



# **Dataklinikka**

## **Eli kuinka hyödyntää Kansalliskirjaston tuottamaa dataa tutkimuksessa**

Jussi-Pekka Hakkarainen  
Tietoasiantuntija  
Erikoiskokoelmat  
Tutkimuskirjasto

**12.02.2019**  
**Oulun yliopisto**

# Dataklinikan tavoitteet ja aiheet

- 1) LUODAAN YLEISKUVA SIITÄ, MILLAISIA AINEISTOJA KANSALLISKIRJASTO TARJOAA JA MISTÄ NE LÖYTÄÄ
- 2) MITEN DATAA TULISI TARKASTELLA JA MILLAISILLA VÄLINEILLÄ PÄÄSEE ETEENPÄIN?

## Ensimmäinen osio

- Digitaaliset ihmistieteet ja Kansalliskirjasto
- Verkkokokoelmat
- Avoin data

## Toinen osio

- Case: Sanomalehtidata NewsEye-hankkeessa
- Tutustuminen Kansalliskirjaston digitoituun lehtiaineistoon harjoitusten avulla

# Digitaalisten ihmistieteiden politiikka

- Digitaalisilla ihmistieteillä tarkoitetaan 1) laskennallisten metodien hyödyntämistä humanistisessa ja yhteiskuntatieteellisessä tutkimuksessa sekä 2) digitaalisuuden ja digitaalisen kulttuurin tutkimusta.
- Kansalliskirjaston kontekstissa *digitaaliset ihmistieteet* tarkoittaa tutkimukseen soveltuvien aineistojen, asiantuntijapalveluiden, infrastruktuurien kehittämistä ja tarjoamista yhteistyössä tiedeyhteisön ja muiden toimijoiden kanssa.
- Kansalliskirjaston DH-politiikka ja [teesit](#)

# Digitaalisten ihmistieteiden politiikka

- DH-politiikan jalkauttamista Kansalliskirjaston omaan toimintaan.
- Politiikassa mainittuja tavoitteita:
  - Avataan tekijänoikeuksien alaisia ja tekijänoikeuksista vapaita digitoituja teksti- ja datamassoja konelukuisessa muodossa.
  - Tarjotaan / avataan rajapintoja ja testataan niitä yhteistyössä tutkijoiden kanssa (Finna,Digi, Doria, metatietovarannot).
  - Avataan verkkoarkisto ja muut e-vapaakappaleet soveltuvasti tutkijakäyttöön.
  - Edistetään verkkoaineistojen keräämistä, arkistointia ja tutkimuskäyttöä koskevaa julkista keskustelua.
  - Tarvittaessa hankitaan ulkopuolisten tuottamaa dataa tutkimusaineistoksi.

# Joitain viimeaikaisia kehityslinjoja

- Digitalia (2015 – 2019)
- Kansalliskirjaston DH-politiikka (2016)
- [digi.kansalliskirjasto.fi](http://digi.kansalliskirjasto.fi):n kehittäminen (2016-)
- Avoin data / datakatalogi (2017)
- Verkkoarkistoinnin ja -haravoinnin hyödyntäminen tutkimuksessa (2018-)
- Massadigitoinnista kohti tutkijalähtöisempää digitointia (2019-)
- Tutkimuksen palvelujen harmonisointi (2019-)

# Digitaalinen ketju



# Valinta – Mitä digitoidaan?



## ■ Perusrahoituksella 2019

- Uudet karttavat sanomalehdet
- Sanomalehdet 1930-1940-luvut
- Aikakauslehdet 1930-1940-luvut
- Pienpainatteista puutarhakuvastot
- Avioliitto- ja seksuaalivalistusoppaat, eli nk. Siveellisyys-signum

## ■ Täydentävällä rahoituksella 2019-2020 (riippuen rahoituksesta)

- Ruotsin vallan ajan kansalliskokoelma loppuun (8000 teosta, Ruotsin kansalliskirjaston kanssa)
- Monrepos'n kartanon arkisto (Venäjän Kansalliskirjaston kanssa)
- SLS:n Skrifter-sarja
- Uusi Suomi

