

# Ajankohtaisia asioita

13.11.2020



# Automaattisen kuvailun kehittäminen Kansalliskirjastolla 2019-2020

1. Kehitetään Annif-työkalua erityisesti sisällönkuvailun laadun näkökulmasta.
2. **Julkaistaan avoin rajapintapalvelu automaattiseen sisällönkuvailuun Finto-palvelun yhteyteen (Finto AI) ja tuetaan sen käyttöönottoa eri organisaatioissa.**
3. Edistetään Annif-avusteisen kuvailun käyttöä Kansalliskirjaston omissa järjestelmissä (julkaisuarkistot, e-vapaakappaleet...).
4. Tutkitaan automaattisen bibliografisen kuvailun mahdollisuuksia kuvailuprosessien kehittämisessä.
5. Viestintä ja yhteistyö.



Yleiskäyttöinen **työkaluohjelmisto**  
automaattiseen sisällönkuvailuun

Tukee monia kieliä ja eri sanastoja

Verkkosivut, GitHub-koodivarasto, demolomake  
ja rajapintapalvelu testikäyttöä varten

Maailmanlaajuinen käyttäjäyhteisö

Tekninen dokumentaatio englanniksi,  
verkkokurssi, käyttäjäryhmä annif-users



Annifiin perustuva automaattisen sisällönkuvailun  
**verkkopalvelu**

Tukee YSO-kuvailua kolmella kielellä (YKL tulossa)

Verkkopalvelu, jossa lomake ja rajapintapalvelu

Kotimaisen kuvailutyön tukipalvelu, Finton sisar

Käyttöönotto-ohjeistus suomeksi, ruotsiksi ja  
englanniksi, Finton tiedotuskanavat

# Finto AI julkaistiin 28.5.2020

## Automaattinen sisällönkuvailu osaksi Finton palveluja

1.6.2020 13:02:40 EEST | [Helsingin yliopisto](#)

Jaa



**Finto laajenee automaattisen sisällönkuvailun palveluihin Finto AI -työkalun ja sen rajapintapalvelun myötä. Palvelu tukee suomen-, ruotsin- ja englanninkielisen tekstin automaattista sisällönkuvailua.**

Finto AI on saatavilla osoitteessa [ai.finto.fi](https://ai.finto.fi), josta löytyvän lomakkeen avulla palvelua voi kokeilla ja käyttää. Automaattinen sisällönkuvailu toimii parhaiten opinnäytetö- ja tietokirja-

<https://www.sttinfo.fi/tiedote/automaattinen-sisallonnkuvailu-osaksi-finton-palveluja?publisherId=3747&releaseId=69882342>

# Annif 0.49

[Edit](#)

 osma released this on Jul 30 · [27 commits](#) to master since this release

This release introduces the hyperopt CLI command for hyperparameter optimization. Initially it can only be used for finding optimal ensemble weights. The Web UI now follows the same visual style as the annif.org website. There are also some improvements to CLI commands, memory optimizations and bug fixes.

New features:

- [#240/#321/#414](#) Hyperparameter optimization of ensemble weights

Improvements:

- [#424/#426](#) New style for Web UI
- [#430](#) Define short form for CLI options and fix some of their docstrings
- [#428](#) Memory optimization: Avoid double allocation of NumPy arrays in eval operation

Maintenance:

- [#437](#) Upgrade TensorFlow to version 2.3.0 (from 2.2.0)

Bug fixes:

- [#431](#) Problem parsing timestamps from Maui Server
- [#432](#) Make modification timestamps timezone-aware

▼ **Assets** 2

 [Source code](#) (zip)

 [Source code](#) (tar.gz)

# Käyttöönotto Kansalliskirjaston järjestelmissä

The screenshot shows the LUT University website with the Doria and Julkari digital publishing systems. The Doria interface includes a search bar with the text "LUTPubin haku- ja käyttöohje" and a search button. Below it, the text "LUTPub on LUT-yliopiston avoin julkaisuarkisto." is visible. The Julkari interface shows a search bar with the text "Julkarin haku- ja käyttöohje" and a search button. Below it, the text "Julkari on Sosiaali- ja terveysministeriön julkaisuista on saatavilla sähköinen versi laitoksen osalta myös julkaisurekisterinä." is visible. The background shows a blue and orange graphic with the text "LUTPub" and "HAE" (Search).

The screenshot shows the Kansalliskirjasto digital publishing system interface. The header includes the text "Digitaalisten julkaisujen arkistointi" and "KANSALLISKIRJASTO". Below the header, there are several sections: "Luovuttajan tiedot" (Author information), "Julkaisun tiedot" (Publication information), and "Julkaisun tekemäärä" (Publication count). The "Luovuttajan tiedot" section includes fields for "Yhteystiedot" (Contact information), "Sähköpostiosoite" (Email address), "Puhelinnumero" (Phone number), and "Organisaatio" (Organization). The "Julkaisun tiedot" section includes fields for "Julkaisun tyyppi" (Publication type) and "Päättämät" (Status). The "Julkaisun tekemäärä" section includes a field for "Luo uuden yhden julkaisun" (Create one new publication) and "Luo uuden useita julkaisuja" (Create several new publications).

## E-vapaakappaleiden vastaanotto ja käsittely

DSpace-pohjaiset julkaisuarkistot: Osuva, Theseus ...

# Annif-verkkokurssit syksyllä 2020



21.9. [Introduction to Annif and automated subject indexing](#) -esitys DCMI Virtual -konferenssissa (Osma Suominen, Koraljka Golub, Annemieke Romein & Sara Veldhoen)

- Lisäksi [itsenäisesti suoritettava verkkokurssi](#) sis. videoita ja harjoitustehtäviä

24.11. Esitys ja tutoriaali myös [SWIB 2020](#) -virtuaalitapahtumassa

# HPD-hanke

## **Tekoäly rikastaa muistiorganisaatioiden digitaalisia aineistoja**

CSC:n, Kansalliskirjaston ja Kansallisarkiston *High-Performance Digitisation* -hankkeessa luodaan älykäs annotaatioputki arkistoitujen aineistojen käsittelyyn ja rikastukseen. Aineistot voivat olla esimerkiksi skannattuja sanomalehtiä, kirjoja, kuvia ja virallisia asiakirjoja. Tekoälyn ja koneoppimisen avulla data laatua voidaan parantaa ja dataa voidaan hyödyntää tehokkaammin.

Ajalle 9/2019 - 9/2020, mutta myönnetty jatkoaikaa vuoden 2020 loppuun

HPD-seminaari tammikuussa 2021



# Muuta ajankohtaista

- Uusi Annifin versio, mallipäivitys
- Finto AI:n kielivalinnat
- [Yake](#)-kokeet
- Artikkelit Annifista

# Seuraava verkoston tapaaminen

- Alkukeväätä 2021?
- Vapaaehtoisia järjestäjiä - esim. Yle tai Kavi?