

# Kansallinen digitaalinen kirjasto: yhteistyötä yli organisaatorajojen

Sisällön- & dokumenttien hallinta 30.-31.8.2011

Esa-Pekka Keskitalo  
Kansalliskirjasto

esa-pekka.keskitalo(AT)helsinki.fi  
[www.kansalliskirjasto.fi](http://www.kansalliskirjasto.fi) - [www.helsinki.fi/~keskital/](http://www.helsinki.fi/~keskital/)

Kansalliskirjasto

Kirjastoverkkopalvelut

## Sisältö

- Taustaa: digitaalinen säilyttäminen Kansalliskirjastossa
- Kansallinen digitaalinen kirjasto –hanke
- Yhteistyön haasteet ja hyödyt
- Keskitetyn PAS-järjestelmän kustannus- ja hyötyanalyysi
- Tuloksia ja opetuksia

Digitaalisella pitkäaikaissäilytyksellä on merkittäviä hyötyjä.

Pitkäaikaissäilytys on ratkaistavissa kustannuksiltaan vakaalla ja hallittavissa olevalla tavalla.

Yhteistyön ja nk. keskitetyn PAS-järjestelmän hyödyt ovat haittoja suuremmat.

Kansalliskirjasto

Kirjastoverkkopalvelut

## Kansalliskirjasto

[www.kansalliskirjasto.fi](http://www.kansalliskirjasto.fi)

Perinteisenä kansalliskirjastona

- kerää ja säilyttää kansallista julkaisuperintöä – painatteita, äänitteitä ja verkkoaineistoja;
- on merkittävä (etenkin humanistinen) tieteellinen kirjasto.

Digitointikeskus Mikkelissä

Palvelukeskus (Kirjastoverkkopalvelut)

- Tuottaa palveluita kirjastoille ja muille organisaatioille:
  - softapalvelut (etenkin korkeakoulukirjastoille);
  - e-aineistojen keskitetty lisensointi (FinELib);
  - Kansallisen digitaalisen kirjaston asiakasliittymä 2012;
  - standardointi, ISBN-numerot yms.
- Helsingin kaupunginkirjasto, yleisten kirjastojen keskuskirjasto, on kirjastoalan toinen merkittävä palvelukeskus.

## Kansalliskirjasto ja digitaalinen säilyttäminen

Kansalliskirjasto:

- Elektroniset vapaakappaleet: velvollisuus kerätä talteen suomalaisia verkkojulkaisuja
  - Verkkoarkistossa 190 milj. tiedostoa
  - Luovutettuja e-kirjoja, äänitteitä verkkokaupoista ym.
- Vanhojen kokoelmien digitointi
- Velvollisuus säilyttää pysyvästi = ikuisesti

Kirjastokentällä tämän lisäksi

- Esim. emo-organisaatioiden digitaaliset julkaisut
- Digitaalinen tieteellinen data ym.

## Kansallinen digitaalinen kirjasto - KDK

[www.kdk.fi](http://www.kdk.fi)

- OKM:n hanke 2008 - 2013
- Tavoittena on varmistaa kulttuurin ja tieteen digitaalisten tietovarantojen tehokas ja laadukas hallinta, jakelu ja pitkäaikaissäilytys.
  - yhteinen asiakasliittymä
  - keskitetty pitkäaikaissäilytysratkaisu
  - tietohallinnon yhteentoimivuuden edistäminen
  - digitoinnin edistäminen



Kansallinen digitaalinen kirjasto

Kansalliskirjasto

Kirjastoverkkopalvelut

## KDK:n kenttä – noin 200 organisaatiota

Kansallisarkisto

Maakunta-arkistot (7)

Kansallinen AV-arkisto  
KAVA

KOTUS, kotimaisten  
kielten tutk.keskus

Yhteiskuntatieteellinen  
tietoarkisto

Yksityiset  
keskusarkistot (9)

Yksityisarkistot

Kansalliskirjasto

Varastokirjasto

Yliopistokirjastot (20)