# MF/Digitoinnin valmistelu/ Konservointi



Analogisena saapuva aineisto

- Mikrokuvaus kirjastojen käyttöön
- Digitoinnin valmistelu

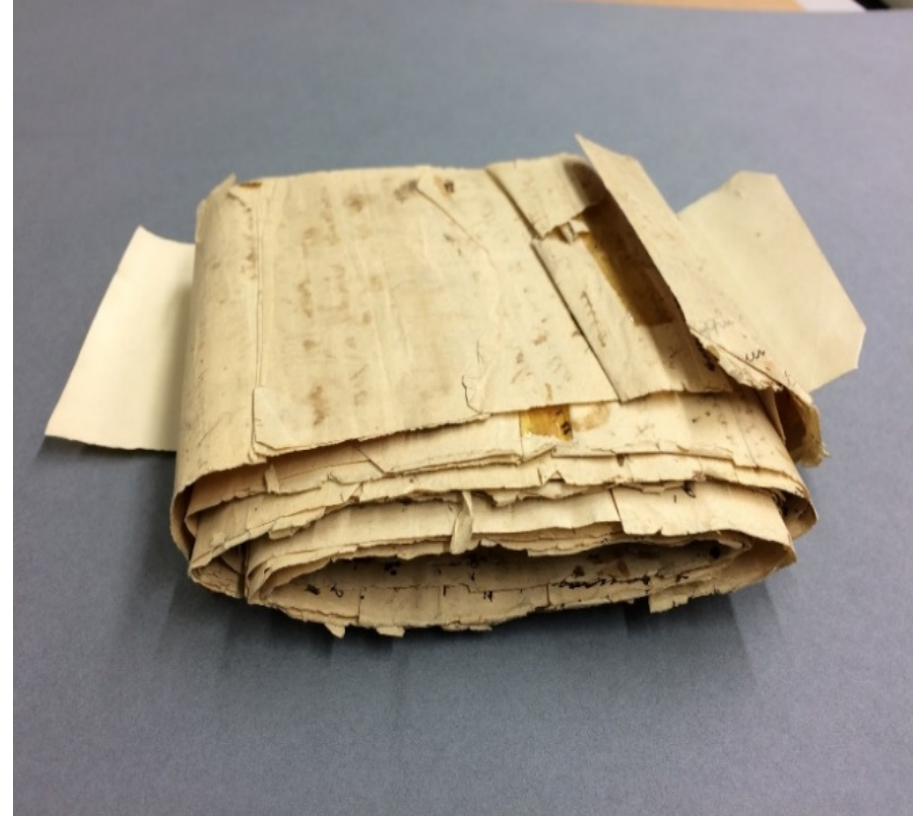
Sähköisesti luovutettu aineisto

- Aineiston tarkistus
- Tulostus mikrofilmille
- Tuonti esitysjärjestelmään



# Maiju Lassilan Liika viisas

- Käsikirjoituskäärö liittyy Maiju Lassilan vuonna 1915 julkaistun Liika Viisas-teoksen käsin kirjoitettuihin lisäyksiin.
- Alkuperäinen teos sekä siihen liittyvät käsin kirjoitetut lisäykset lahjoitettiin Kansalliskirjastolle vuonna 2015, irrallinen käärö vuonna 2005.
- Liika Viisas-teos lisäyksineen konservoitiin vuonna 2016 ja käärö vuonna 2018.



# Maiju Lassilan Liika viisas

- Konservoinnin aikana käärö paljastui 12 metriä pitkäksi yhtenäiseksi, toisiinsa liimatuista paloista kootuksi liuskaksi.
- Käärö on kirjoitettu hiokepaperille, joka on hyvin kellertävää, hapanta, haurasta helposti kosketettaessa ja käsiteltäessä vaurioituvaa.
- Muotonsa vuoksi herkkä lisävaurioille.
- Suositellaan käärön digitointia, jonka jälkeen alkuperäisen kohteen käyttö nollaantuu ja käärön ulkoasun sekä tekstin sisältämän informaation säilyminen taataan.

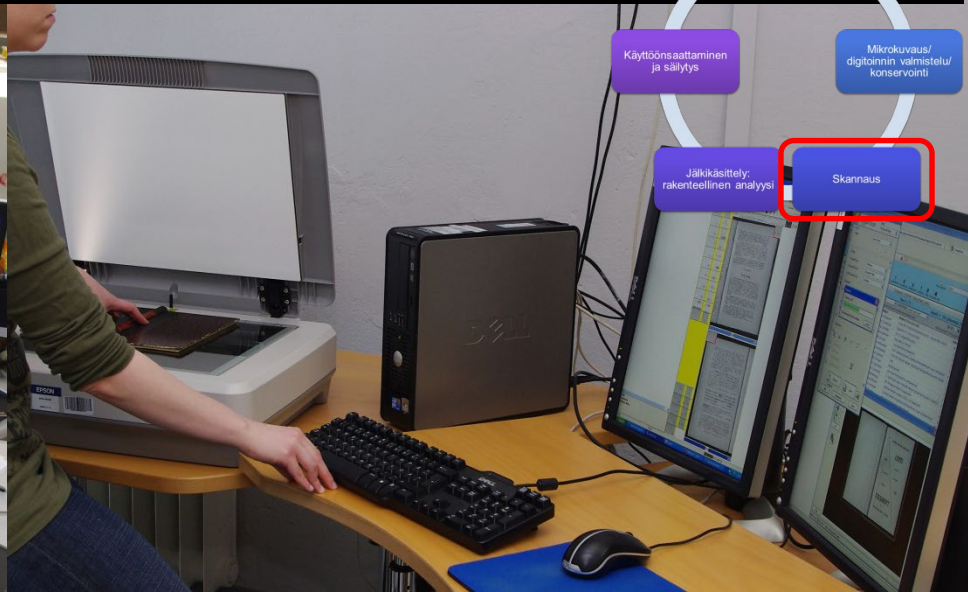


# Skannaus



- Aineistoon ja toivottuun laatutasoon sopiva skannerit
- Automaattiskanneri, jos aineisto soveltuu
- Erikoislaitteet

# Skannaus – aineistot, laitteet, osaajat



# Jälkikäsitteily

- Skannatut sivukuvat käsitellään DocWorks ohjelman kautta
  - Tehdään tekstintunnistus (ALTO XML)
  - Alueiden tunnistus (txt,kuva,otsikko...)
  - jne
- Kulminoituu arkistopakettiin, jossa METS XML kuvaa paketin sisällön:
  - Näyttökuvat, säilytyskuvat, METS/ALTOt ja PDF niteestä.



