

SIGNUM

SUOMEN TIETEELLINEN KIRJASTOSEURA

SIGNUM SIIRTYY VERKKOON

Jäsenkysely antoi tukea
verkkojulkaisemiselle 2022

TEKOÄLY TULEE

Kirjaston roolit suhteessa
uuteen teknologiaan

KONE KUVAILUA OPPIMASSA

Sisällönkuvailun
automatisointia kehitetään
Kansalliskirjastossa

AJATUKSIA METRIIKASTA

Eva Isaksson käy haastattelussa
läpi työuraansa ja visioi tulevaa

NORDILL 2021

Konferenssi toi esiin
pohjoismaisten kirjastojen
ajankohtaisia teemoja



Julkaisija

Suomen tieteellinen kirjastoseura
(STKS)

<http://www.stks.fi/>

Signum verkossa

<http://journal.fi/signum>

Puheenjohtaja

Päivi Kaiponen

Helsingin yliopiston kirjasto
PL 56, 00014 Helsingin yliopisto
palvi.kaiponen@helsinki.fi
Puh. +358 05 415 1671

Sihteeri

Tuula Hurskainen

Suomen ympäristökeskus /
Tietopalvelu
PL 140, 00251 Helsinki
tuula.hurskainen@syke.fi
Puh. 0295 251 191, 0400 148 862

Signumin päätoimittaja

Juuso Ala-Kyyny
juuso.ala-kyyny@helsinki.fi

Toimitussihteeri

Minna Marjamaa
minna.marjamaa@laurea.fi

Kannet, ulkoasu ja taitto

Petri Fills
mail@petrills.com

Tilaushinnat

Suomi 50€
Muut maat 60€
Irttonumero 6€, lähetysmaksu 4€

Tilaukset

Tarja Hämäläinen
tarja.hamalainen@helsinki.fi

Paino

Grano Oy
ISSN 0355-0036

SIGNUM

4/2021

Tässä numerossa

Pääkirjoitus

3

Reittejä on monta –
tieteellisten kirjastojen
käsitäyksiä tekoälystä

9

Tekoälyohjelma Iris.Aln
alkutaival, käyttäjäkokemukset
ja uudet kilpailijat

21

NordILL 2021 -konferenssi

31

Uutisia lyhyesti

43

Jäsenistö kertoi
näkemysensä Signumista
– verkkojulkaisua
kokeillaan vuonna 2022

4

Automaattisen
sisällönkuvailun ohjelmiston
rakentaminen – case Annif

14

Tekoälyn lukutaito –
valmiuksia, joita jo tarvitsemme

26

Eva Isakssonin
haastattelu

37



SUOMEN TIEDELLINEN KIRJASTOSEURA

FINLANDS VETENSKAPLIGA BIBLIOTEKSSAMFUND THE FINNISH RESEARCH LIBRARY ASSOCIATION

TEKOÄLY JA TIETEELLISET KIRJASTOT



Tekoälyn ja sen keskeisen osa-alueen, koneoppimisen, historia juontaa vuosikymmenien taakse. Hypetermi tekoälystä on tullut viimeisen kymmenen vuoden aikana, kun laskentatehon kasvamisen myötä tekoälyn kehitystyö on laajentunut voimakkaasti.

Tekoälyn älykkyys perustuu todennäköisyyksien laskemiseen annetun datan pohjalta, ja algoritmin justeeramiseen datan muuttuessa. Arkiset nykysovellukset ovat ehdotuksia käyttäjille: Ostit juuri muistitikun, ja käyttäjätiedotamme perusteella saattaisit olla kiinnostunut myös sankakuulokkeista tai säiliöpakastimesta.

Verkkomaailmassa tekoälysovelluksia on jo kaikkialla. Esimerkiksi hakukoneissa ja sosiaalisen median palveluissa käyttäytymistämme ohjataan varovaisin ehdotuksin tai suoranaisiin käyttöä koskeviin rajauksiin. Siellä missä on paljon dataa, on ainakin potentiaalia tekoälysovelluksille. Tämän takia esimerkiksi kirjastot ovat luontevia kumppaneita teknologiayrityksille

* * *

Tekoälyllä on historiaa myös kirjastoissa. Heli Kautosen (Suomalaisen Kirjallisuuden Seura) ja Andrea Gasparinin (Oslo yliopiston kirjasto) Liber Quarterlyn tekemässä laajassa – vielä julkaisemattomassa – kirjallisuuskatsauksessa varhaisin kirjaston tekoälyvisio oli vuodelta 1976.

