



Open source tools for SKOS vocabulary services: Skosmos and Annif

Osma Suominen, Joeli Pokkinen, Mona Lehtinen, Juho Inkinen

NKOS 2025, Tampere, Finland

23 September 2025

Part I:

Skosmos and Finto: publishing SKOS controlled vocabularies on the Web

Skosmos

finto

Osma Suominen

Skosmos

- Web-based SKOS vocabulary browser and publishing tool
 - SKOS is a Linked Data standard for representing thesauri, subject headings, taxonomies, classifications, lightweight ontologies...
- Facilitates subject indexing, information retrieval and vocabulary development
- Developed by the National Library of Finland
 - with contributions from many others
- Open source (MIT license)

Finto service of the National Library of Finland

- Finto started as a project in 2013, financed by Finnish Ministry of Finance and the Ministry of Education and Culture
- Launched to the public in 2014
- Permanent service at the National Library of Finland since 2019
- Service and infrastructure of linked data vocabularies for the *whole Finnish Public Sector* (not just libraries, museums and archives)



YSO - General Finnish ontology

Content language English Search

A-Z

Hierarchy

Groups

New

- events and action
- objects
 - abstract objects
 - physical objects
 - inanimate objects
 - abacuses
 - adhesive tapes
 - admission tickets
 - amigurumi
 - animal bodies
 - armours**
 - articles (inanimate objects)
 - artificial nails
 - artificial organs
 - audience areas
 - balloons
 - banderoles
 - baskets
 - bathtubs
 - bird tables
 - birdhouses
 - bits (bridles)
 - blinds
 - bookmarks
 - booms
 - braids and weaves
 - bridge floors
 - bridles
 - briquets
 - brooms and brushes
 - buckles
 - busses (computing)
 - buttons (clothing)
 - candles
 - candlesticks
 - cardboard
 - cards
 - cart structures
 - celestial bodies
 - ceramics
 - chains (objects)

OF FINLAND

objects > physical objects > inanimate objects > armours

PREFERRED TERM

armours

TYPE

General concept

BROADER CONCEPT

[inanimate objects](#)

RELATED CONCEPTS

[military uniforms](#)

BELONGS TO GROUP

[67 Warfare. Military Technology. Defence. Weapons](#)

IN OTHER LANGUAGES

[haarniskat](#)

Finnish

[harnesk](#)

Swedish

URI

<http://www.yso.fi/onto/ysop14728>

Download this concept:

[RDF/XML](#) [TURTLE](#) [JSON-LD](#)

Last modified 5/10/17

EXACTLY MATCHING

CONCEPTS

[armours](#)

KOKO Ontology

[haarniskat](#) (fi)

YSA - General Finnish thesaurus

[harnesk](#) (sv)

Allärs - General thesaurus in Swedish

Images indexed with the term in Finna 245

Image



Liivimaan
kenraaliku
sotamarsa
Da
Museovira



valtaneuvos Pontus De la



Kustaa Eriksson Vaasa



Ruotsin kuningas Kaarle XI



Ruotsin kuningas Fredrik I

Where we publish thesauri, classifications, ontologies etc. for use by libraries, archives, museums, media, students...



Finto.fi

Skosmos installations around the world

UNESCO Thesaurus (UNESCO, UN)



Vocabularies About Feedback Help | Interface language: English ▾

UNESCO Thesaurus

Content language

English ▾



Search

Alphabetical

Hierarchy

Groups

A Å B C D E F G H I J K L M N
O P Q R S T U V W X Y Z

Aboriginals → Indigenous peoples
Abortion
Absenteeism → Leave
Abstract journals → Abstracts
Abstract reasoning → Reasoning
Abstracting
Abstracting and indexing services → Bibliographic services
Abstracts
Abuse of human rights → Human rights violations
Abuse of power → Oppression
Academic achievement
Academic admission → Admission requirements
Academic buildings
Academic degrees → Degrees
Academic equipment → Educational equipment
Academic facilities → Educational facilities
Academic failure
Academic fraud
Academic freedom
Academic grouping → Educational grouping
Academic laboratories → University laboratories
Academic libraries
Academic management → Educational management
Academic misconduct → Academic fraud
Academic performance → Academic achievement
Academic qualifications → Educational qualifications
Academic recognition → Equivalence between diplomas

Libraries > Academic libraries

PREFERRED TERM

Academic libraries

Search in UNESDOC

BROADER CONCEPT

[Libraries](#)

RELATED CONCEPTS

[Academic buildings](#)

[Higher education institutions](#)

[School libraries](#)

[Universities](#)

ENTRY TERMS

University libraries

BELONGS TO GROUP

[Information and communication > Documentary information systems](#)

IN OTHER LANGUAGES

مكتبات أكاديمية

Arabic

مكتبات جامعية

Bibliothèque universitaire

French

Bibliothèque académique

Академические библиотеки

Russian

Университетские библиотеки

Biblioteca universitaria

Spanish

Biblioteca académica

URI

<http://vocabularies.unesco.org/thesaurus/concept1060>

Download this concept:

[RDF/XML](#) [TURTLE](#) [JSON-LD](#)

Last modified 12/15/19

AGROVOC Thesaurus (FAO of the UN)



Food and Agriculture Organization
of the United Nations

AGROVOC About Feedback Help

| Interface language: English ▾

AGROVOC Multilingual Thesaurus

Content language

English ▾

×

Search

Alphabetical

Hierarchy

Groups

Leptocarpum
Leymus
Lolium
Melinis
Melocanna
Microstegium
Miscanthus
Muhlenbergia
Nardus
Oryza
Oryza glaberrima
Oryza longistaminata
Oryza perennis
Oryza punctata
Oryza rufipogon
Oryza sativa
Oryzopsis
Oxytenanthera
Panicum
Pascopyrum
Paspalidium
Paspalum
Pennisetum
Perotis (Spermatophyta)
Phalaris
Phleum
Phragmites
Phyllostachys
Pleioblastus
Poa
Pseudocereus

... > Liliopsida > Poales > Poaceae > Oryza > Oryza sativa

PREFERRED TERM

① Oryza sativa

BROADER CONCEPT

Oryza (en)

ENTRY TERMS

① Oryza indica (en)

① Oryza japonica (en)

HAS DISEASE

rice blast disease (en)

tungro disease (en)

HAS PEST

Nilaparvata lugens (en)

IS USED AS

starch crops (en)

PRODUCES

rice (en)

IN OTHER LANGUAGES

① رز مزروع

Arabic

① رز هندي

① رز ياباني

① рыс пасяўны

Belarusian

① 稻

Chinese

① 糯米

① 日本稻

① Oryza sativa

Czech

① Oryza indica

① Oryza japonica

① Oryza sativa

Danish

ELSST Thesaurus (CESSDA)



[Documentation](#) [About](#) [Feedback](#) [Help](#)

ELSST Thesaurus (Version 5 - 2024)

Content language

English ▾



Search

Alphabetical

Hierarchy

New and
Deprecated

- POLITICAL SYSTEMS
- POLITICAL, SOCIAL AND ECONOMIC ISSUES
- POLITICS
- POPULATION
- POPULATION MIGRATION
- POSTWAR MEASURES
- PRESTIGE
- PRODUCTS
- PROPERTY, OWNERSHIP AND TENURE
- PSYCHOLOGICAL EFFECTS
- PSYCHOLOGY
- PUBLIC FIGURES
- PUBLIC OPINION POLLS
- PUBLIC SERVICES
- QUALIFICATIONS
- QUALITY
- RELIGION
- RESEARCH
 - AGRICULTURAL RESEARCH
 - CULTURAL RESEARCH
 - ECONOMIC RESEARCH
 - EDUCATIONAL RESEARCH
 - MARKET RESEARCH
 - PSYCHOLOGICAL RESEARCH