AMK-kirjastot (19)

Erikoiskirjastot (>100)

Yl. kirjastojen  
keskuskirjasto

Maakuntakirjastot (19)

Yleiset kirjastot (312)

Museovirasto

Kansallismuseo

Luonnontieteell.  
keskusmuseo

Valtion taidemuseo

Valtakunnalliset  
erikoismuseot (14)

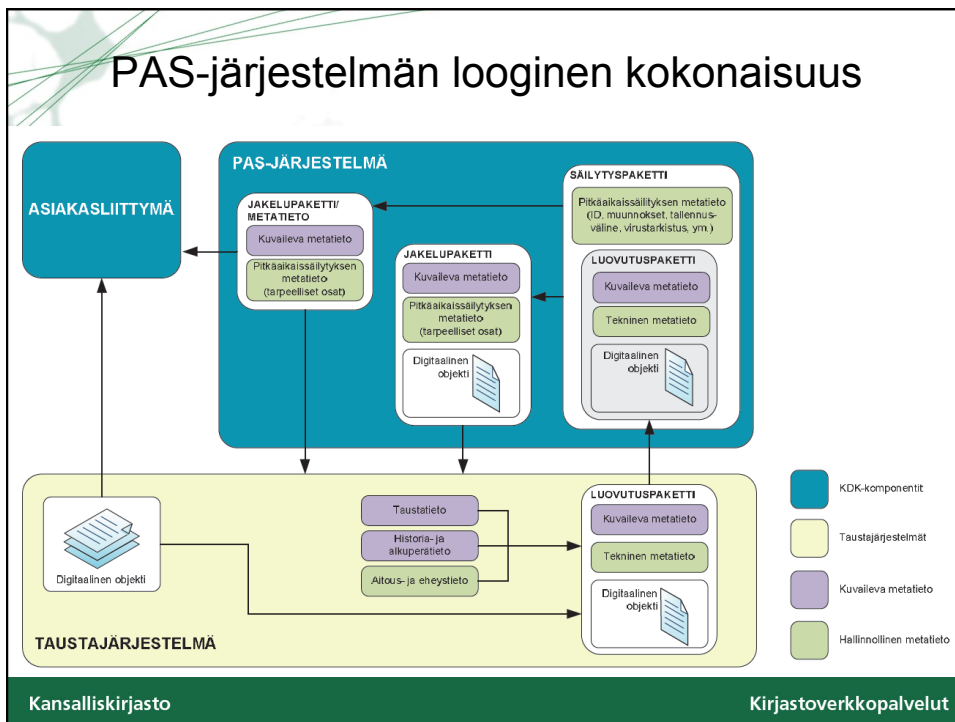
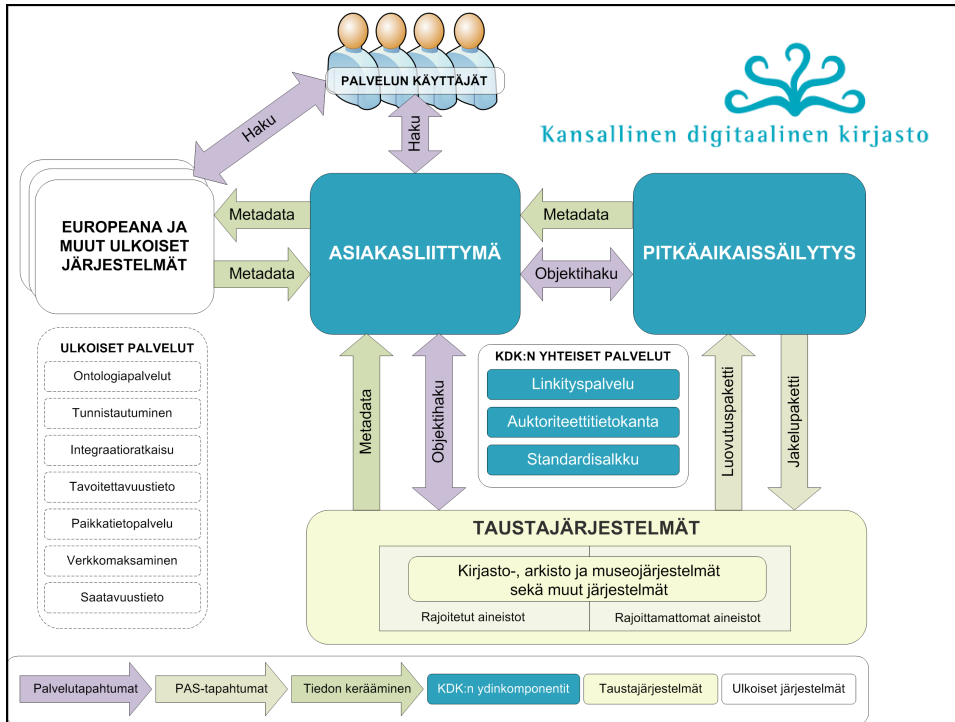
Maakuntamuseot (22)