Tekoäly on mukana yhä useammassa tieteellisten kirjastojen palveluissa, mutta kenties vielä enemmän puhuttaa tekoälyn tulevaisuus kirjastoissa. Miten siihen pitäisi suhtautua?

Lyhyt, välitön vastaus voisi olla: pragmaattisesti ja ymmärtäen, mistä tekoälyssä on kyse. Kirjaston näkökulmasta tekoälyyn voi liittyä mahdol-

lisuuksia esimerkiksi kokoelmatyön, palveluiden ja kirjastotilojen kehittämisessä, ja tämän voi nähdä luontevana osana digitaalisen kirjaston ja kirjastoautomaatiikan jatkumoa.

Yhtä lailla on tärkeää ymmärtää tekoälyn ohjaamaan tai rajaamaan käyttöön liittyvät ongelmat, kuten erilaisten vinoumien ja stereotyyppien vahvistaminen. Tekoäly ei vain lue dataa, se myös tuottaa ja uusintaa todellisuutta.

* * *

Edellä kirjoitetut asiat nousevat esiin toisaallakin tässä lehdessä, sillä vuoden viimeinen numero on omistettu neljän artikkelin osalta tekoälylle. Otos kattaa kirjastokentän eri näkökulmista.

Heli Kautonen ja Pirjo Kangas johdattelevat teemaan, ja pohtivat kirjaston mahdollisia rooleja ja asennoitumisia suhteessa tekoälyyn. Taustalla on muun muassa tuo edellä mainittu kirjallisuuskatsaus.

Mikko Lappalainen ja kumppanit luovat pintaa syvemmän katsauksen koneoppimisen soveltamiseen kirjastotyön ytimessä eli sisällönkuvaailussa, ja taustoittaa samalla automaattisen sisällönkuvaailutyökalun Annifin rakentamista.

Kirsi Salmi, Jukka Englund ja Maarit Putous tarkastelevat tutkijoiden tiedonhakuun suunniteltua Iris.ai-tekoälyohjelmaa niin yrityksen ja kirjaston välisen kumppanuuden kuin käytännön tiedonhaunkin näkökulmasta.

Heidi Enwald ja Noora Hirvonen pohtivat tekoälyn lukutaitoa käsittelevässä artikkelissaan, millaisia näkökulmia tekoälyteknologioiden käytön ja vaikutusten tunnistamiseen liittyy kirjastojen kannalta.

Johanna Lahikainen ja Juuso Ala-Kyyny

Jäsenistö kertoi näkemyksensä Signumista – verkkojulkaisua kokeillaan vuonna 2022

Suomen tieteellinen kirjastoseura (STKS) toteutti touko-kesäkuussa kyselyn, jossa kartoitettiin jäsenistön näkemyksiä seuran lehdestä, Signumista, erityisesti sen julkaisemiseen painettuna tai elektronisena. Vastausten perusteella lehteä pidettiin tärkeänä jäsenetuna, mutta sen julkaisumuotoa ei nähty yhtä keskeiseksi asiaksi esimerkiksi jäsenyyden kriteerinä. Jäsenkyselyn vastausten perusteella STKS:n hallitus päätti syksyllä, että elektronista julkaisemista kokeillaan vuonna 2022 siten, että Signum ilmestyy verkkolehtenä (kolme numeroa) ja painettuna lehtenä (yksi numero).

Viime vuosina Suomen tieteellisen kirjastoseuran (STKS) hallituksessa on keskusteltu Signum-lehden tulevaisuuden vaihtoehtoista. Jäsenistön keskuudessa Signum on nähty korkeatasoisena kirjasto- ja informaatioalan julkaisuna ja tärkeänä jäsenetuna.