# Jälkikäsittely

Import > Cropping > Zoning > 4 Structure > 5 Output


Text List Image Text Properties

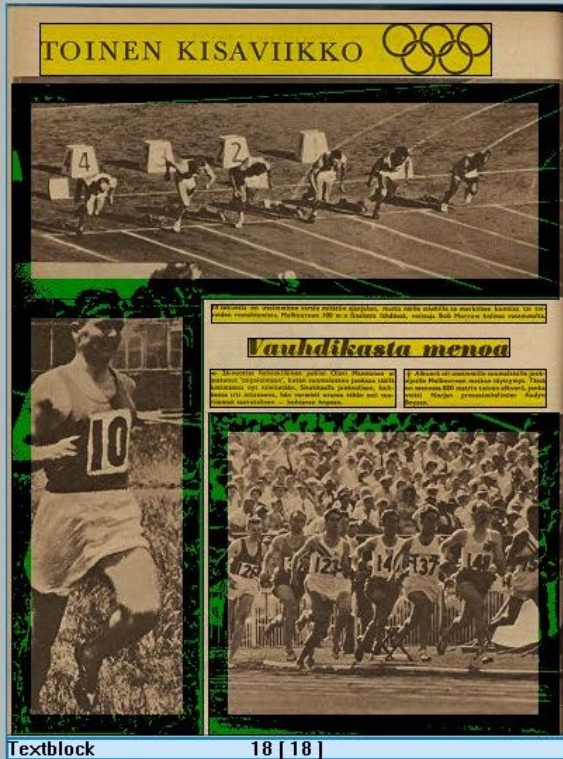
doc

docContent: Hierarchy

Suomen Kuvalehti		
[ 1 ]: Front	0	
[ 2 ]: Issue	49	08.12.1956 40
[ 1 ]: Front	3	1 [1]
[ 1 ]: CoverPage	1 [1]	
[ 2 ]: imagePage	2 [2]	
[ 3 ]: imagePage	3 [5]	
[ 2 ]: Main	60	4 [4]
[ 11 ]: imagePage	14 [14]	
[ 12 ]: imagePage	15 [15]	
[ 13 ]: imagePage	16 [16]	
[ 14 ]: TitlePage	17 [17]	
[ 15 ]: imagePage	18 [18]	
[ 16 ]: imagePage	19 [19]	
[ 17 ]: imagePage	20 [20]	
[ 18 ]: imagePage	21 [21]	
[ 19 ]: imagePage	22 [22]	
[ 20 ]: imagePage	23 [23]	
[ 3 ]: Back	1	64 [64]
[ 1 ]: imagePage	64 [64]	
[ 3 ]: Issue	50	15.12.1956 40
[ 1 ]: Front	4	65 [1]
[ 2 ]: Main	51	69 [5]
[ 3 ]: Back	1	120 [56]
[ 4 ]: Issue	51-52	22.12.1956 40
[ 1 ]: Front	10	121 [1]
[ 2 ]: Main	53	131 [11]
[ 3 ]: Back	1	184 [64]
[ 5 ]: Back	0	

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32


TOINEN KISAVIIKKO 



**Vauhdikasta menoa**

18 [18]

Textblock

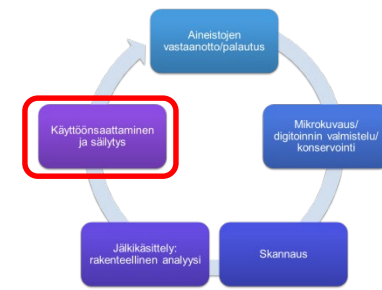


**Yksi kulta ja pronssia**

19 [19]

Illustration

# Käyttöönsaattaminen ja säilytys

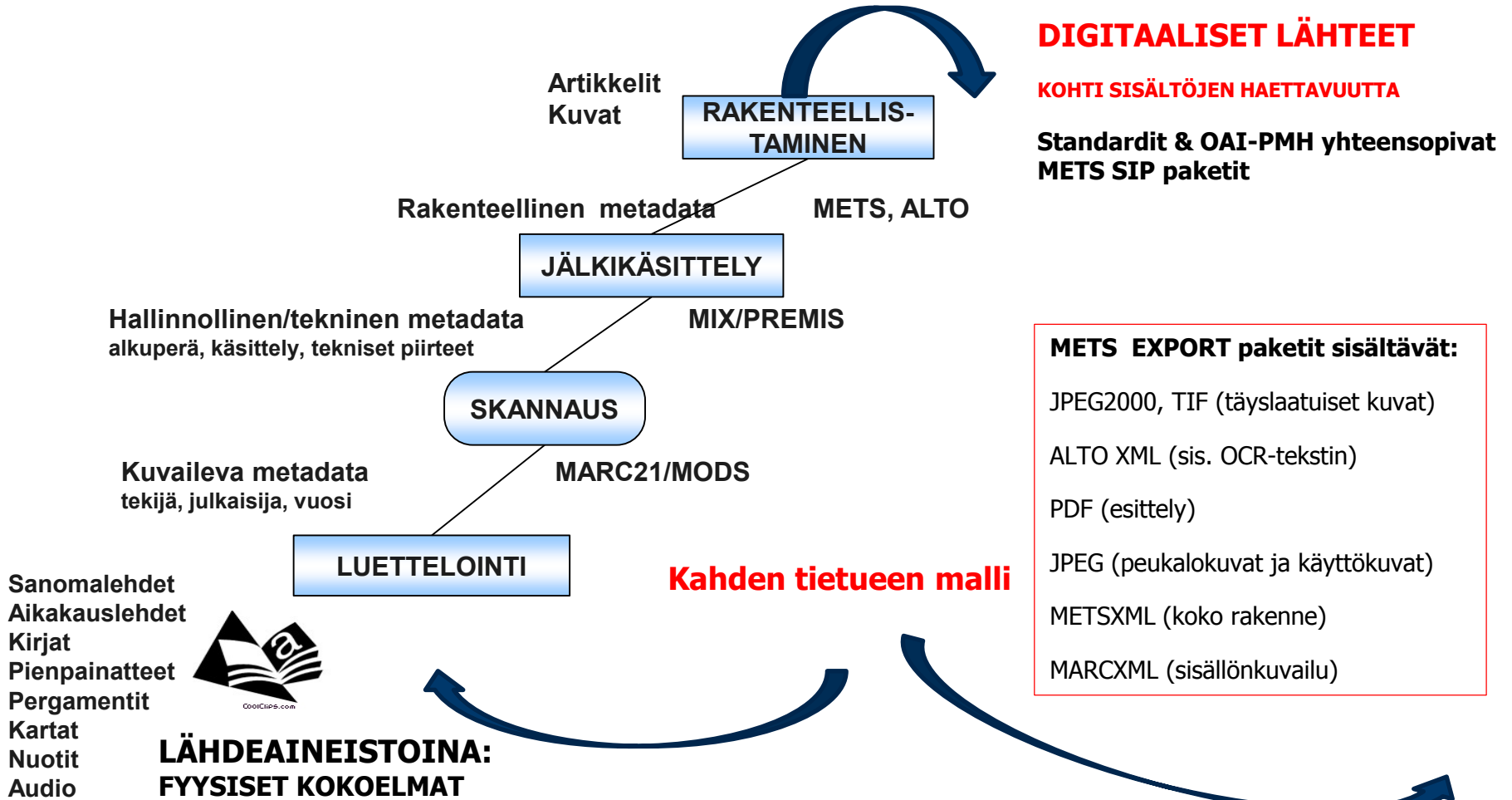


- Pitkäaikaissäilytys (PAS)
  - Varmuuskopiot aineistosta eri paikkoihin
- Esitysjärjestelmä
  - <https://digi.kansalliskirjasto.fi>
- Aineistojen siirto kolmansiin järjestelmiin
  - <https://korp.csc.fi/>