[RESEARCH](#) > SOCIAL RESEARCH

PREFERRED TERM

SOCIAL RESEARCH

BROADER CONCEPT

RESEARCH

RELATED CONCEPTS

COMMUNICATION RESEARCH

ENTRY TERMS

SOCIAL SCIENCE RESEARCH

IDENTIFIER

urn:ddi:int.cessda.elsst:fccdd028-7e40-406b-9b96-cbf71d2c334d:5

IS A VERSION OF

<https://elsst.cessda.eu/id/fccdd028-7e40-406b-9b96-cbf71d2c334d>

DATE ISSUED

2024-09-23

KEYWORD SEARCH

[Search for 'SOCIAL RESEARCH' in CESSDA Data Catalogue](#)

HAS PREVIOUS VERSION

[SOCIAL RESEARCH](#) (ELSST Version 4)

IN OTHER LANGUAGES

[SOCIÁLNÍ VÝZKUM](#)

Czech

[SOCIAAL ONDERZOEK](#)

Dutch

[YHTEISKUNTATIETEELLINEN
TUTKIMUS](#)

Finnish

Loterre (ISTEX, France)

[Home](#) / [Signal theory and processing](#)

Signal theory and processing

English ▾

×

Search

Help

ALIGN

ANNOTATE

DOWNLOAD

Alphabetical

Hierarchy

Groups

A B C D E F G H I J K L M N O P

Q R S T U V W X Y Z 0-9

a posteriori estimation
a priori estimation
A/D conversion → AD conversion
AAM → active appearance model
absolute error
absorption wave
abstracting
access time
ACO optimisation → **ACO optimization**
ACO optimization
acoustic analysis
acoustic convolution
acoustic filter
acoustic image
acoustic imaging
acoustic noise
acoustic radar → sodar
acoustic receiver
acoustic resonator filter
acoustic signal
acoustic signal detection
acoustic signal processing
acoustic transmitter
acoustic wave filter
acoustic wave interference
acoustic optical modulation

PREFERRED TERM

ACO optimization 

SYNONYM(S)

ACO optimisation
ant colony optimisation
ant colony optimization

BELONGS TO GROUP

Treatment

IDENTIFIER

<http://data.loterre.fr/ark:/67375/Signal/1747>

IN OTHER LANGUAGES

optimisation ACO	French
<i>optimisation par colonies de fourmis</i>	
optimización ACO	Spanish
<i>optimización en colonias de hormigas</i>	

URI

<http://data.loterre.fr/ark:/67375/SN8-CCBZN9N6-C> 

DOWNLOAD THIS CONCEPT:

[RDF/XML](#) [TURTLE](#) [JSON-LD](#)

Datu (EHES, France)



Thésaurus Savoirs

[Vocabularies](#)[About](#)[Feedback](#)[Help](#)[en - fr](#)Content language **French**[x](#) [Search](#)[Alphabetical](#)[Hierarchy](#)[Groups](#)

- ▶ acteurs de savoir
- ▶ construction des savoirs
- ▶ espaces savants
- ▶ inscription des savoirs
- ▶ matérialité des savoirs
- ▶ pratiques savantes
- ▶ typologie des savoirs
 - ▶ disciplines
 - ▶ divisions historiques des savoirs
 - ▶ sciences appliquées
 - ▶ sciences formelles et expérimentales
 - ▶ informatique
 - algorithmique et combinatoire
 - architecture des machines et des systèmes
 - communication humain-machine
 - génie logiciel et programmation
 - imagerie numérique
 - informatique quantique
 - informatique théorique
 - intelligence artificielle

[typologie des savoirs](#) > [disciplines](#) > [sciences formelles et expérimentales](#) > [informatique](#)
> **recherche opérationnelle**

PREFERRED TERM**recherche opérationnelle** **BROADER CONCEPT**[informatique](#)**IDENTIFIER**

doi:10.34817/sv-oyamg5jb

IN OTHER LANGUAGES

Operations Research

investigación operativa

URI<http://data.ehess.fr/thes/savoirs/doi:10.34817/sv-oyamg5jb> **Download this concept:**[RDF/XML](#) [TURTLE](#) [JSON-LD](#)

Created 09/02/2020, last modified 09/02/2020

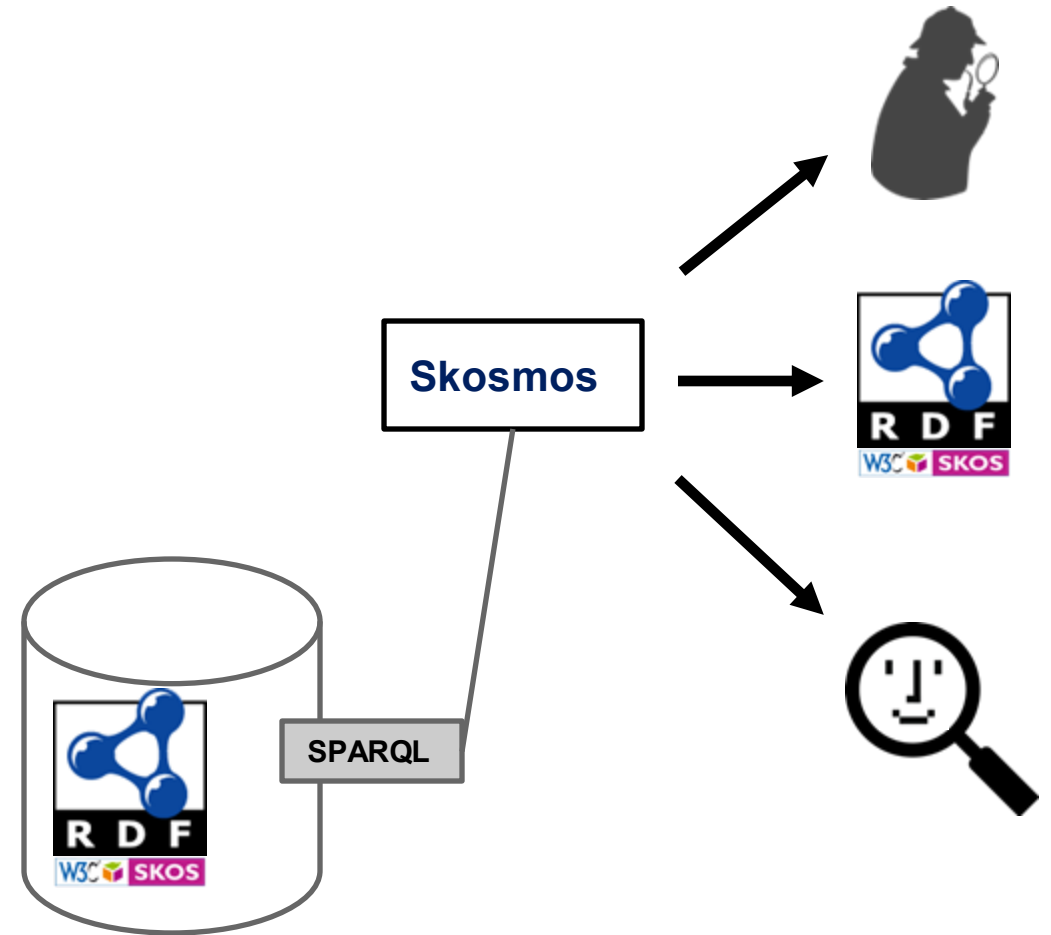
ALTERNATIVE URI<https://doi.org/10.34817/sv-oyamg5jb>

doi.org

Architecture

- Database: RDF triple store with SPARQL
 - usually Apache Jena Fuseki
 - text index (jena-text) for efficient searching
- PHP backend
 - generates HTML pages for humans/browsers and search engines
 - provides RDF linked data
 - REST API for integrations
- JS/HTML/CSS frontend
 - client side dynamic functionality

Skosmos is a traditional web application with server side generation. Not a Single Page Application (SPA)!