Alati nousevat postitus- ja painatuskustannukset rasittavat kuitenkin seuran taloutta, ja tämä kehitys – yhdessä jäsenmäärän kehityksen kanssa – on saanut hallituksen pohtimaan siirtymistä pelkästään verkossa ilmestyvään lehteen. Puheenjohtaja Päivi Kaiponen ja varapuheenjohtaja Sinikka Luokkanen avasivat tilannetta keväällä 2021: ”Suurin osa työelämän ulkopuolella olevan jäsenen jäsenmaksusta ja yli puolet varsinaisen jäsenen jäsenmaksustakin menee jäsenlehti Signumin

kustannuksiin.” (Kaiponen & Luokkanen 2021)

Jäsenkysely touko-kesäkuussa

Hallitukselle oli selvää, että mahdolliset tulevat päätökset tarvitsevat jäsenistön kannatuksen. Hallitus halusi selvittää kyselyllä lehden merkitystä seuran jäsenille sekä jäsenistön mielipiteitä lehden ilmestymismuodosta. Olisiko jäsenistön toiveissa saada jatkossakin painettu Signum vai saisiko vain sähköisessä muodossa ilmestyvä lehti enemmän kannatusta? Signumia koskeva jäsenkysely toteutettiin 26. 5.–30. 6. 2021 ja siihen saatiin 79 vastausta. Kiitos kaikille vastanneille!

Kyselystä viestittiin seuran verkkosivuilla, sähköpostilistalla, seuran sosiaalisen median palveluissa (Facebook ja Twitter) sekä päätoimittaja Juuso

Ala-Kyynyn kirjoittamassa Signumin pääkirjoituksessa numerossa 2/2021. Kyselyssä tiedusteltiin seuraavia asioita:

1. Kuinka kauan olet ollut STKS:n jäsen?
2. Oletko työelämässä vai työelämän ulkopuolella?
3. Kuinka tärkeänä jäsenetuna pidät Signum-lehteä?
4. Kuinka tärkeänä jäsenetuna pidät kotiin kannettua, painettua Signum-lehteä?
5. Mikäli Signum olisi vain elektroninen lehti, miten se vaikuttaisi haluusi pysyä seuran jäsenenä?
6. Olisitko valmis maksamaan korkeampaa jäsenmaksua, jos saat painetun Signum-lehden?
7. Terveisesi Signumin toimitukselle?
8. Terveisesi STKS:n hallitukselle?

Kyselyn tulokset

Kyselyn keskeiset tulokset on koottu seuraavalle sivulle. Kyselyyn vastaajilta pyydettiin taustatiedoiksi tieto jäsenyyden pituudesta ja siitä, oliko vastaaja työelämässä vai sen ulkopuolella. Vastaajista 50 henkilöä (63,3 %) oli ollut seuran jäsenenä yli kymmenen vuotta, 17 henkilöä (21,5 %) viidestä kymmenen vuotta ja 12 henkilöä (15,2 %) alle viisi vuotta. 74 vastaajaa (93,7 %) oli työelämässä ja viisi (6,3 %) työelämän ulkopuolella.

Signum-lehteä piti erittäin tärkeänä 26 vastaajaa (32,9 %) ja melko tärkeänä

jäsenetuna 40 vastaajaa (50,6 %). Yhteensä siis lehteä piti joko erittäin tai melko tärkeänä 66 vastaajaa (83,5 %). Yhdeksälle vastaajalle (11,4 %) lehti ei ollut kovin tärkeä ja kahdelle (2,5 %) se ei ollut lainkaan tärkeä. Kaksi vastaajista (2,5 %) valitsi ”en osaa sanoa”.

Seuraavaksi kysyttiin painetun, koettiin kannettavan Signum-lehden merkityksestä. Erittäin tärkeänä jäsenetuna sitä piti 11 vastaajaa (13,9 %) ja melko tärkeänä 19 (24,1 %). Ei kovin tärkeäksi sen näki 25 vastaajaa (31,6 %) ja ei lainkaan tärkeäksi 15 (19 %). Yhdeksän vastaajaa (11,4 %) valitsi vaihtoehdon ”en osaa sanoa”.

