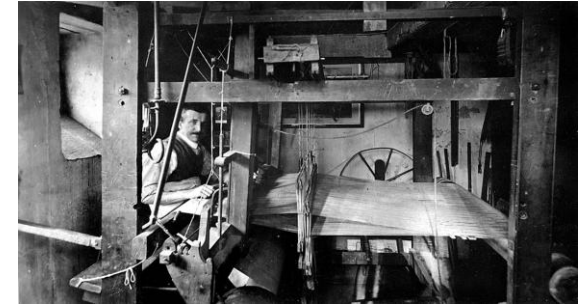
- *Suomessa on selkeä tietopolitiikka, jota yhteiset sähköiset palvelut tukevat.*
- *Julkisen sektorin hallinnoimat tietoaineistot sekä julkisen tutkimusrahoituksen tuella syntyneet tietoaineistot ovat lainsäädännön ja yhtenäisten käyttöehtojen ohjaamina pääsääntöisesti maksutta yhteiskunnan hyödynnettävissä.*
- *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.*

# Toimenpiteitä tietoaineistojen hyödyntämisen parantamiseksi



# Tutkimustiedon infrastruktuurin täytyy tukea:

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



# Tutkimusaineistojen hallinta

- Systemaattinen hallinta oleellista, jotta tieto ja sen evidenssiketju voidaan turvata.
  - Tukee tutkimuksen ydintoimintoja.
  - Vaikuttaa moniin tutkimusprosessin vaiheisiin.
  - Mahdollistaa luotettavuuden arvioinnin, ja siten tuo uskottavuutta tutkijalle, tutkimusryhmälle, organisaatiolle, rahoittajille.

# 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 löytyminen epävarmaa

Säilymistodennäköisyys kasvaa, löydettävyys kasvaa



# TTA -hankkeen rakenne

Koordinaatioryhmä

Ohjausryhmä

Työryhmät

Metatieto

Tieto-  
infrastrukturi

TURE-  
pienryhmä

PAS

Tiedon  
saatavuus

Toteuttavat  
projektit

Tutkimus-IDA

Datakatalogi

Datan  
hallinnan  
pilotti

# TTA-palvelut

- IDA -tallennuspalvelu
- KATA -aineistokatalogi 2013
- PAS -pitkäaikaissäilytysratkaisu 2015

# IDA tallennuspalvelu

- Yhteinen tallennuspalvelu tutkimuksen digitaalisten tietoaineistojen säilyttämiseksi ja uudelleenkäytön lisäämiseksi.
- Tietoaineistojen 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.

# TTA -hankkeen piiriin kuuluvat tietoaaineistot

- Suomen Akatemian tutkimusrahoituksen (akatemiahankeet, huippuyksiköt, tutkimusohjelmat) sekä tutkimusympäristörahoituksen (tutkimusinfrastruktuurit) projektien tuottamat tietoaaineistot
  - Alussa 1 PT tallennuskapasiteetti
- Soveltuvin osien korkeakoulujen, ammattikorkeakoulujen ja tutkimuslaitosten tuottamat aineistot
  - Alussa yhteensä 1 PT tallennuskapasiteetti
- Erillisrahoitetut ESFRI-hankkeet (esim. BBMRI, CLARIN)
- Yhteiskuntatieteellisen tietoarkiston aineistot
- Soveltuvin osin säätiöiden ja SHOK-yhteistyön piirissä olevat aineistot

# Käyttöoikeudet: korkeakoulut

Yliopisto	Osuus	Teratavuina
Aalto-yliopisto	0,12	120
Helsingin yliopisto	0,27	270
Itä-Suomen yliopisto	0,08	80
Jyväskylän yliopisto	0,07	70
Kuvataideakatemia	0	
Lapin yliopisto	0,01	10
Lappeenrannan teknillinen yliopisto	0,03	30
Oulun yliopisto	0,09	90
Sibelius-Akatemia	0	
Svenska handelshögskolan	0,01	10
Tampereen teknillinen yliopisto	0,06	60
Tampereen yliopisto	0,06	60
Teatterikorkeakoulu	0	
Turun yliopisto	0,14	140
Vaasan yliopisto	0,01	10
Åbo Akademi	0,04	40