History of Skosmos

- 2011 Development of ONKI Light started
at the Semantic Computing Research Group (SeCo)
(predecessors ONKI 1, ONKI 2 and ONKI 3 were developed earlier)
- 2012 First publication "ONKI Light on SPARQL" (EKAW 2012)
- 2013 Maintenance handed over to the National Library of Finland.
- 2014 Finto.fi launched. Renamed to Skosmos. Code moved to GitHub.
- 2015 Skosmos 1.0 released. Skosmos-users group started.
- 2018 Skosmos 2.0 released: new configuration format and Docker support.
- 2022 Planning for Skosmos 3 started. New visual design created.
- 2023 First Skosmos 3 development sprints

Why build Skosmos 3, a new major version?

1. Many UI components used in Skosmos 2 are no longer maintained, so we need to replace them
2. The Skosmos 2 frontend doesn't use any framework except jQuery, which has led to a messy JavaScript codebase
3. Changing the above requires a full front-end rewrite
4. This is an opportunity for improvements in many areas:
 - a. Reduce dependencies for easier maintenance
 - b. Better JS, HTML, CSS and PHP code quality
 - c. Introduce automated frontend testing
 - d. Improve UX and accessibility

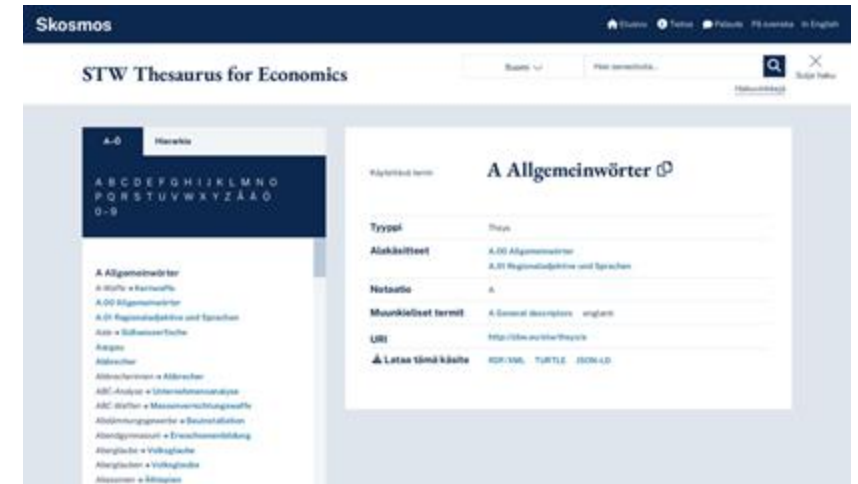
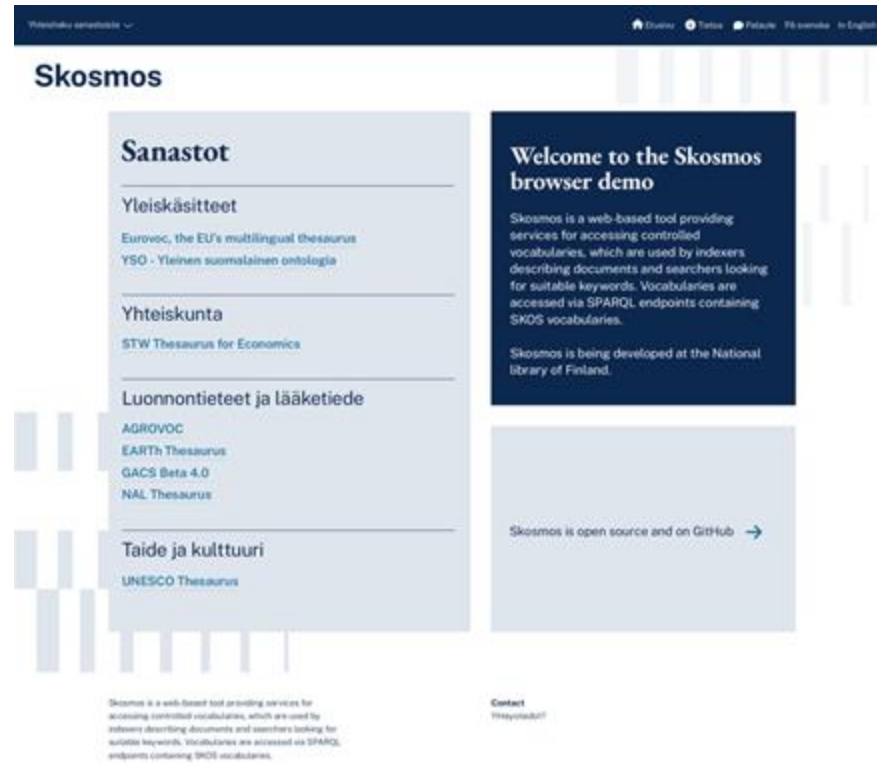
Overview of changes in Skosmos 3

1. New visual style
2. Reimplemented frontend using Vue.js
3. Small changes to the PHP backend
4. Localization using Symfony Translations and Lokalise

What hasn't changed:

- **core functionality remains the same**
- **architecture remains the same**

New visual style and logo



Demo time!

Current Skosmos 3 development version (tracking the main branch) always available at <https://test.dev.finto.fi>

Yhteishaku skosmosista v

Vocabularies About Feedback suomi en fi svenska skosmos

Skosmos

Available vocabularies and ontologies

General

- Allärs - General thesaurus in Swedish
- HCLCS - Helsinki City Library Classification System
- KOKO II/22
- KOKO III/22
- KOKO Ontology
- Metatietosanasto
- OMA Ontology (M-Cult)
- OMA Ontology (M-Cult) - test
- PLC - Finnish Public Libraries Classification System
- Pondus categories

Welcome to the Skosmos browser demo

Skosmos is a web-based tool providing services for accessing controlled vocabularies, which are used by indexers describing documents and searchers looking for suitable keywords. Vocabularies are accessed via SPARQL endpoints containing SKOS vocabularies.

Skosmos is being developed at the National library of Finland.