## Kehityshankkeet

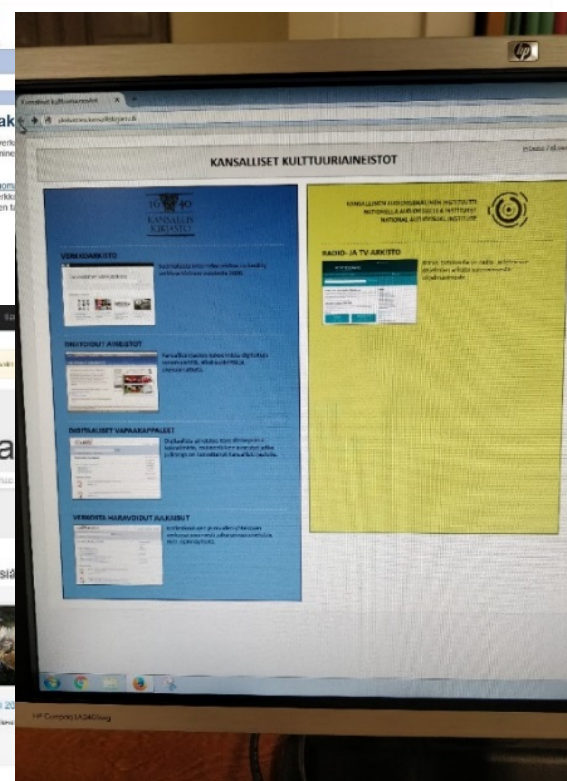
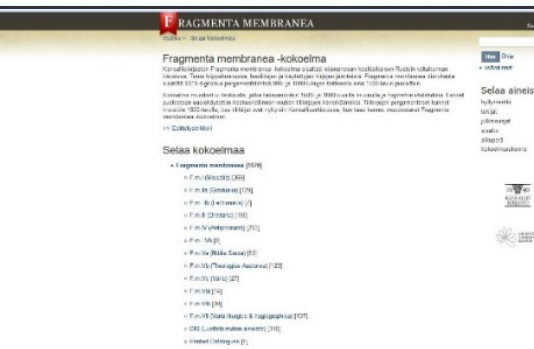
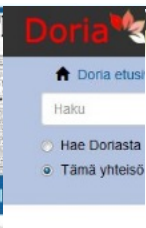
- Tekijänoikeus, ja tietosuoja-asiat, pilotoidaan laajempaa käyttöä => Esim. Aviisi ja Haka-pilotti (<https://blogs.helsinki.fi/digiaviisi/> )
- (<https://www.kansalliskirjasto.fi/fi/projektit/sanoma-ja-aikakauslehtiaineiston-tutkimus-ja-opetus kayton-edistaminen> )
- Tiedonloughinta, ja digitaalinen humanismi => [Digitalia](#)

# METS – ALTO – PAS





# Verkkokoelmat



- UUTUUSIKKUNA
- Historia
  - Filosofia
  - Yhteiskuntatieteet
  - Kielen ja kirjallisuuden tutkimus
  - Taiteiden tutkimus
  - Musiikin tutkimus
  - Venäjän ja itä-Euroopan tutkimus



- ## Kansalli
- Yhteisö sisältää alayhteisöt**
- Elektra [37177]
  - Kansalaisuusa
  - Kansalliskirjasto
  - Kartat [1615]
  - Kirjat [10453]
  - Kortitot ja kol
  - Kuvat [1513]
  - Kasikirjoitukse
  - Pienpainatteet
  - Yhteisö sisältää**
  - Lautapelit [21]
  - Raita - musiikki
  - Urho Kekkose



# Verkkokoelmat : Digi



DIGITAALISET AINEISTOT BETA

DIGI.KANSALLISKIRJASTO.FI

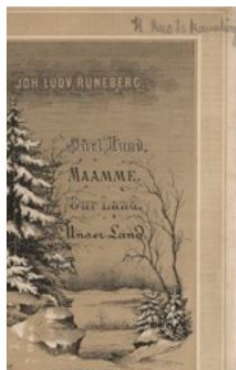
Palautte

SUOMEKSI  
PÅ SVENSKA  
IN ENGLISH

Kirjautu

HAKU KOKOELMAT LEHDET PÄIVÄN LEHDET LEIKKEET

Hae digitoiduista aineistoista



Klassikkokirjasto



Varhaisimmat suomalaiset nuottijulkaisut



Käkisalmen ja Laatokan alueen lehdet



Maantiede ja matkailu



Sanomalehdet

Tutustu historiaan helposti laajojen digitaalisten aineistojen avulla!

25.01.2019

15 178 371

sivua aineistoja



KANSALLISKIRJASTO

<https://digi.kansalliskirjasto.fi/etusivu>

# Verkkokokoelmat : Digi

4161 osumaa

« 1 2 3 4 5 6 7 »



Osuviimmat ensin ▾



27

Flugs vampen

nro 1, 1908  
Helsinki, tuntematon



Tuulispää

nro 7, 14.2.1908  
Helsinki, Suomalaisen Sanomalehtimiesten ...