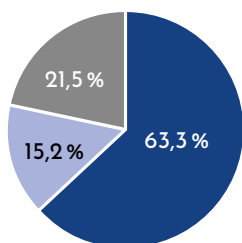
Kyselyllä haluttiin selvittää myös sitä, vaikuttaisiko Signumin mahdollinen muuttuminen vain sähköiseksi lehdeksi vastaajien haluun jatkaa seuran jäsenenä. 49 vastaajaa (62 %) kertoi, ettei se vaikuttaisi jäsenyyshalukkuuteen. 12 henkilöä (15,2 %) koki sen vaikuttavan erittäin myönteisesti ja kymmenen melko myönteisesti. Seitsemän jäsentä (8,9 %) kertoi, että se vaikutti haluun pysyä jäsenenä melko kielteisesti ja yksi (1,3 %) erittäin kielteisesti.

Seuran hallitus halusi tietää, olisivatko jäsenet valmiita maksamaan korkeampaa jäsenmaksua painetusta lehdestä. 52 jäsentä (65,8 %) kertoi, ettei olisi valmiita maksamaan enempää. 23 vastaajaa (29,1 %) oli valmis maksamaan kymmenen euroa enemmän ja neljä (5,1 %) voisi maksaa 15 euroa enemmän, jotta saisi lehden painettuna.

Signum-jäsenkyselyn 2021 vastaukset

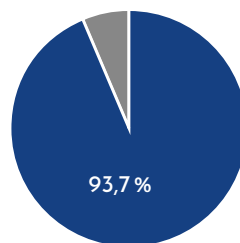
Kysely toteutettiin STKS:n jäsenille 26. 5.–30. 6. 2021,
ja siihen saatiin 79 vastausta.

Kuinka kauan olet ollut STKS:n jäsen?



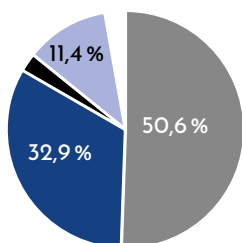
- Alle 5 vuotta
- 5-10 vuotta
- yli 10 vuotta

Oletko



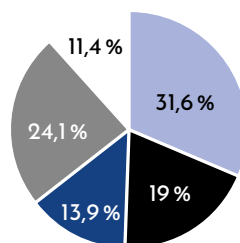
- Työelämässä
- Työelämän ulkopuolella

Kuinka tärkeänä jäsenetuna pidät Signum-lehteä?



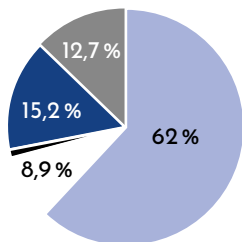
- Erittäin tärkeänä
- Melko tärkeänä
- Ei kovin tärkeänä
- Ei lainkaan tärkeänä
- En osaa sanoa

Kuinka tärkeänä jäsenetuna pidät kotiin kannettua, painettua Signum-lehteä?



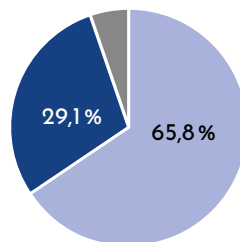
- Erittäin tärkeänä
- Melko tärkeänä
- Ei kovin tärkeänä
- Ei lainkaan tärkeänä
- En osaa sanoa

Mikäli Signum olisi elektroninen lehti, miten se vaikuttaisi haluusi pysyä seuran jäsenenä?



- Erittäin myönteisesti
- Melko myönteisesti
- Ei vaikuttaisi
- Melko kielteisesti
- Erittäin kielteisesti

Olisitko valmis maksamaan suurempaa jäsenmaksua, jos saat painetun Signum-lehden?



- Kyllä, olisin valmis maksamaan 10 € enemmän
- Kyllä, olisin valmis maksamaan 15 € enemmän
- En olisi valmis maksamaan enempää

Terveisiä Signumille ja STKS:n hallitukselle

Monivalintakysymysten jälkeen oli kaksi kysymystä, joihin sai vastata vapaasti kirjoittamalla. Niissä pyydettiin (1) terveisiä Signumin toimitukselle ja (2) terveisiä STKS:n hallitukselle.

Näitä avovastauksia tuli ilahduttavan paljon: yhteensä 48 kappaletta, joista toimitukselle 32 ja hallitukselle 16.