Skosmos

Vocabularies About Feedback suomi en fi svenska skosmos

YSO - General Finnish ontology

English Search

Alphabetical Hierarchy Groups New and Deprecated

- conferences
- conflicts (societal events)
- congestion
- conscriptions
- consolidations of municipalities
- corrosion facts
- counter-reformation
- coups d'état
- creep tests
- crises
- cultural events
- dangerous situations
- dear's visitations
- decolonisation
- delay
- demonstrations (activities)
- demonstrations (presentations)
- disintegration (groups and communities)
- economic events
- elections
- electric power failures
- emergency landings
- emergency situations
- error conditions
- Euroaidan
- events related to mutual activity
- examinations (education)
- examinations (evaluation)
- exchanges of territory
- exhibitions
- experiments
- exploratory expeditions
- fires and explosions

events and action • events • societal events • demonstrations (presentations)

Preferred term

demonstrations (presentations)

Type

General concept

Broader concept

societal events

Related concepts

presentations (introductions)

Belongs to group

GO General Terms

In other languages

demonstration	Finnish
demonstrational	Northern Sami
demonstrationer	Swedish
(törelvinnings)	

URI

<http://www.ys.fi/ontology/2052>

Download this concept

RDF/XML Turtle JSON-LD

Created 5/4/05, last modified 5/8/23

Closely matching concepts

demonstration	www.wikidata.org
---------------	--

Exactly matching concepts

demonstrationer	Allärs - General thesaurus in Swedish
(törelvinnings) (sv)	

Localization: Skosmos UI in 21 different languages

- English is the base/default language
- Finnish, Swedish and Northern Sámi translations maintained by NLF
- translations for other languages crowdsourced using Transifex (so far)

For Skosmos 3, we're setting up **Lokalise** for managing translations.



Skosmos-3

DONE 92%

BASE WORDS 451

TEAM 1

KEYS 223

QA ISSUES 5843

OTA USAGE (GB) -

APPS

📁

🔄

📦

Install apps

📧

👤

⬇️

⋮

🗑️

English

DONE 58%

UNVERIFIED 224

Arabic

DONE 89%

WORDS TO DO 46

Chinese Simplified

UNVERIFIED 48

DONE 132%

Danish

UNVERIFIED 3

DONE 89%

WORDS TO DO 46

UNVERIFIED 49

Dutch

DONE 96%

WORDS TO DO 17

UNVERIFIED 21

Finnish

DONE 100%

UNVERIFIED 5

French

UNVERIFIED 50

DONE 96%

WORDS TO DO 17

UNVERIFIED 21

Frisian

DONE 100%

UNVERIFIED 14

German

DONE 100%

UNVERIFIED 12

Italian

DONE 79%

WORDS TO DO 91

UNVERIFIED 53

Latvian

UNVERIFIED 14

DONE 100%

Northern Sami

DONE 100%

UNVERIFIED 9

Norwegian Bokmål

DONE 96%

WORDS TO DO 17

UNVERIFIED 22

Norwegian Nynorsk

DONE 96%

WORDS TO DO 17

UNVERIFIED 22

Persian

UNVERIFIED 22

DONE 89%

WORDS TO DO 46

UNVERIFIED 50

Polish

DONE 81%

WORDS TO DO 85

UNVERIFIED 50

Portuguese

DONE 100%

UNVERIFIED 12

Portuguese (Brazil)

DONE 98%

WORDS TO DO 6

UNVERIFIED 15

Russian

UNVERIFIED 30

DONE 64%

WORDS TO DO 161

UNVERIFIED 87

Spanish

DONE 64%

WORDS TO DO 161

UNVERIFIED 87

Swedish

DONE 100%

UNVERIFIED 6

Add language

Skosmos on GitHub

Code repository

- main branch is Skosmos 3
- skosmos-2 branch is Skosmos 2

Issues and pull requests

Documentation in the wiki

The image shows two screenshots of the Skosmos GitHub repository. The top screenshot is the repository's main page, displaying the repository name 'Skosmos', a search bar, and navigation links like Code, Issues, Pull requests, Discussions, Actions, Projects, Wiki, Security, Insights, and Settings. It also shows statistics: 130 branches, 74 tags, 244 issues, 28 pull requests, 98 forks, and 224 stars. A list of files is visible, including 'github', 'ansible', 'bin', 'config', 'dockerfiles', 'plugins', 'resource', 'src', 'tests', '.codeclimate.yml', '.dockerignore', '.env', and '.gitignore'. The bottom screenshot is the README page, which provides an overview of Skosmos as a web-based tool for accessing controlled vocabularies. It mentions that Skosmos is used as a basis for the Finto vocabulary service and is implemented using PHP, Twig, jQuery, and jsTree. It also notes that the code is open source under the MIT license and provides links for installation, release notes, and reporting issues.

Development of Skosmos 3

- Most work is done in Scrum-style **sprints**
 - 2 weeks (only Tuesdays to Thursdays) per sprint + planning & review
 - five sprints in 2023, six in 2024, six in 2025 (two to go)
- Development **team**
 - currently 4 developers (Joeli, Mika, Osma, Unni)
 - occasional consultants (Bruno, Henri)
 - product owner (previously Mikko, now Mona)
 - user representatives (Mirja, Tuomas, Maria) in sprint planning & reviews
- Coordinated using GitHub issues and project boards

Roadmap 2025–2026

3.0-alpha.1 Released in March 2025

3.0-alpha.2 Released in May 2025

3.0-beta.1 Released in August 2025

3.0-beta.2 To be released in November 2025

3.0 To be released in December 2025, with core
functionality

2026: ...more functionality...

Customization and plugins needed for Finto.fi

What does it take to deploy Skosmos?

Hardware & Software:

- a Linux server with modest specifications
- basic web server software: Apache, PHP
- Apache Jena Fuseki triple store (open source)
- git version control utilities

Data:

- one or more controlled vocabularies as SKOS/RDF files

Skills:

- Installing & maintaining the above
- Basic RDF skills, e.g. loading SKOS data into the triple store
- optional: web skills (e.g. CSS) for customizing Skosmos
- being able to read our excellent wiki documentation ;)

Skosmos-users Google Group

Welcome to the [Skosmos](#) users' mailing list / web forum! This list can be used for

- general discussion about Skosmos, its features and usage scenarios
- asking for help with installing or running Skosmos
- future directions for Skosmos
- announcements for new versions and other Skosmos-related news



	Osma Suominen	ANN: Skosmos 3.0-beta.1 released — Skosmos 3.0-beta.2 has been released! https://github.com/NatLibFi/S...	28.8.	
	Natuurfilmpjes	Comparing skosmos library against other libraries — Hi, I am using the Skosmos REST API to get vocabulari...	7.8.	
	Omid Ghasvand	Translations not passed — Hello, I am using the old interface of Skosmos. Everything works fine, but the prob...	20.7.	
	Osma Suominen	ANN: Skosmos 3.0-alpha.2 — Skosmos 3.0-alpha.2 has been released! https://github.com/NatLibFi/Skosmo...	26.5.	
	Osma Suominen	ANN: Skosmos 3.0-alpha.1 — Skosmos 3.0-alpha.1 has been released! https://github.com/NatLibFi/Skosmo...	20.3.	
	Osma Suomin... , ... Joeli Taka... 3	Re: Custom properties - not able to display — Hello Uldis, No not really. Skosmos won't show the hierarchy if ...	5.3.	
	Osma Suominen	Apologies for the spam — Hi all, apologies for the spam messages that were posted through skosmos-users ...	24.1.	
	opo...@gmail.... , Osma Suomi... 2	Hierarchy disappears when groups are introduced — Hi Oddrun, can you please show the configuration you u...	23.1.	

