

Annifin kuulumisia & automaattisen kuvailun vuosi 2022

Automaattisen kuvailun verkosto 18.5.2022

Annifin kuulumisia

- Mitä viimeisimmän tapaamisen jälkeen?
- Toimintasuunnitelma 2022
- Uudistuksia jne.

Artikkeli J LIS.it-lehdessä julkaistu 1/2022

J LIS.it vol. 13, no. 1 (January 2022)
ISSN: 2038-1026 online
Open access article licensed under CC-BY
DOI: 10.4403/jlis.it-12740

J LIS.it

Annif and Finto AI: Developing and Implementing Automated Subject Indexing

Osma Suominen^(a), Mona Lehtinen^(b), Juho Inkinen^(c)

a) National Library of Finland, <http://orcid.org/0000-0003-0042-0745>

b) National Library of Finland, <http://orcid.org/0000-0002-4735-0214>

c) National Library of Finland, <http://orcid.org/0000-0002-6497-6171>

Contact: Osma Suominen, osma.suominen@helsinki.fi; Mona Lehtinen, mona.lehtinen@helsinki.fi;

Juho Inkinen, juho.inkinen@helsinki.fi

Received: 5 May 2021; **Accepted:** 21 May 2021; **First Published:** 15 January 2022

Annifin / Finto AI:n käyttöönottoja

Finto AI-palvelun käyttö

Omatoiminen käyttö verkkoselaimella

API-integraation kautta:

- **Julkaisuarkistot:** JYX (JYU), Osuva (Vaasan yo), Theseus (amk:t), Trepo (Tampereen yo), Taju (Taideyliopisto), Lauda (Lapin yo) ...
- 40 **museon** Collecte-kokoelmanhallinta
- **Kirjavälitys Oy**, ennakkotiedot uutuuskirjoista
- KK:n **E-vapaakappaleiden** vastaanotto

Itsenäinen Annif-asennus

- **Yle** / verkkouutisten asiasanoitus
- Saksan kansalliskirjasto **DNB**
- Saksan taloustieteen instituutti **ZBW**
- Alankomaiden kansalliskirjasto **KB.nl**
- Alankomaiden tilastokeskus **CBS**
- Puolan kansalliskirjasto
- Dissem.in -palvelu

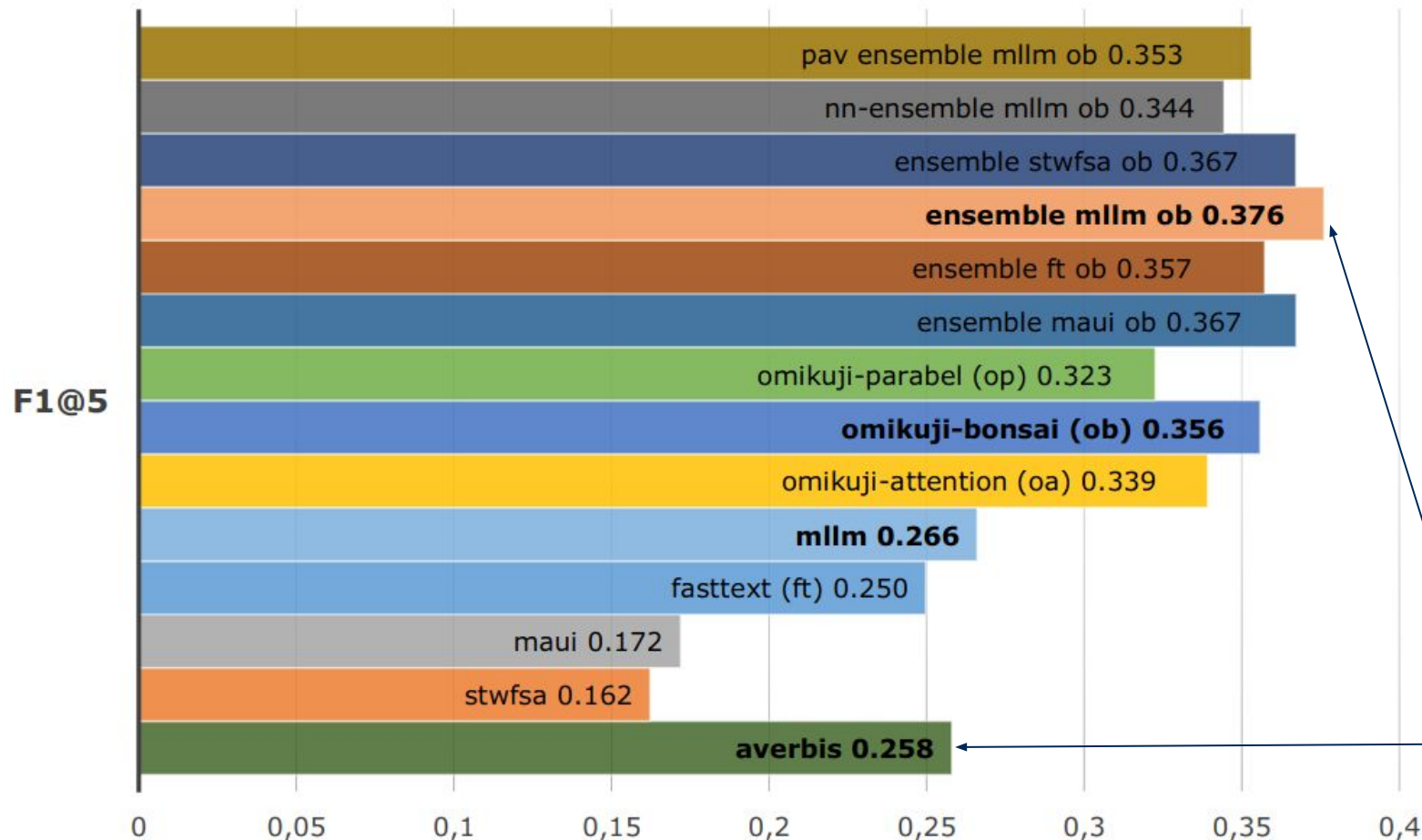
... ja muita, joista emme välttämättä edes tiedä!