Aluetaidemuseot (16)

Muut museot

Kansalliskirjasto

Kirjastoverkkopalvelut



## KDK ja pitkäaikaissäilytys

- 2007: OKM:n työryhmä ”Sähköisen aineiston pitkäaikaissäilytys ja käyttö”
  - K: Miten PAS pitäisi järjestää OKM:n hallinnonalalla?
  - V: ”Yhteisellä PAS-järjestelmällä voidaan tukea organisaatioiden toimintamalleja, lisätä organisaatioiden yhteisiä toimintamalleja ja edistää julkisen hallinnon tuottavuutta.”
- Toteuttaminen liitettiin osaksi KDK-hanketta 2009.
- Tällä hetkellä päävastuullisena CSC – Tieteen tietotekniikan keskus

## PAS-yhteistyön haasteet

- Erilaiset asiakkaat (tutkija – viranomainen – kansalainen)
  - Monesti joka tapauksessa yhdellä organisaatiolla erilaisia rooleja
- Erilaiset aineistot ja niiden hallinnan periaatteet
- Erilaiset organisaatiot, riippuvuudet, rahoitusmallit
- Sektorien erilaiset luonteet
  - Yhteistyön määrä (kirjastoissa paljon, museoissa lisääntymässä)
  - Sääntelyn määrä (arkistoissa suuri)
  - Järjestelmätilanne (kirjastoissa harvoja isoja, museoissa useita pieniä)
  - Omat terminologiat
- Vertikaalinen vs horisontaalinen nojautuminen (onko yleinen kirjasto osa kirjastoverkkoa vai osa kuntaa?)

## Yhteistyötä puoltavat tekijät

- Yhteiset asiakkaat
- PAS:n periaatteet ovat pitkälti samoja aineistosta riippumatta
- Avautumisen tarve, olipa KDK:ta tai ei
  - KDK:n edellyttämät uudistukset lisäävät aineistojen näkyvyyttä ja käytettävyyttä myös muissa konteksteissa.
- Teknologian mahdollisuudet
  - Esim. museoissa siirrytään yhteiseen LIDO-vaihtoformaattiin
- Valtion ohjaus (ValTIT, SADE, tietohallintolaki)
- Kansainvälinen trendi

## PAS:n kustannus- ja hyötyanalyysi

- PAS-järjestelmä = hallinnon ja rahoituksen malli, joka varmistaa aineistojen säilymisen käytettävissä
- Toteutuksen tulee tasapainottaa kustannukset ja hyödyt pitkällä aikavälillä
- Elinkaarikustannusten ymmärtäminen on välttämätöntä
  - Asetettava kriteerit, taustaoletukset, rajaukset
  - Tunnistetaan kustannustekijät; lasketaan kustannukset; vertaillaan vaihtoehtoja
  - Aloitusvaiheen kustannukset; käyttökustannukset; ylläpitokustannukset