Vastaajat antoivat avovastauksissa kehuja ja kiitosta päätoimittajalle ja toimitukselle. Risujakin tuli: toivottiin etteivät kirjoittajat olisi niin usein samoja henkilöitä. Lehden ulkoasua sekä kiitettiin että moitittiin. Osasta se oli vanhanaikainen ja tylsä, osa taas koki sen olevan uudistunut ja tyylikäs.

Lehden tehtävää ja sisältöä pidettiin enimmäkseen tärkeänä. Yksi vastaajista kirjoittaa:

Lehden perustehtävän (= tiedon välittäminen) lisäksi Signum tarjoaa suomalaisille tieteellisten kirjastojen ammattilaisille mahdollisuuden reflektoida työtään. Kynnys suomenkielisen tekstin kirjoittamiseen (puoli)tutuille kollegoille on paljon matalampi kuin tekstin sorvaaminen kv-journaliin.

Yksi vastaaja kommentoi, että jos seuran jäsenetuna ei enää ole painettua Signumia, tarvitaan jotain muuta etua tilalle.

Verkkoon siirtyminen jakoi mielipiteitä:

Digilehti riittäisi hyvin ja olisi nykyaikaa.

Lehden sähköistäminen sinänsä OK, mutta kiva olisi esim. yksi painettu koostenumero tai erikoisnumero vuodessa tms., eli olisi harmi, jos painettu lehti kokonaisuudessaan loppuisi.

Pidän paperista Signumia sähköisen ohella arvokkaana jäsenetuna, ja surisin, jos lehti muuttuisi kokonaan elektroniseksi.

Ymmärrän hyvin, jos joudutte luopumaan painetusta.

En tule lukeneeksi verkkolehtiä, en yksinkertaisesti jaksa. Enkä ole edes iäkäs vielä vaan 40 v. Toivottavasti lehti ilmestyy edelleen painettunakin!

Toimitusta ja hallitusta kannustettiin ideoimaan villisti Signumin tulevaisuutta, hankkimaan sponsoreita, etsimään uusia jäseniä erityisesti opiskelijoista ja satsaamaan näkyvyyteen verkossa ja sosiaalisessa mediassa.

Verkkolehti nähtiin mahdollisuutena uudenslaisiin sisältöihin. Kommentteissa ehdotettiin myös julkaisemista verkossa artikkeli kerrallaan, uutiskirjeen aloittamista ja uudesta lehdestä ilmoittamista sähköpostitse.



Signum ilmestyy verkkolehdenä 2022


Jäsenkyselyn vastausten perusteella STKS:n hallitus teki syksyllä päätöksen, että Signum ilmestyy vuonna 2022 verkkolehdenä (kolme numeroa) ja painettuna lehtenä (yksi numero). Jäsenkyselyn katsottiin tukevan elektronisen julkaisemiseen siirtymistä, ja ensi vuosi toteutetaan verkkojulkaisemisen pilotoinnin merkeissä.

”Arvostetun julkaisumme siirtymistä verkkolehdeksi ei tehty kevein mielin, ja hallitus punnitsi tarkoin eri vaihtoehtoja. Katsoimme tärkeäksi, että jäsenet saavat jatkossa välittömästi tiedon, kun uusi Signum on luettavissa verkossa, ja painetun numeron kaavailaan olevan tuhti lukupaketti vuoden aikana esiin nousseista teemoista. Jäsenkyselyn antia hyödynnetään lehden kehittämisessä uuden päätoimittajan

aloittaessa tehtävänsä ensi vuonna”, STKS:n hallituksen puheenjohtaja Päivi Kaiponen sanoo.

Tarkoitus on ensi vuoden aikana kerätä tilastopohjaista dataa verkkolehden käytöstä ja palautetta suoraan käyttäjiltä. Samalla pyritään kehittämään myös tiedotukseen ja viestintään liittyviä keinoja, joilla elektroninen lehti löytää lukijansa.

Verkkolehden numerot ilmestyvät nykyisellä Journal.fi-alustalla osoitteessa <https://journal.fi/signum/>. Painetun lehden luonteesta hallitus kävi vilkasta keskustelua. Keskeisenä ajatuksena on, että painettu lehti toimisi hieman vuosikirjan tapaan kooten yhteen keskeisiä aiheita ja artikkeleita vuoden ajalta.