Saksan kansalliskirjasto: Annif tuotantoon 4/2022

<https://swib.org/swib21/slides/03-02-uhlmann.pdf>

6

Results GND descriptors



"Our thanks go to you and the whole Annif team for your great software. Annif does what we need, and it is reliable, well maintained and continuously being improved. No wonder we use it ;-)" – [Claudia Grote](#), DNB EMa team

1261 Online publications

F1-score

(n=5 descriptors)

paras Annifin algoritmiyhdistelmä

nykyinen kaupallinen työkalu



Dalai

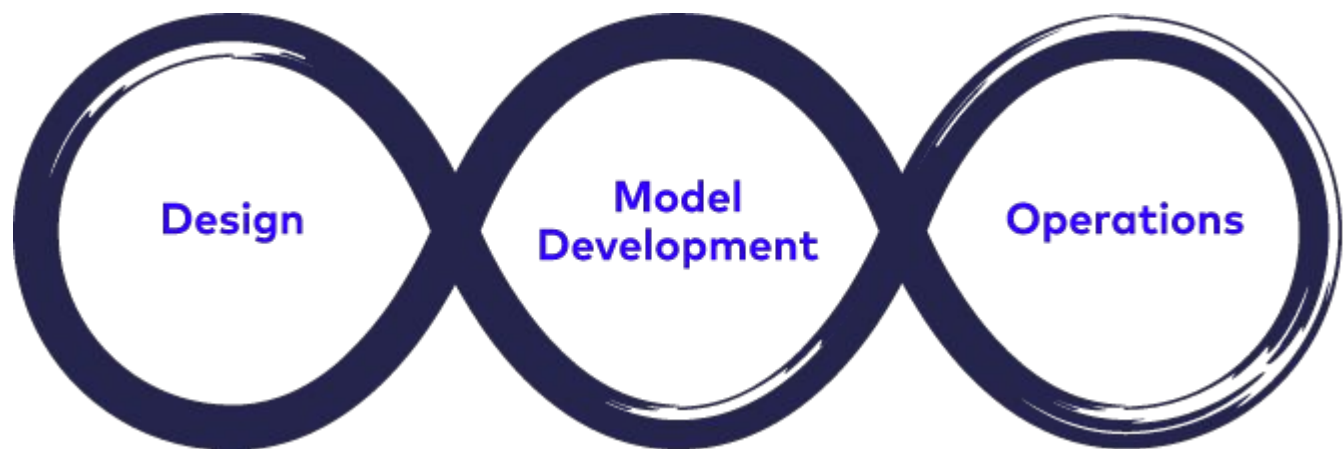
- <https://kansallisarkisto.fi/dalai>
 - Digitaalisten aineistojen laadun ja käytettävyyden parantaminen tekoälyavusteisesti
- Annifia kokeiltu hankkeen tiimoilta