Näyttämö

nro 3, 1924  
Helsinki, Suomen näyttämöiden liitto



Raataja

nro 4, 21.2.1908  
Helsinki, tuntematon



Fyren

nro 16, 16.4.1910  
Helsinki, tuntematon



Tidskrift utgiven av Pedagogiska fö...

nro 4, 1903  
Helsinki, tuntematon



Finsk Tidskrift

nro 1, 4.1902  
Föreningen Granskaren



Finsk Tidskrift

nro 1, 1.1905  
Föreningen Granskaren

## Apua Digin käyttöön

- Katso opasvideot
  - [Video-ohje](#)
  - [Kirjautuminen](#) Haka-tunnuksin
  - [Haku](#) aineistosta
  - Haun [tarkemmat rajaukset](#)
  - Hakutulosten [käsittely](#)
  - Haku ja [leikkeen teko](#)
  - [Leikekirja](#)
- **LAAJA HAKUOHJE = [Käyttöohjeita ja niksejä!](#)**



KANSALLISKIRJASTO

[Aineistojen selaus](#)

# Verkkokokoelmat : Doria

## Doriassa eri kokoelmia

- [Aarteet-kokoelma](#) sisältää digitoituja kirjoja, karttoja ja käsikirjoituksia
- [Kansalaissota ja itsenäistyminen](#)
- Kirjahistoria sisältää kolme kokoelmaa
  - [Kaunokirjallisuus ja lastenkirjat](#)
  - [Arkkiveisut](#)
  - [Kieliopit](#)
- [Klassikkokirjasto](#) sisältää 1800-luvun sekä 1900-luvun alun klassisen suomalaisen kaunokirjallisuuden teosten digitaalisen kokoelman.
- Valikoima Suomea koskevaa tai suomalaisten kirjoittamaa [maantiedettä ja matkailua käsittelevää kirjallisuutta](#) 1600-luvulta 1900-luvun alkuun.
- [Nordenskiöldin karttakokoelma](#)
- [Pienpainate-kokoelmassa](#) on digitoituja esitteitä, mainoksia, kuvastoja ja hinnastoja, opas- ja ohjekirjasia, tiedotteita ja muita pienimuotoisia painotuotteita.



# Verkkokokoelmat : Raita

## Raita - musiikkia vanhoilta äänilevyiltä

Linkki tälle sivulle: <http://www.doria.fi/raita>

Raita sisältää tekijänoikeusvapaita kotimaisia äänitteitä.

Kansalliskirjasto on digitoinut yli 10 000 kappaletta, jotka kaikki ovat kuunneltavissa kirjaston tiloissa.

Raita on osa *Vuosisadan äänet* –hanketta. Lisätietoja kotimaisista äänitteistä:

- Kansalliskirjaston *Viola*-tietokannassa ja
- Suomen äänitearkisto ry:n luetteloissa.

## Gamla Finska skivorna

Raita innehåller digitaliserade gamla inhemska ljudinspelningar, som kan offentliggöras iakttagande upphovsrätten. Nationalbiblioteket har digitaliserat cirka 10 000 musikstycken, som alla kan lyssnas i bibliotekets lokaler.

Raita är en del av projektet "Sektlets toner". Omfattande uppgifter om inhemska ljudinspelningar finns i *Viola* - samkatalog över musikmaterial.

## Old Finnish Recordings

Raita is a collection of digitized early Finnish sound recordings. The Music Library of the National Library of Finland has digitized over 10,000 pieces of music that all may be listened to at the premises of the library. - Exhaustive information of Finnish recordings can be found in

- *Viola*, the Union Catalogue of Finnish Music;
- in catalogues of the [Finnish Institute of Recorded Sound](#);

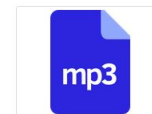


Kansalliskirjasto on luopunut kaikista tekijänoikeuksista ja niiden lähioikeuksista, joita sillä mahdollisesti on ollut Raita-kokoelman teoksiin. Kokoelma on julkaistu Suomessa.

Kuula, Toivo (säv.)

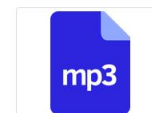
Näytä tarkemmat rajaukset

Villitteet 1-10 / 22



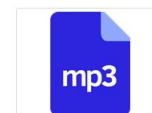
### Marjatan laulu

Kuula, Toivo (säv.); Leino, Eino (san.); Granfelt, Hanna (esitt.); Pergament-Parment, Simon (esitt.) (1929)



### Niin kauan minä tramppaan

Kuula, Toivo (säv.); Rautavaara, Eino (esitt.); Merikanto, Oskar (esitt.) (1909)



### Syystunnelma

Kuula, Toivo (säv.); Leino, Eino (san.); Bäckman, Edvin (esitt.) (1909)



KANSALLISKIRJASTO

<http://www.doria.fi/raita>

# Verkkokokoelmat : Fenno-Ugrica

## Fenno-Ugrica

Suomeksi In English По-русски

Fenno-Ugrica etusivu

Hae Hakuohjeet



Kansalliskirjaston Fenno-Ugrica on uralilaisilla kielillä painettujen julkaisujen digitaalinen kokoelma. Kokoelma koostuu 20 eri kielellä julkaistuista monografioista ja sanomalehdistä. Digitoituja monografianimekkeitä on kokoelmassa noin 1500 ja kausijulkaisuja yli 110 nimekettä.

Uralilaisilla kielillä painettujen aineistojen lisäksi Fenno-Ugrica sisältää kuusi erikoiskokoelmaa: saamelaiskielisen Lapponican, romanikielisen Zingarican ja jiddišinkielisten Hebraican sekä Viron kielen instituutin liivinkielisten ja Komin tasavallan kansalliskirjaston kominkielisten aineistojen kokoelmat. Lisäksi erikoiskokoelmiin kuuluvat Suomen Muinaismuistoyhdistyksen ja Suomalais-Ugrilaisen Seuran digitoituja julkaisuja.