Skosmos online workshop at SWIB25

"From chaos to Skosmos: publishing controlled vocabularies made easy"

Monday November 17th, 9:00–13:00 CET (probably a bit shorter)

More information and registration:

<https://forum.swib.org/t/from-chaos-to-skosmos-publishing-controlled-vocabularies-made-easy/1454>

The target audience of this workshop are people working with vocabulary services and knowledge organization systems such as thesauri, classifications, or ontologies. We welcome current Skosmos users as well as those who would be interested in taking Skosmos into use.



Questions about Skosmos?

before we move on to Annif

Part II:

Annif and Finto AI: automated subject indexing and classification

annif

fintoai

Mona Lehtinen



developed since 2017

- General purpose open source **tool** for automated subject indexing and classification using multiple backends
- Multilingual, supports many vocabularies
- Code on GitHub, website with test form and API (Apache 2.0, except some dependencies)
- Global development and user community; user forum **annif-users** on Google Groups



launched in 2020

- Automated subject indexing **service** for production use, based on Annif. Web user interface and API service
- Supports indexing with the General Finnish Ontology YSO & PLC – Finnish Public Libraries Classification System (in Fin, Swe & Eng) as well as KAUNO – ontology for fiction (in Finnish)
- Intended to support subject cataloguers in Finland regardless of institution (GLAMs, public administration); sister project to the Finto vocabulary service

Various algorithms

Lexical approaches (MLLM)

Look for matches between terms in text and vocabulary

***“Renewable resources** are a part of Earth's **natural** environment and the largest components of its ecosystem.”*

yso:p14146
“renewable natural resources”

Less training data needed

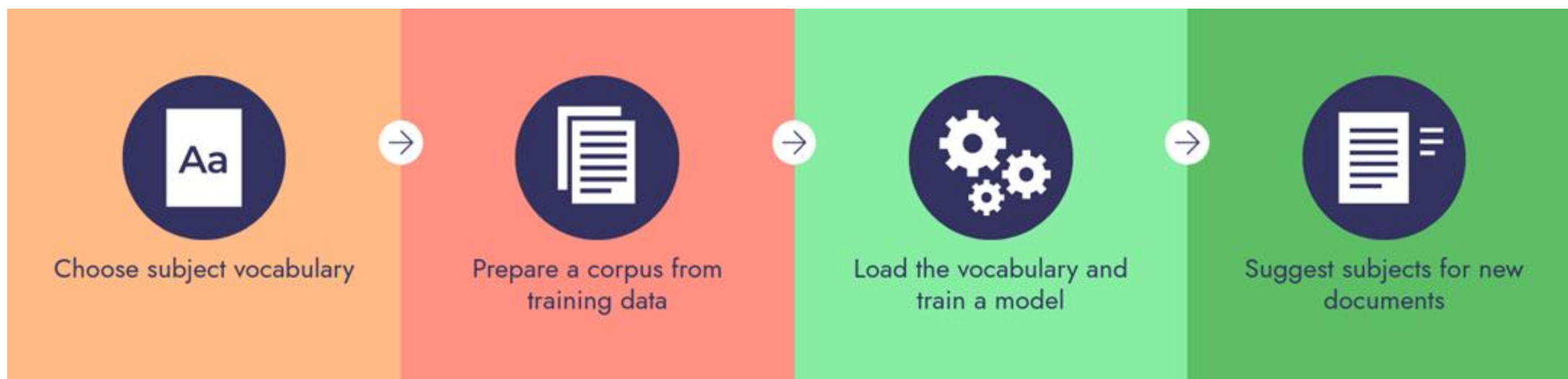
Associative approachest (TF-IDF, fastText, Omikuji)

Learn statistical correlations between terms in the text and vocabulary



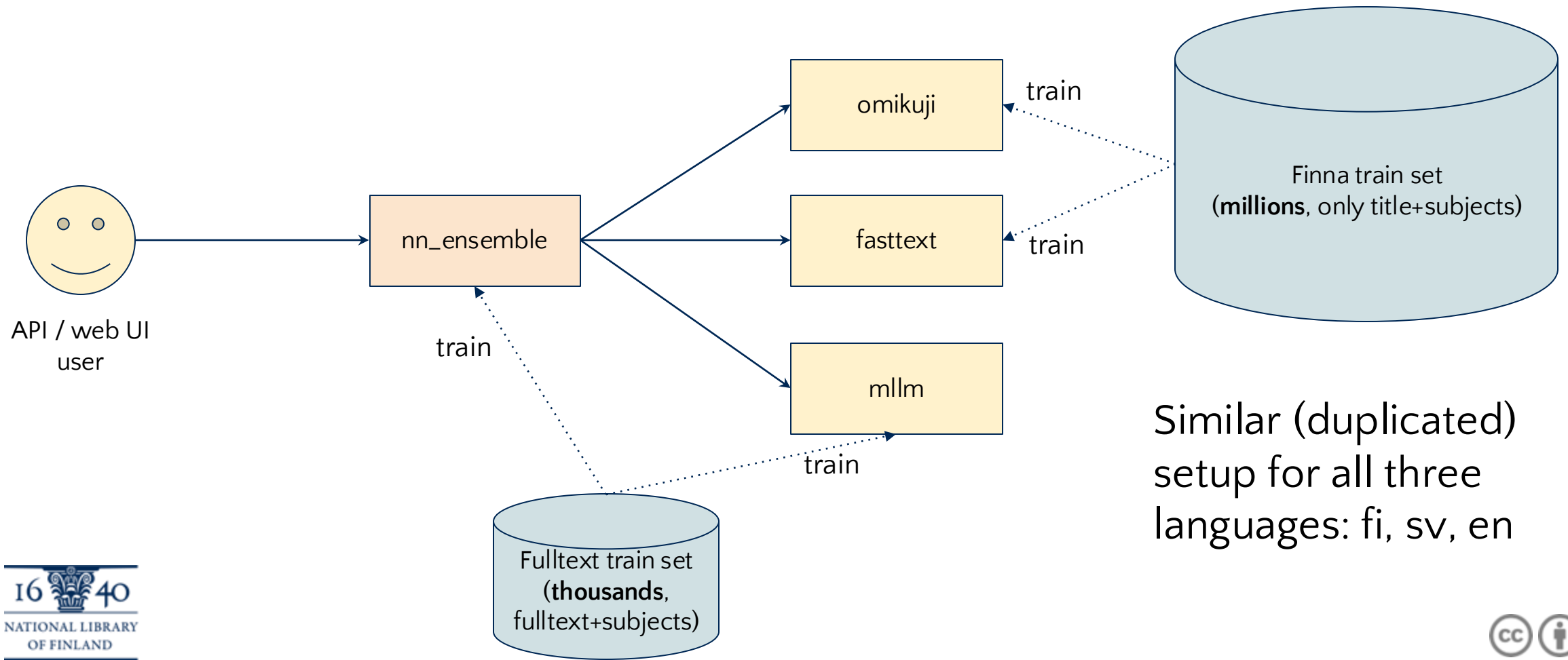
More training data needed

What would it take to use Annif?