Automaattisen kuvailun projektisuunnitelma 2022

- Parannetaan automaattisen kuvailun laatua mm. hyödyntämällä koneoppimista ja kieliteknologiaa sekä tekemällä yhteistyötä
- Laajennetaan automaattista sisällönkuvailua luokituksiin (YKL, Thema) ja kaunokirjallisuuteen (KAUNO)
- Kehitetään Finto AI -palvelun toiminnallisuuksia, käytettävyyttä ja saavutettavuutta
- Edistetään automaattisen kuvailun käyttöönottoja Suomessa ja maailmalla

MLOps ja Data Version Control

Helpotetaan laatutyötä – tehostetaan koneoppimismallien kehitystä



Tämänhetkistä kehitystyötä

- Luokitukset: osoitteessa <https://dev.annif.org/> on testattavissa YKL- ja Thema-luokitusta Annifilla
- Mallipäivityskierros tekeillä, päivitetyt mallit julkaistaan pian

Automaattisen bibliografisen kuvailun kehitys

Hyödyllistä mm. e-vapaakappaleiden vastaanotossa ja julkaisuarkistoissa

- **PDF-tiedosto** → **perusmetatiedot** (tekijä, julkaisuaika, nimeke, sarja...)

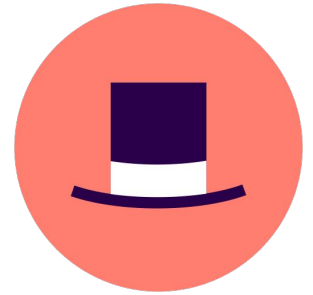
Aiemmin testattu **Grobid**-ohjelmaa, joka valitettavasti ei toiminut hyvin.
Käynnissä pilotti **Lingsoft Oy:n** kanssa.

Puoliautomaattisen sisällönkuvailun tekeminen Finto AI:lla – käyttöopas

- Yhdessä Sisällönkuvailun asiantuntijaryhmän kanssa
- Yleisellä tasolla puoliautomaattisesta sisällönkuvailusta
 - Miksi? Hyödyt ja sudenkuopat?
- Esimerkkinä Finto AI:n lomakekäyttö

Annif tutorial

Annifin käytännön kurssi – 5 kertaa vuonna 2021



LES FUTURS
FANTASTIQUES

SWIB21

Semantic Web in Libraries

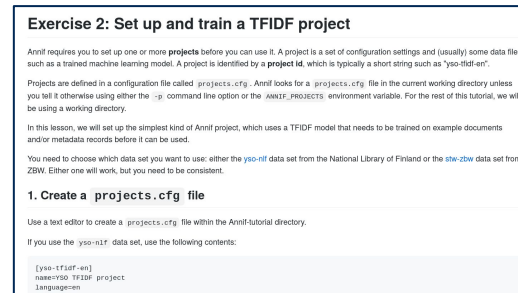
SWIB20

Semantic Web in Libraries

DCMI Virtual, 2020
September 14th-25th, 2020

SWIB19

Semantic Web in Libraries



Videos and exercises freely available on YouTube & GitHub!



Finto AI:n parannuksia

1. Perustettiin oma GitHub-varasto Finto AI:lle, ml. julkinen kehitysjojo <https://github.com/NatLibFi/FintoAI>
2. Toteutetaan käyttäjien toiveita:
 - PDF-tiedoston lataaminen Finto AI:hin -> asiasanaehdotukset
 - kopiointinappuloiden parannukset (esim. kopioi kaikki kerralla)
 - saavutettavuuden parantaminen edelleen

Kuvailtava teksti

Syötä tiedosto

Syötä URL

Valitse tiedosto

Selaa

Tuetut tiedostomuodot: .txt, .pdf, .doc, .docx, .odt, .rtf, .pptx, .epub, .html

Voit myös kopioida tekstin tähän. Ehdotukset saat painamalla "Anna aihe-ehdotukset"-nappia



Tai voit pudottaa tiedoston tai URL:n tähän.

Sisällönkuvailu

Sanasto ja tekstin kieli

YSO suomi (2021-09-23)

Ehdotusten enimmäismäärä

10 15 20

Aihe-ehdotusten kieli

sama kuin tekstin kieli

Anna aihe-ehdotukset

Kiitos!



Juho Inkinen



Mona Lehtinen



Osma Suominen