## Käytetyt työkalut

- LIFE-elinkaarimalli (Lifecycle Information For E-Literature)
  - Digitaalisen aineiston elinkaaren vaiheet
  - Tavoitteena mahdollisen täydellinen tietämys toiminnon koko elinkaaren kustannusvaikutuksista (Life Cycle Costs)
- Kustannushyötyanalyysi
  - Vaihtoehtojen vertailu, hyötyjen ja haittojen tunnistaminen
- Kansalaishyödyt: Standard Cost Model (SADE-hankkeesta)

Figure 4 - The LIFE Model v2

Lifecycle Stage	Creation or Purchase	Acquisition	Ingest	Bit-stream Preservation	Content Preservation	Access
Lifecycle Elements	....	Selection	Quality Assurance	Repository Administration	Preservation Watch	Access Provision
	....	Submission Agreement	Metadata	Storage Provision	Preservation Planning	Access Control
	....	IPR & Licensing	Deposit	Refreshment	Preservation Action	User Support
	....	Ordering & Invoicing	Holdings Update	Backup	Re-ingest	
	....	Obtaining	Reference Linking	Inspection	Disposal	
	....	Check-in				

Kuva 4: LIFE-projektin laatima sähköisen aineiston elinkaarimalli, versio 2

## Taustaselvityksiä

- Preservation and Use of Digital Materials. Memorandum of the Working Group on Long-Term Preservation and Usability of Digital Materials. Ministry of Education: Reports 2008:2.
- "Identifying Benefits of the LTP Project and Increasing Organizations' Preparedness", December 2009. Jaakko Asplund (Salivirta & Partners)
- Digital Materials in Archives, Libraries, and Museums. December 2009. Juhani Koivunen (Netum Konsultointi)

## Taustaoletuksia: datan määrä

Aineisto	Määrä, teratavuja			
	2008	2009	2010	2011
Tekstidokumentit	64	164	190	216
Kuvat	6	10	17	25
Liikkuva kuva	11	25	31	37
Ääni	17	23	28	33
Viitetietueet	<10	<10	<10	<10
Verkkoarkisto	8	17	26	35
Radio- ja TV-arkisto	2	57	112	167
Yhteensä	108	296	404	513

**Arviot ovat osoittautuneet liian pieniksi!**



## Taustaoletuksia: kasvuvauhti

Vuosi	Organisaatioita	Järjestelmiä	Aineistoa, Tt
1			
2			
3	5	5	300
4	20	10	700
5	80	20	1,000
6	140	30	1,400
7 - 12	209	40	4,000

## Oletuksia: palvelun rajaus

- Varauduttava erilaisiin organisaatioihin ja aineistoihin
- Organisaatio vastaa aineistojen valmistelusta säilytyskuntoon
  - Standardien mukaisuus
  - Metadatan vähimmäisvaatimukset
  - METS-paketointi
  - PAS-järjestelmä voi tarjota neuvoja ja palveluita (esim. METS-paketoinsofta)
- Organisaatioiden rooli voi vaihdella (asiakas – aktiivinen kehittäjä)
- Säilytettävän aineiston jakelu viime kädessä organisaation asia.

## Oletuksia: toteutus

- Aineistosta useita kopioita, eri medioilla, eri paikoissa
  - esim. kopio levyllä (S-ATA), kaksi kopiota nauhoilla (varmistuskopio, arkistokopio)
- Kaksi palvelu- ja säilytys pistettä eri paikkakunnilla; sekä kolmannessa paikassa arkisto
- Arvio: hintojen lasku ja tallennustilan kasvu kumoavat toisensa = tallennuksen kustannukset pysyvät suunnilleen samoina

## Oletuksia: tekniikan elinkaaret

Levytilan lisäys	1
Levykehikot	3
Palvelimet	3 - 5
Nauharobotti	5 - 8
Nauhurit	3 - 5
Tietoliikenne	3 - 5
Hallinta- ja valvontaohjelmistot	3 - 8
PAS-ohjelmisto	5 - 10
Formaattirippuvat ohjelmistot	3 - 5