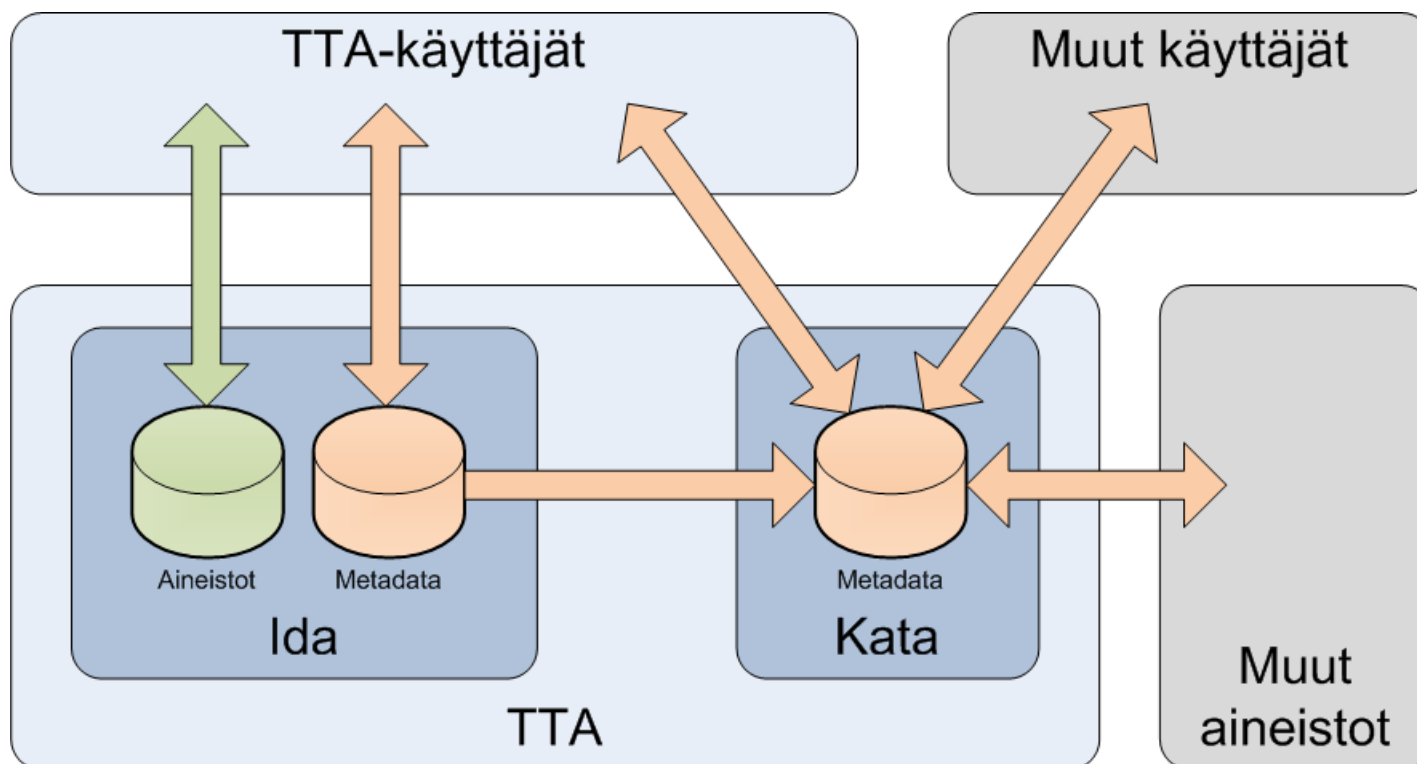
Ammattikorkeakouluille varataan 10 TT kiintiö ammattikorkeakoulujen yhteisiin tarpeisiin.

# 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 tietoaineistoja **pitkäaikaissäilytykseen**
- Edellyttää hankkeilta sitoutumista sovittaviin periaatteisiin (kuvailu, avoimuus)