Fenno-Ugrican aineisto on tuotettu Kansalliskirjaston toteuttamassa Sukukielten digitointiprojektissa vuosina 2012–2015 ja Vähemmistökielten digitointiprojektissa vuonna 2016. Projektit on toteutettu Koneen Säätiön rahoituksella.

Fenno-Ugrican aineisto on linkitetty myös Uralica-portaaliin, joka on avoin tietojärjestelmä eri kirjastoissa digitoiduille suomalais-ugrilaisille kirjoille, kartoille ja äänitteille.

Hankkeen yksityiskohdista voi lukea projektin blogista.

Yhteydenotot ja tiedustelut: [kk-palvelu@helsinki.fi](mailto:kk-palvelu@helsinki.fi)

## Kokoelmat

- [Erikoiskokoelmat](#) [1692]
- [Kausijulkaisut](#) [29651]
- [Monografiat](#) [1283]
- [Sanalijat](#) [26]

## Selaa Fenno-Ugricaa

- [Nimekkeet](#)
- [Tekijät](#)
- [Julkaisuajat](#)
- [Asiasanat](#)
- [Uusimmat](#)
- [Selaa kielen mukaan](#)
- [Julkaisutyypit](#)
- [Sivukartta](#)

## Omat tiedot

- [Kirjaudu sisään](#)
- [Rekisteröidy](#)

KONEEN SÄÄTIÖ



KANSALLISKIRJASTO

<http://fennougrica.kansalliskirjasto.fi/>

# Verkkokokoelmat : Fragmenta Membranea

## F RAGMENTA MEMBRANEA

Suomeksi In English

etusivu > Selaa kokoelmaa

### Fragmenta membranea -kokoelma

Kansalliskirjaston Fragmenta membranea -kokoelma sisältää leijonanosan keskiaikaisen Ruotsin valtakunnan itäosissa, Turun hiippakunnassa, laadittujen ja käytettyjen kirjojen jäänteistä. Fragmenta membranea -tietokanta sisältää 9319 digitoitua pergamenttilehteä, 900- ja 1000-lukujen taitteesta aina 1500-luvun puoliväliin.

Kokoelma muodostuu teoksista, jotka takavarikoitiin 1500- ja 1600-luvuilla kruunulle ja hajotettiin itolehdiksi. Lehdet puolestaan uusiokäytettiin keskushallinnon voutien tilikirjojen kansikääreiksi. Tilikirjojen pergamenttiset kannet irrotettiin 1800-luvulla; itse tilikirjat ovat nykyisin Kansallisarkistossa, kun taas kannet muodostavat Fragmenta membranea -kokoelman.

>> Esittelyartikkeli

### Selaa kokoelmaa

- **Fragmenta membranea [1620]**
  - F.m.I (Missalia) [369]
  - F.m.IIa (Gradualia) [129]
  - F.m.IIb (Lectioaria) [7]
  - F.m.III (Breviaria) [190]
  - F.m.IV (Antiphonaria) [203]
  - F.m. IVb [9]
  - F.m.Va (Biblia Sacra) [53]
  - F.m.Vb (Theologiae Auctores) [129]
  - F.m.Vc (Varia) [27]
  - F.m.VIa [14]
  - F.m.VIb [30]
  - F.m.VII (Varia liturgica & hagiographica) [137]
  - DIG (Luetteloimaton aineisto) [318]
  - Printed Catalogues [5]

Hae Ohje

+ lisävalinnat

### Selaa aineistoa

hyllymerkki

tekijät

julkaisuajat

sisältö

alkuperä

kokoelmarakenne



KANSALLISKIRJASTO

<http://fragmenta.kansalliskirjasto.fi/>

# Tiedonlouhinta

- Sanoma- ja aikakauslehtien suhteen kaksi (kolme) tapaa:
- 1.1) **Valmiit ladattavat datapaketit:**  
<https://digi.kansalliskirjasto.fi/opendata/submit>
  - Aineistot vuoteen 1917 saakka.
  - Paketit sisältävät sivukohtaisesti sivun metatiedot, AltoXML-standardin mukaisen rakenneanalyysin ja sivun raakatekstin.
  - Kysymme käyttötarkoituksestasi jonka jälkeen saat henkilökohtaisen latauslinkin valitsemiisi aineistoihin.
- 1.2) **Räätälöitävät datapaketit:**  
<https://wiki.helsinki.fi/display/Comhis/Digi.kansalliskirjasto.fi+Data>



# Tiedonlouhinta

- 2) **Haravoitavat:** Haka-pilotin puitteissa Kansalliskirjasto ei saa jakaa räätälöityjä, mutta saamme ohjeistaa, kuinka käyttäjät voivat tehdä sen itse: <https://natlibfidigi.gitlab.io/gitbook/>
- Lisää ohjeita esim. **COMHIS**in sivuilla:
  - <https://wiki.helsinki.fi/display/Comhis/Digi.kansalliskirjasto.fi+Data>
  - <https://wiki.helsinki.fi/display/Comhis/Interfaces+of+digi.kansalliskirjasto.fi>
- Sovelluksia:
  - Löydät pari apuskriptiä [Githubista](#), joilla esimerkiksi voit yhdestä XML-tiedostosta poimia pelkän tekstin.
  - Java ja Php versiot apuskripteistä löytyvät tästä: [java](#), [php](#).