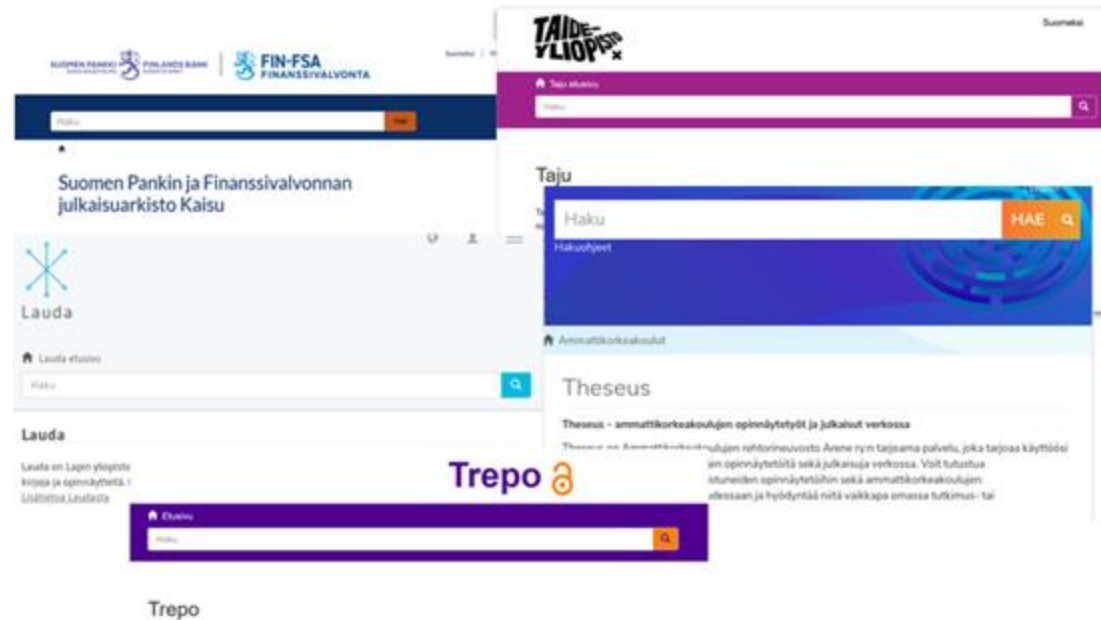
Setting up, training and evaluating models

Current Finto AI set of projects/models for YSO subject vocabulary:



Annif & Finto AI development in-house

- Within the library, Annif development has ties to
 - Ontologies and vocabularies: Finto
 - Repository services
 - Legal deposits
 - Finna discovery service
 - National Bibliography



Annif & Finto AI development & community

- Work with external partners:
 - EU project High Performance Digitisation (Connecting Europe Facility, 2018–2020) implemented in the supercomputing environment provided by the IT Centre for Science CSC → led to major improvement in Annif
 - ZWB Leibniz Information Centre for Economics: contributions to Annif, Annif-tutorial
 - We also work closely with other cultural heritage and research organisations regarding AI

Annif & Finto AI development & community

- We participated in the LLMs4Subjects workshop
- Idea: LLMs for efficient subject indexing, competition for different teams
- In the 1st round, Annif was 1st/2nd in quantitative evals, 4th in qualitative eval (of 14 teams)
- In the 2nd round Annif got 1st place (emphasis on resource consumption)
- we used LLMs for **translation** and **synthetic data generation**
- [SemEval proceedings](#); [our paper](#); [our poster](#); [our 2nd paper preprint](#)
→ New ideas and lessons for Annif development

Annif & Finto AI community & support

- For hands-on details of working with Annif, see the [Annif tutorial](#)
 - The next Annif tutorial / online workshop: [SWIB25](#)
 - November 17th 2025, 18:00–22:00 CET
 - more info and registration: <https://forum.swib.org/t/introduction-to-the-annif-automated-indexing-tool/1456>
 - Materials (videos and exercises) are available for self-study on GitHub and YouTube.
- Annif documentation on [GitHub](#)
- Annif-users community: [Google forum](#)

Annif-users Google Group

Welcome to the [Annif](#) users' mailing list / web forum! This list can be used for

- general discussion about Annif, its features and usage scenarios
- asking for help with installing or running Annif
- future directions for Annif
- announcements for new versions and other Annif-related news



	juho.i...@hels..., ... Christoph ... 5	ANN: Annif 1.4 released — Dear Annif-team, thank you for releasing the new Annif Version. Including the fully...	4.9.	
	Annif Users	Proceedings and Continuation of the LLMs4Subjects Competition — Dear all, As you may remember, we part...	28.8.	
	Thomas Guignard	Annif use cases in Switzerland — Dear Annif community I'm preparing a short presentation on Annif for the L...	21.8.	
	Ball, Lakshmi rb 2	Xtransformer backend — Hi Lucas, Thanks for testing the Xtransformer backend and reporting this. The issu...	15.8.	
	Teemu Nuut..., ... Osma Suom... 3	WSGI service setup fails — Hi Teemu and Uldis, you are both correct, the wiki page on WSGI is outdated and t...	29.7.	
	Jan Ja..., ... anna.k...@google... 4	practice paper 'Automatic Subject Cataloguing at the German National Library' — Dear João Lima, you might...	29.7.	
	Robina M, juho.i...@helsinki.fi 2	XML Support and aksing for best practces — Dear Robina, 1.) Using corpus in XML-format is not possible, y...	16.7.	
	Sven Sass, juho.i...@helsinki.fi 6	X-Transformer backend in nn-ensemble — Hello Juho, just a short follow up - it might be interesting for every...	15.7.	

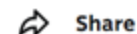
Annif Users



See <https://annif.org> for more

Use case: YLE

- Suominen, O. & Virtanen, P. Yle meets Annif – an open source tool for automated subject indexing. [Presentation](#) at EBU MDN Workshop 2020, 10 June 2020.
- Operates with own installation & vocabulary
 - Weekly updates
- Integrated into workflow in which a journalist submits an article
- Commissioning project took 1-2 years (with a small project group)
- News item: <https://yle.fi/aihe/a/20-10001817>



Yle has started using the Annif tool for the computer-assisted tagging of article content in Finnish and Swedish.

The logo for Annif, featuring the word "annif" in a dark blue, lowercase, sans-serif font. The letter "i" is stylized with a green dot above it.

Use cases & contributors: ZBW and DNB

- Kasprzik, A. (2025). Transferring applied machine learning research into subject indexing practice. In E. Balnaves, L. Bultrini, A. Cox, & R. Uzwyszyn (Eds.), *New horizons in artificial intelligence in libraries* (pp. 199–212). De Gruyter Saur.
<https://doi.org/10.1515/9783111336435-015>
- Kasprzik, A. (2023). Automating subject indexing at ZBW – making research results stick in practice. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, 33(1).
<https://doi.org/10.53377/lq.13579>
- [Team page](#)
- [DNB launches it's cataloguing machine \("Erschließungsmaschine"\)](#)
- Poley, C., Uhlmann, S., Busse, F., Jacobs, J.-H., Kähler, M., Nagelschmidt, M., & Schumacher, M. (2025). Automatic Subject Cataloguing at the German National Library. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, 35(1), 1-29.
<https://doi.org/10.53377/lq.19422>
- Kluge, L., & Kähler, M. (2025). DNB-AI-Project at SemEval-2025 Task 5: An LLM-Ensemble Approach for Automated Subject Indexing. arXiv.
<https://doi.org/10.48550/arXiv.2504.21589>

Use case: DESKRYPTOR / National Library of Poland

**e-usługa
DESKRYPTOR**

**1 | WKLEJ LUB WPISZ
TEKST** ? ✕

W tym miejscu wprowadź tekst do
otagowania

2 | WYBIERZ TRYB ?

pojedynczy zestaw tagów ▼

3 | WYBIERZ MODEL SI ?

nn-ensemble-pl-lem (uniwersalny) ▼

4 | OKREŚL LICZBĘ TAGÓW ?