Seuran hallitus ja Signumin toimistunkunta kiittävät lukijoita palautteesta ja toivovat saavansa sitä jatkossakin. Kehitetään lehteä yhdessä! 

Lähde:

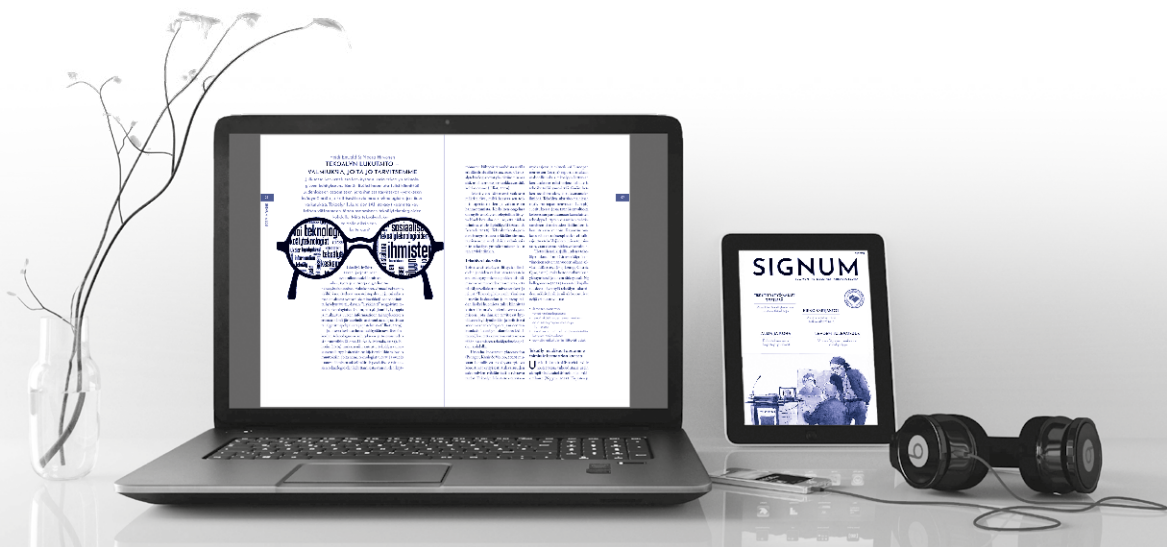
KAIPONEN, PÄLVI & LUOKKANEN, SINIKKA, 2021. Katsaus seuran toimintaan talouden näkökulmasta. Julkaisussa *Signum* 1/2021, <https://doi.org/10.25033/sig.107803>.

JOHANNA LAHIKAINEN

Helsingin yliopiston kirjasto
johanna.lahikainen@helsinki.fi

JUUSO ALA-KYVY

Helsingin yliopiston kirjasto
juuso.ala-kyyny@helsinki.fi



REITTEJÄ ON MONTA – TIETEELLISTEN KIRJASTOJEN KÄSITYKSIÄ TEKOÄLYSTÄ

STKS:n Uuden teknologian työryhmän jäsenet pohtivat syksyllä 2021 pitämässään työpajassa kirjastojen tämänhetkisiä käsityksiä tekoälystä. Uusi ilmiö voi hämmentää, koska siihen liittyy monta eri näkökulmaa, monta roolia ja monta lähestymistapaa. Jostain kannattaa kuitenkin aloittaa ja kokeilemalla oppii. Tätä kautta kirjasto oppii paremmin tunnistamaan prosesseja, joita voi tekoälysovellutusten avulla kehittää.

”Tekoäly... Toivon, että se tekee kaiken tyhjän työn jonain päivänä, niin että kivat työt jää meille.” Näin kuvasi tekoälyä yksi STKS:n Uuden teknologian työryhmän jäsen elokuun työpajassa, jossa pohdittiin kirjastojen tämänhetkisiä käsityksiä tekoälystä. Tässä artikkelissa esitämme synteisin työpajasta ja siellä esitellyn kirjallisuuskatsauksen tuloksista.