## Oletuksia: henkilökulut

- Arvioitiin johdon, asiantuntijoiden ja tukipalveluiden keskipalkat
- Arvioitiin vuosikasvu
- Työnantajan kustannukset (30 % palkasta)
- Tilakustannukset (750 € / kk / hlö)
- 10 asiantuntijaa kohti 1 johtaja ja 0,5 htv tukipalvelua
- Palkkasumman 1/10 ulkoistettuun koulutukseen, auditointiin yms.

## Kustannustekijät: henkilötyö

Creation	Acquisition, selection, cataloguing	Ingest	Bit level preservation	Content preservation	Access
Digitization	Selection	Enrich metadata	Systems maintenance	Monitoring	Dissemination
R&D	Acquisition	QA	New media	Operations	Access ctrl
	Cataloguing	Transfer	Backups	Error mgmt	Statistics
	Management	Systems integration	Monitoring	R&D	Services
	Systems maintenance	R&D	Security		R&D
	DRM		R&D		
	Contracts				
R&D					

## Kustannustekijät: investoinnit

Creation	Acquisition, selection, cataloguing	Ingest	Bit level preservation	Content preservation	Access
Creation systems' - HW - SW - Building phase	Background systems' - SW - HW - Building phase		LTP system's - HW - SW - Building phase	SW	User interface's - HW - SW - Building phase

Kansalliskirjasto Kirjastoverkkopalvelut

## Kustannustekijät: kiinteät ja muuttuvat kustannukset

Creation	Acquisition, selection, cataloguing	Ingest	Bit level preservation	Content preservation	Access
- Insurance - Rents - Basic fees for energy, broadband			As in Creation		As in Creation
- Licences - Media - Leasing - Energy - Broadband - Outsourced services		Outsourced services	As in Creation	Outsourced services	As in Creation

Kansalliskirjasto Kirjastoverkkopalvelut

## PAS-järjestelmän pääpiireet

- 200 palveluja käyttävää organisaatiota (monet pieniä).
- Alussa 700 Tt, kasvaa 4 000 teratavuun 7 vuodessa.
- Suunnittelu ja käyttöönotto vie 6 vuotta.
- 2 pääasiallista sijoituspaikkaa + 3. säilytyspaikka.
- Aineistosta useita kopioita, eri medioilla, eri paikoissa
- Suorittaa riskien hallintaa ja säilyttämistoimenpiteitä
- Conducts risk management and preservation operations
- 10 % aineistoista muunnetaan vuosittain
- Asiakkaalla aineistojen omistus ja viimesijainen kontrolli
- Erilaisia palveumalleja eri organisaatiotyypeille