10 15 20

OTAGUJ TEKST

API | DOKUMENTACJA | MLUMEN | BIBLIOTEKA NARODOWA | KONTAKT | DEKLARACJA DOSTĘPNOŚCI | MAPA SERWISU |

☐ Zamień kolory

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego



Finto AI at NatLibFi

The collage displays five different institutional repository interfaces:

- Kaisu:** Suomen Pankin ja Finanssivalvonnan julkaisuarkisto Kaisu. Features logos for Suomen Pankki, Finland's Bank, and FIN-FSA. Includes a search bar and a home button.
- Lauda:** Lauda on Lapin yliopiston kirjoja ja opinnäytteitä. Li: Lisätietoa Laudasta. Includes a search bar and a home button.
- Trepö:** Trepö on Tampereen yliopiston avoin julkaisuarkisto. Trepö sisältää Tampereen yliopiston avoimia julkaisuja: yliopiston henkilökunnan kirjoittamien tieteellisten artikkelien rinnakkaisallentia ja yliopiston muita avoimia julkaisuja sekä opinnäytteitä. Tampere University Pressin Open Access -kirjat ovat Trepossa omana TUP OA Books -kokoelmana.
- Theseus:** Theseus - ammattikorkeakoulujen opinnäytetyöt ja julkaisut verkossa. Theseus on Ammattikorkeakoulujen rehtorineuvosto Arene ry:n tarjoama palvelu, joka tarjoaa käyttöösi en opinnäytetöitä sekä julkaisuja verkossa. Voit tutustua stuneiden opinnäytetöihin sekä ammattikorkeakoulujen idessään ja hyödyntää niitä vaikkapa omassa tutkimus- tai
- Taju:** Taju etusivu. Includes a search bar and a home button.

Legal deposits

Use case: Finto AI at Finnish museums

 Collecte

Esine

Arkisto

Junat

Haku

Lisää tiedostoja

Hankintaerät

Korit

Osapuolet

Kokoelmat

Muu

Asetukset

Käyttäjät

Lomakkeet

Ontologiat

Organisaatiot

Pudotusvalikot

Roolien oikeudet

Säilytyspaikka

Sarjanumerointi

Statistikka

Muokkaa aineistoa - Junat: JUN-1-1

Tiedostot



Hr1_1021_Rm.jpg

Oletuskuva

Poista

Lisää tiedostoja

Numero *

Tekstina

Sarjallinen

Hankinta

JUN-1.1 (Sarjanumero olemassa)

Nimi

Ukko-Pekka

Kunto

Uudenveroinen

Valmistuspaikat

Objektin kuvailu

Raskas pikajunaveturi Hr1 (ennen vuoden 1942 nimimuutostusta P1) eli tuttavallisemmin Ukko-Pekka oli Valtionrautateiden voimakkain henkilöliikenteen höyryveturi. Tyyppiä rakennettiin kaiken kaikkiaan

Asiasanat

Annif-asiasanat

Jos asiasanoja kirjoitetaan manuaalisesti, lisää vain yksi asiasana per rivi.

veturit

rautatiekalusto

junat

rautatiet

nopeus

höyryveturit

Tampere

henkilöliikenne

raiteet

TONNI

Annif-asiasanat

Hankintaera

Kokoelma

Materiaalit

Lisää uusi materiaali

1640

NATIONAL LIBRARY OF FINLAND



Annif users survey

- Started in Nov 2024, 20 responding organizations
- The overall satisfaction with Annif was 4.4 out of 5
- Satisfaction to documentation gained the highest score of 4.5
- Satisfaction to suggestion quality was 3.6
- Major challenge: gathering and preparing data for training and evaluating
- Inkinen, J., Lehtinen, M., & Suominen, O. (2025). *Annif users survey: Understanding usage and challenges*. Kansalliskirjasto.
<https://urn.fi/URN:ISBN:978-952-84-1301-1>

Evaluation approaches (Golub et al. 2016), **emphasis** ours

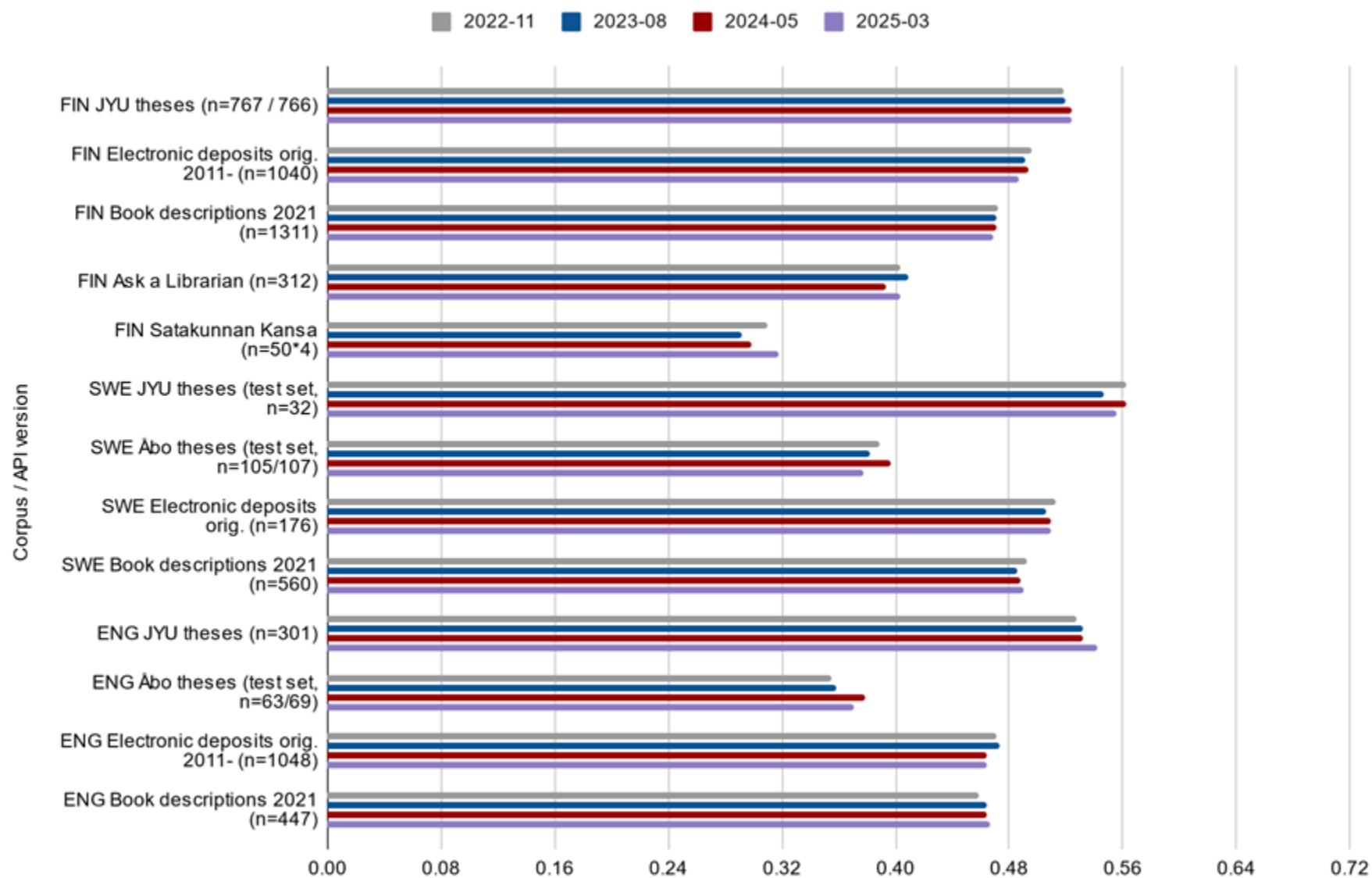
1. Evaluating indexing quality directly through **assessment by an evaluator** or by **comparison with a gold standard**.
2. Evaluating indexing quality directly **in the context of an indexing workflow**.
3. Evaluating indexing quality indirectly through retrieval performance.