# IDA-KATA -rajapinnat



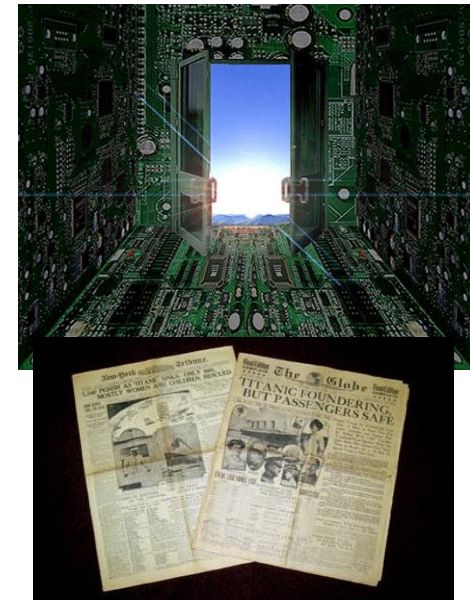


# PAS: Tutkimusaineistojen pitkäaikaissäilytys

- Tutkimustiedon **kumuloituminen** on tutkimuksen ja tutkimusorganisaatioiden toiminnan ydin.
  - Nykyisin ei ole käytössä hallittua ja toimivaa tapaa huolehtia digitaalisesta tiedosta pitkällä aikavälillä.
- Digitaalisten aineistojen pitkäaikaissäilytys tarkoittaa digitaalisen informaation **luotettavaa säilyttämistä kymmenien tai jopa satojen vuosien ajan**.
  - Laitteet, ohjelmistot ja tiedostomuodot vanhenevat, mutta tästä huolimatta informaation täytyy säilyä ymmärrettävänä.
- Tutkimusaineistojen pitkäaikaissäilytyksessä tehdään tiivistä **yhteistyötä** Kansallisen Digitaalisen Kirjaston (KDK) kanssa.

# 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.



# Pitkäaikaissäilytyksen haasteet

## 1. Mitä pitää säilyttää?

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

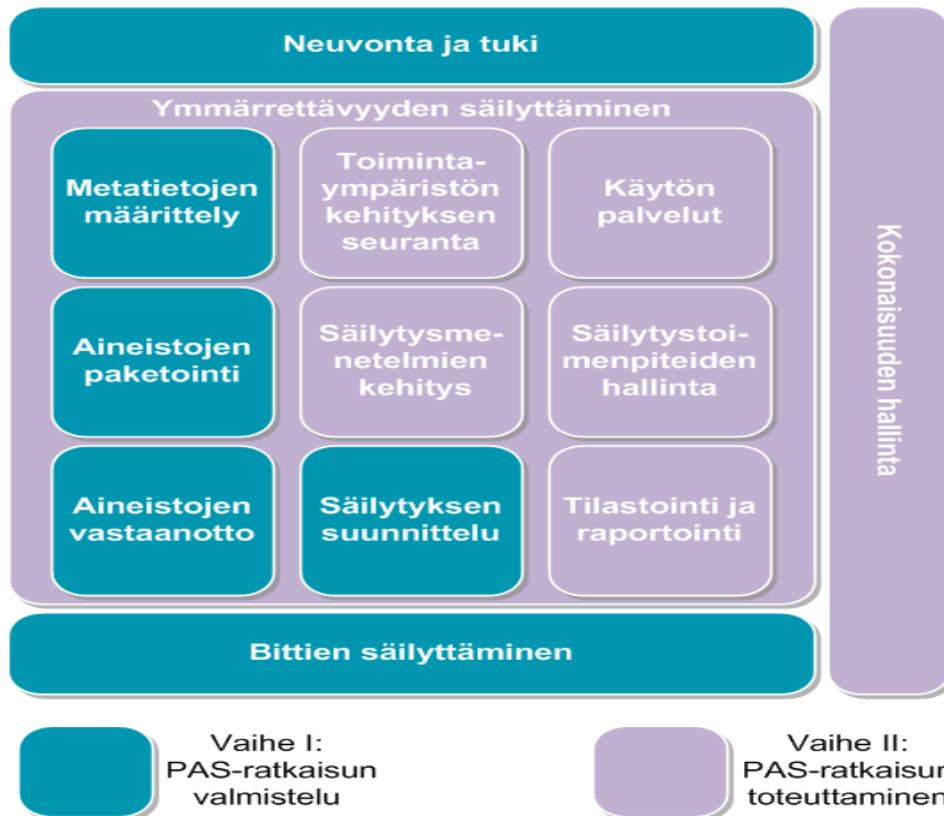
## 2. Kuinka se tulisi säilyttää?

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

## 3. Kuka maksaa?

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

# PAS-ratkaisun valmistelu 2012 - 2013



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

# Lisätietoja

- [www.csc.fi/tta](http://www.csc.fi/tta)
- [www.csc.fi/ida](http://www.csc.fi/ida)