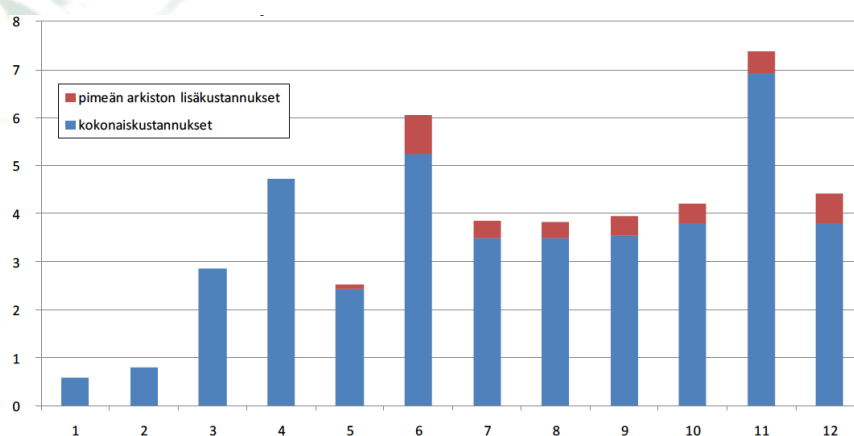
## Tunnistetut laadulliset hyödyt

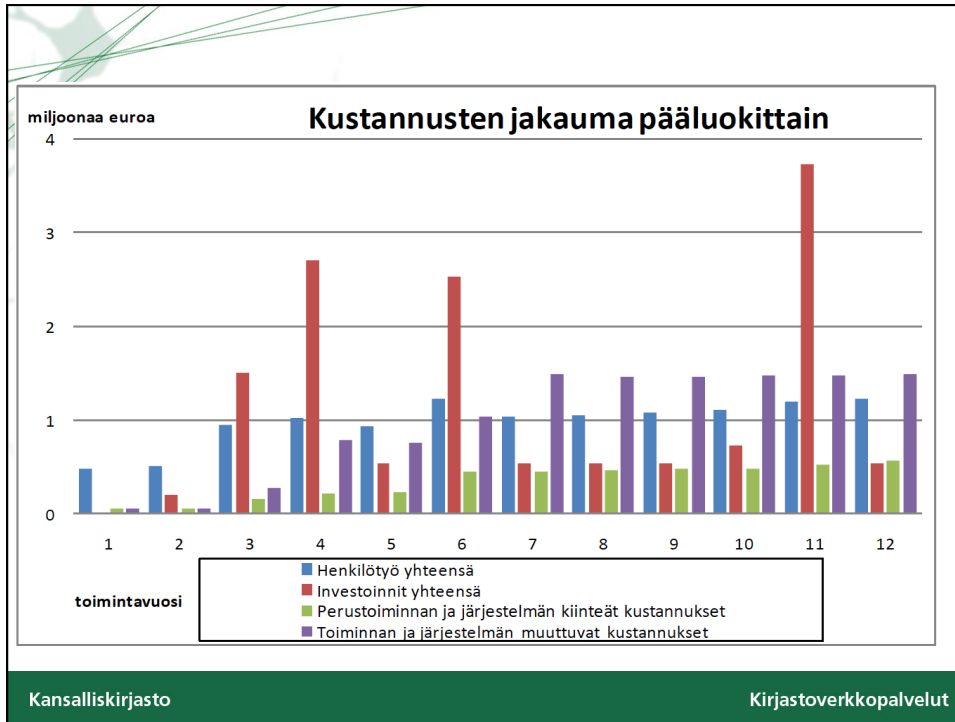
- Organisaatiot voivat täyttää lain velvoitteet
- Poistaa päällekkäisi toimintoja
- Yhteistyö ja resurssien yhteiskäyttö yli organisaatorajojen
- Osaamisen kehittäminen henkilöriippumattomasti
- Keskittyminen ydinosaamiseen
- Paremmat prosessit ja palvelut
- Aineistot hyödynnettävissä pitkällä aikavälillä

## Kustannusarvio

- 42 M€ vuosina 1-12 (sisältää suunnittelun ja käyttöönoton)
- Plus 3,4 M€, jos toteutetaan ”pimeä arkisto”
- Vuoden 12 jälkeen vuosittaiset kustannukset 4 M€

## Kustannukset vuosina 1 - 12





### PAS-järjestelmän laadulliset hyödyt

Hyöty	Hyötyindeksi (/100)
Organisaatio voi täyttää lakisääteiset velvollisuutensa.	81
Organisaatio voi keskittyä ydintoimintoihin.	81
Päällekkäisyydet vähenevät.	63
Osaaminen ja sen säilyminen turvataan.	63
Aineistojen säilyminen varmistuu.	63
Paremmat prosessit ja palvelut	63
Mahdollistaa tuleville sukupolville aineiston hyödyntämisen.	63
Yhteistyö ja yhteiskäyttö yli organisaatorajojen	38