The different evaluation approaches are complementary.

Not a good idea to look at just a single measure.

Golub, K., Soergel, D., Buchanan, G., Tudhope, D., Hiom, D., and Lykke, M. 2016. A framework for evaluating automatic indexing or classification in the context of retrieval. *Journal of the Association for Information Science and Technology*, 67(1): 3–16.

F1@5 scores



Evaluation in the workflow context: Theseus

Wrong / irrelevant
suggestions can be left
out

Missing relevant
terms can be added

Annif-ehdotukset

☒ working abroad	☐ families (groups)
☒ career development	☒ managers and executives
☒ career	☐ human resources
☒ adaptation (change)	☐ work
☒ expatriates	☐ returnees (immigrants)

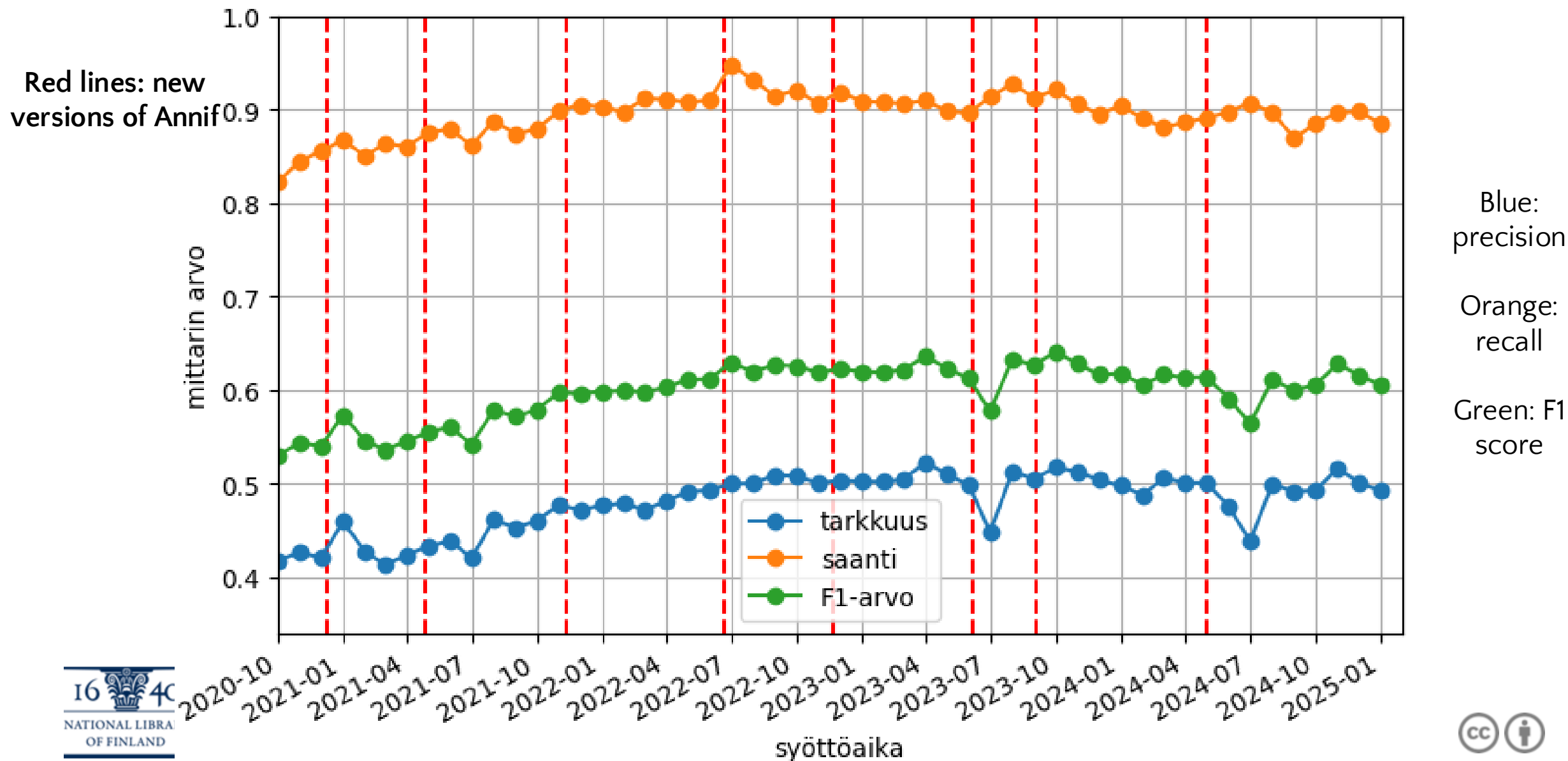
Lisää

working life, teamwork

Lisää

Syötä asiasanat, jokainen asiasana omaan kenttäänsä. Paina siis jokaisen asiasanan jälkeen Lisää-nappia. Kirjoita tarvittava määrä asiasanan alkua, jolloin ennakoiva tekstinsyöttö ehdottaa asiasanoja. Muista myös valita yllä olevasta laatikosta Annif-ehdotukset, jotka perustuvat edellisessä vaiheessa syöttämäsi kokotekstin sisältöön.

Evaluation in the workflow context: Theseus



Try Finto AI!

<https://ai.finto.fi>

.....fintoai

Finto.fi Feedback User guide | suomi på svenska

Finto AI – a service for automated subject indexing based on artificial intelligence. Finto AI suggests subject headings for texts from a vocabulary. The subjects can be utilized to support information retrieval.

Enter text to be indexed

Enter text Upload file Enter URL

Copy or write text here and press the button "Get subject suggestions" ✕

☁️ You can also drop a file or a URL here

Powered by Annif v1.1.0

Subject indexing

Vocabulary and text language

YSO Finnish (2024.2.Ikhwan-al-Safa) ▾

Vocabulary: YSO – General Finnish ontology ↗

Maximum # of suggestions

10 15 20

Suggestions language

Same as text language ▾

Get subject suggestions

Lessons learned

- Implementing semi-automated indexing with an API service is easy; explaining it to users can be more challenging
- Community & collaboration are key (engage with end-users and other developers)
- Traditional ML approaches are still valid

Skosmos

finto

annif

fintoai

Thank you!

Osma Suominen, Joeli Pokkinen, Mona Lehtinen, Juho Inkinen

[firstname.lastname]@helsinki.fi