# Tiedonlouhinta

- Doriassa / Fragmentassa / Fenno-Ugricassa olevien aineistojen tiedonlouhinta rajapintojen kautta
- Kansalliskirjaston julkaisuarkistoissa on käytössä seuraavat OAI-PMH rajapinnat:
  - dublin core
  - qualified dublin core
  - Kansalliskirjaston xml-formaatti, joka on qualified dublin coren laajennus sisältäen myös tiedostolinkit ja tietoa kokoelmista
  - Finnaa varten laadittu formaatti qdc\_finna, joka on räätälöity Finnan tarpeisiin
  - Esim.: [http://fennougrica.kansalliskirjasto.fi/oai/request?verb=ListRecords&metadataPrefix=oai\\_dc](http://fennougrica.kansalliskirjasto.fi/oai/request?verb=ListRecords&metadataPrefix=oai_dc)

# Tiedonlouhinta

- Doriassa / Fragmentassa / Fenno-Ugricassa olevien aineistojen tiedonlouhinta rajapintojen kautta
- OpenSearch-hakurajapinnan avulla voidaan hakea URL:ään upotettujen Lucene-hakusyntaksilla tehtyjen hakujen mukaisia kyselyjä ja vastaus saadaan XML:nä.
  - OpenSearch -rajapinnan kautta julkaisuarkistossa tehtyjen hakujen tuloksia saa vastauksena RSS- tai Atom-syötteinä
  - Esim.: [https://fennougrica.kansalliskirjasto.fi/open-search/?query=Stalin&sort\\_by=3&order=desc&start=0&rpp=50&format=atom&scope=10024/61201](https://fennougrica.kansalliskirjasto.fi/open-search/?query=Stalin&sort_by=3&order=desc&start=0&rpp=50&format=atom&scope=10024/61201)

# Kansalliskirjasto ja avoin data

- Kansalliskirjaston itse tuottamaa dataa ja metadataa voivat kaikki hyödyntää vapaasti. Tarjoamme lukuisia tietovarantoja ja rajapintoja jotka on julkaistu avoimena datana cc0-lisenssillä. Lisätiedot aineistoista ja niiden käyttömahdollisuuksista on koottu datakatalogiin.
- Kansallisbibliografia Fennican data on hyödynnettävissä useilla eri tavoilla. Fennican tietoja pääsee selaamaan avoimen datan päälle rakennetun käyttöliittymän avulla osoitteessa <http://data.nationallibrary.fi/>.
- Kansallisbibliografian [avaaminen](#) avoimena datana, joulukuu 2017.

# Kansalliskirjasto ja avoin data

- Avoimen datan lähteet Kansalliskirjastossa
  - **Datakatalogi:** <http://data.nationallibrary.fi>
  - **Linkitetty data:** <http://data.nationallibrary.fi/bib/me/CFENNI>
  - **Datapaketit (Kokoteksti / metadata):**  
<https://www.kiwi.fi/display/Datacatalog/Data+sets>
  - **API:t:** <https://www.kiwi.fi/display/Datacatalog/APIs>
  - **Sanastopalvelut:** <http://finto.fi/fi/>

## Open Data Service

Access open data sources maintained by the National Library of Finland

### Data Catalog

Our data catalog showcases the open data sets and APIs of the National Library.



#### Data sets

Open data sets available for reuse.



#### APIs

Interfaces that provide access to our data.

### Linked Data

#### Linked Bibliographic Data

The national bibliography Fennica is available as RDF Linked Open Data.

#### Browse our Linked Data

##### Collections

- Fennica

#### Access

- SPARQL endpoint
- Linked Data Fragments server
- Downloadable data dumps

#### Documentation

- Fennica RDF data model documentation

Vocabularies About Feedback Help | suomeksi på svenska

from all English Search

## Welcome to the Finto service!

Finto is a Finnish thesaurus and ontology service, which enables both the publication and browsing of vocabularies. The service also offers interfaces for integrating the thesauri and ontologies into other applications and systems.

### Available vocabularies and ontologies

**GENERAL**

- Allärs - General thesaurus in Swedish
- Finnish Corporate Names
- KOKO Ontology
- Metatietosanasto
- Pondus categories
- Suomalainen lajityyppi- ja muutosanasto
- UCUM - The Unified Code for Units of Measure
- YSA - General Finnish thesaurus
- YSAn ja YSO:n käsite-ehdotukset
- YSO - General Finnish ontology

**SOCIETY**

- IPTC NewsCodes
- JUHO - Julkishallinnon ontologia
- Julkisten palvelujen luokitus
- JUPO - Finnish Ontology for Public Administration Services
- KEKO - Ontology for Education for Sustainable Development
- Korkeakoulujen tutkimustiedonkeruussa käytettävä tieteenalaluokitus
- Lapponica
- LIITO - Liiketoimintaontologia
- MERO - Merenkulualan ontologia
- OIKO - Oikeushallinnon ontologia
- PUHO - Puolustushallinnon ontologia

Tweets by @Fintopalvelu

# Linkitetty Data

- Fennica on julkaistu avoimena datana CC0-lisenssillä, joka antaa mahdollisuuden sen vapaaseen hyödyntämiseen esimerkiksi sovelluksissa ja datan visualisoinneissa.
  - <http://data.nationallibrary.fi/bib/me/CFENNI>



National Library of Finland  
Open Data and Linked Data Service

Search works, persons, organizations and subjects:

Concept: [miehitys](#) (p39)  
Work: [Lempeä miehitys](#) (Wo1152041700)  
Work: [Yrityksen organisaation rakenne, miehitys ja toiminta](#) (Wo00511642)  
Work: [Yrityksen organisaation rakenne, miehitys ja toiminta](#) (Wo06590950)  
Work: [Miehitys ja sotavuodet Islannissa 1940-45 = : Ockupation och krigstid](#)  
Work: [Kriisien ja sodan maailma : sota etsii toista rintamaa - Tanskan mieh](#)