Kansalliskirjasto Kirjastoverkkopalvelut

## Määrälliset hyödyt



### Tärkeimmät hyötytekijät

1. Alhaisemmat järjestelmien kustannukset
2. Aineistojen halvempi hallinta
3. Säilyttämisen parempi onnistuminen

## Alhaisemmat järjestelmien kustannukset

- Suunnittelu ja käyttöönotto: 30 miljoonaa €
- Tuotantovaihe: 8 miljoonaa € / v
- Verrattuna tilanteeseen, jossa 5 PAS-järjestelmää
- Säästöt syntyvät henkilökuluissa ja investoinneissa



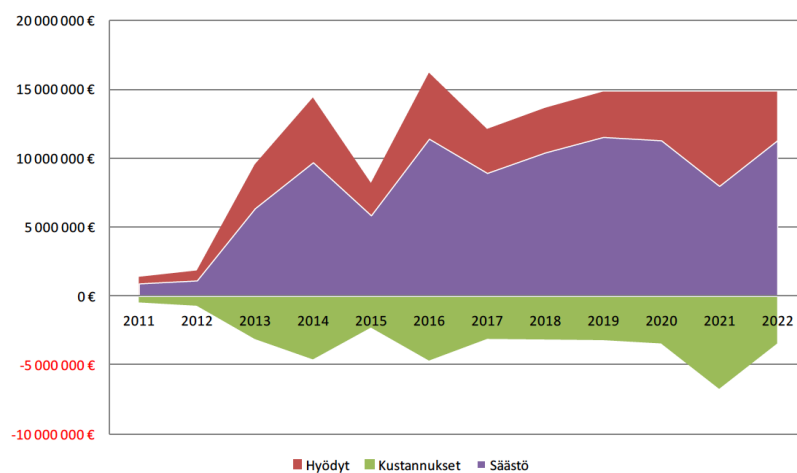
## Aineistojen halvempi hallinta

- Tuotantovaiheessa 2,5 miljoonaa € / v
- Verrattuna tilanteeseen, jossa aineistojen hallinnointi, laadunhallinta ja käyttöön asettaminen tehtäisiin organisaatioissa

## Säilyttämisen parempi onnistuminen

- Digitoidut aineistot
  - 0,5 milj. € / v
  - Vertailutilanne: 2 % digitoidusta aineistosta vaurioituu ilman PAS:ia
  - Digitointikustannukseksi laskettu: 1,3 € / objekti
- Syntyjään digitaaliset aineistot
  - ”Miljoonia euroja vuodessa”

## PAS: hyötytase



Kansalliskirjasto

Kirjastoverkkopalvelut

## Hyötyjen ja kustannusten yhteenveto

- Kustannukset:
  - 42 milj. € vuosina 1-12 (sis. Suunnittelun ja käyttöönoton).
  - Plus 3,4 mij. €, jos toteutetaan ”pimeä arkisto”.
  - Sen jälkeen juoksevat kulut vakaasti 4 milj. € / v.
- Yhteisen järjestelmän rahalliset hyödyt
  - 30 milj. € vuosina 1-12
  - Sen jälkeen 10 milj. € / v.
- Kumulatiivinen hyöty vuosina 1 -12: 100 milj. €

Kansalliskirjasto

Kirjastoverkkopalvelut

## Kustannus- ja hyötyanalyysin tulokset

- Kulujen kasvu ajan myötä on siedettävä
- Tutkitussa skenaariossa kulut ovat hallittavissa
- PAS:sta on hyötyjä ja ne voi ilmaista rahassa
- Yhteinen PAS-järjestelmä on varteenotettava vaihtoehto

## Prosessin aikana opittua

- On olemassa menetelmiä, joita voidaan soveltaa pitkäaikaissäilytyksen kustannusten ja hyötyjen analyysiin.
- Vastaavia analyyseja, joiden tuloksista saisi tietoja, on kuitenkin tehty varsin vähän.
- Epävarmat luvut: parempi kuin ei lainkaan lukuja
- Arvot voi ja ne täytyy ilmaista rahassa.