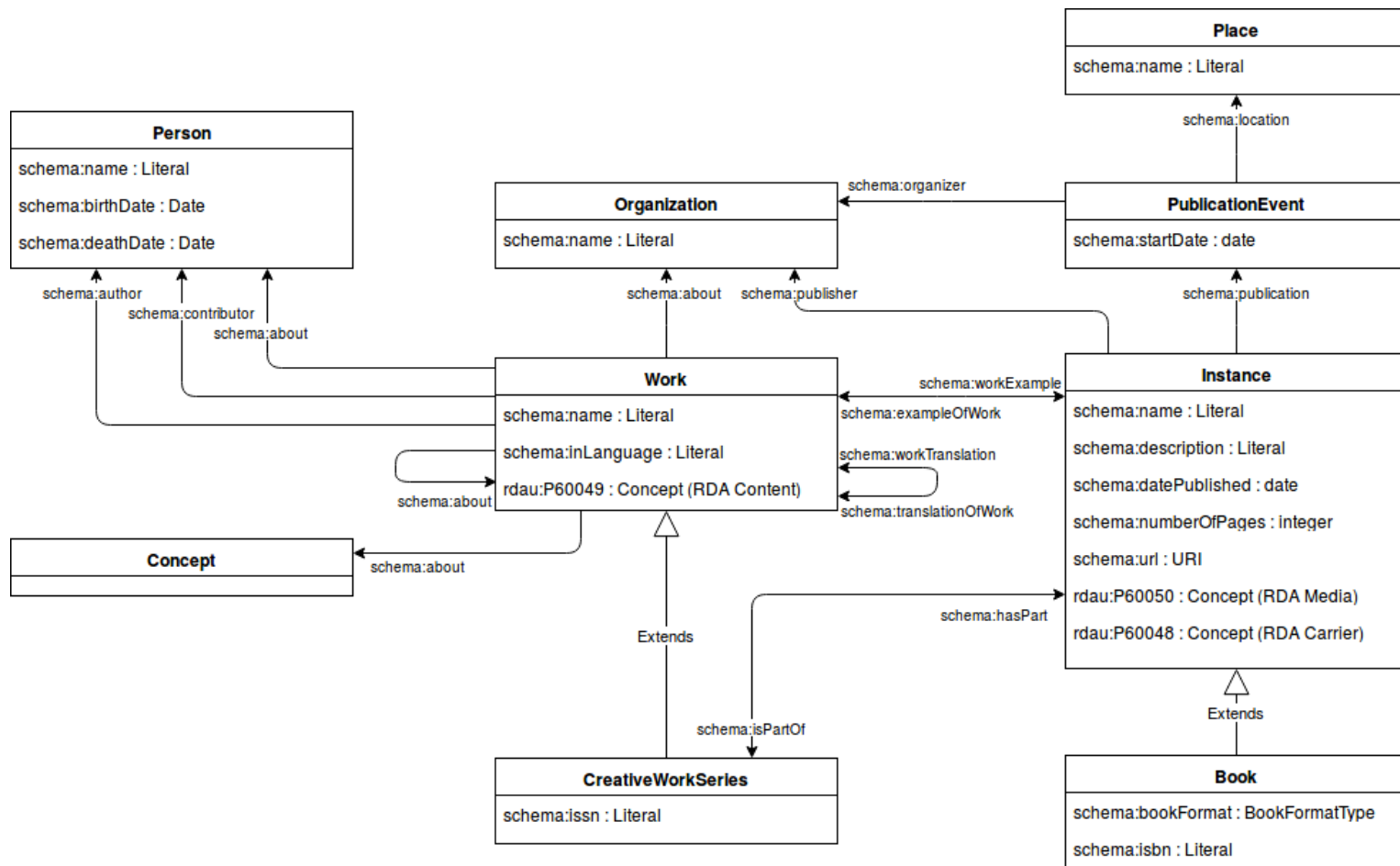
## Search results: 'miehitys'

1. [miehitys](#)  
Concept
2. [Lempeä miehitys](#)  
Work
3. [Yrityksen organisaation rakenne, miehitys ja toiminta](#)  
Work
4. [Yrityksen organisaation rakenne, miehitys ja toiminta](#)  
Work
5. [Miehitys ja sotavuodet Islannissa 1940-45 = : Ockupation och krigstid på Island 1940-45](#)  
Work
6. [Kriisien ja sodan maailma : sota etsii toista rintamaa - Tanskan miehitys ja Norjan sota - Saksan voitto lännessä : ilmasota ja merisota : Itä-Afrikan ja Libyan taistelut : Saksan uudestijärjestely Balkanilla : Itärintaman taistelut v.1941 : 2, Maailman tapahtumat vv. 1940-1942](#)  
Work



KANSALLISKIRJASTO

# Linkitetty Data



# Finto

- Finto on Kansalliskirjastossa ylläpidetty keskitetty ontologia- ja sanastopalvelu. Finton avulla käyttäjä saa yhdestä paikasta käyttöönsä [hyvin laajan kirjon](#) erilaisia metatietovarantoja.
- Finto tarjoaa paitsi yleissanastoja (Yleinen suomalainen ontologia - YSO, KOKO) myös monia erikoisalojen sanastoja ja ontologioita: esimerkiksi julkishallinnon, liiketalouden, lääke- ja kasvitieteen sekä valokuvauksen aloilta.
- Finto on vapaasti kaikkien käytettävissä ja selailtavissa.



# Avoim Data

- Mitä avoimella datalla voi saada aikaiseksi?
  - [COMHIS ja Fennica](#)
  - [Bibliographic Data Science and the History of the Book \(c. 1500–1800\)](#)
  - [Wide Hackahton](#)
    - [YSO ja YSO-Paikat](#)
    - FennicaTrends –[visualisointi](#)
  - Annif
    - Automaattisen sisällönkuvailun [väline](#)
    - <http://annif.org/>

Seuraa myös Twitter-kanavaa: <https://twitter.com/NatLibFiData>

# Kiitos! Ja kysymyksiä?

[www.kansalliskirjasto.fi](http://www.kansalliskirjasto.fi)  
[kk-tutkijapalvelut@helsinki.fi](mailto:kk-tutkijapalvelut@helsinki.fi)

Jussi-Pekka Hakkarainen  
Tuula Pääkkönen  
Jukka Kervinen  
Pirjo Karppinen  
Osma Suominen  
Mikko Lappalainen  
Jaakko Tahkokallio