Työpajan järjestivät Andrea Gasparini Oslon yliopiston kirjastosta ja Heli Kautonen Suomalaisen Kirjallisuuden Seuran kirjastosta. He esittelivät työpajan aluksi edellisenä vuonna tekemänsä ja pian julkaistavan kirjallisuuskatsauksensa (Gasparini & Kautonen, 2021/2022). Tämän jälkeen työpajaan osallistuneet Uuden teknologian työryhmän jäsenet keskustelivat katsauksessa esiin nousseista teemoista: Miten tekoäly ymmärretään tieteellisissä kirjastoissa? Millaisia rooleja kirjastoilla tai kirjastoammattilaisilla voi ja tulisi olla suhteessa tekoälyyn? Mikä rooli on käyttäjillä, teknologialla itsellään tai suunnittelulla?

Työpajassa esitetyt näkemykset heijastavat samaa ajatusta, jota professori Teemu Roos korosti syksyn 2021

Kirjastoverkkopäivillä: Tekoälyä ei voi eikä kannata määritellä tarkkaan. Yhdelle tekoäly tarkoittaa odotettua kirjastotyön tehostajaa, toiselle hiukan ärsyttävää hypetystermiä. Miksi siis ylipäätään puhua tekoälystä eikä mieluummin yksiselitteisesti ja tarkemmin määriteltävistä sovelluskohteista kuten suosittelupalveluista, kuvantunnistuksesta tai chattiboteista?

Andrean ja Helin kirjallisuuskatsaus osoitti, että tekoäly on käyttökelpoinen yleiskäsite, johon keskustelu uudesta teknologiasta on helppo kiinnittää. Kirjallisuuskatsausta varten tehtiin tiedonhaku laajalla käsitteellä ja se toi analysoitavaksi paljon lähteitä, joista pääosa oli tieteellisiä tutkimusartikkeleita, mutta joukossa oli myös kir-

Role of library/librarians: Main findings

Roles vary from neutral investigator (6/88) or a dutiful professional learning to use new tools (43/88), to a responsible guardian of free and unbiased knowledge (7/88), or even a agent of desirable change (12/88).

“Library can be seen as a valued partner and collaborator, instead of being seen solely as a service provider.” (Burton et al. 2018)

“We must continue to find ways to make the machines work for us.” (Arlitsch & Newell 2017)



Kirjallisuuskatsauksen tuloksia esiteltiin STKS:n uuden teknologian ryhmän syksyn työpajassa.

jastoalan ammattilaisille suunnattuja kannanottoja, uutisia ja projektiraportteja. Yli puolet (93/126) teksteistä oli julkaistu viimeisen kymmenen vuoden sisällä, mutta olipa joukossa varhainen kirjaston tekoölyvisio vuodelta 1976. Näin runsaasta aineistosta oli mahdollista piirtää kattava kokonaiskuva eri suuntiin polveilevista, uusia teknologioita koskevista toiveista, peloista, käytännön kokeiluista ja strategiapohdinnoista. (Gasparini & Kautonen, 2021/2022.)

Kaikkien ei tarvitse olla koodareita

Tekoölyn kaltaisen yleiskäsitteen käyttö jättää tilaa sille, että kirjastoissa voidaan lähestyä aihetta erilaisista rooleista käsin. Teknologiyritykset ja koodarit ovat toki ydinroolissa tekoölyn kehityksessä, mutta myös muiden alojen osaajien kannattaa ottaa asia omakseen. Tekoölyn ja algoritmien ymmärrystä voi tarvita myös tekoölyä hyödyntävien järjestelmien hankintapäätösten tekijä, käyttäjä tai käytön ohjaaja.

Kirjallisuuskatsauksen tulokset osoittavat, että tieteellisten kirjastojen kirjastoammattilaisille hahmotellaan aihetta käsittelevissä teksteissä erilaisia mahdollisia rooleja. Rooleja voi hahmotella toisaalta suhteessa tekoölyyn ja toisaalta suhteessa muutokseen.

Suhteessa tekoölyyn, tulevaisuuden kirjastolaiset voidaan nähdä esimerkiksi ammattilaisina, jotka *käyttävät* tekoölyä hyödyntäviä työkaluja tai palveluja, tai ehkä jopa *kehittävät* niitä. Myös sitä pidetään mahdollisena, että kirjastolaiset kuuluvat ammattiryhmiin, jotka

tulevat ainakin joiltain osin tekoölyn korvaamiksi. (Gasparini & Kautonen, 2021/2022.)

Skenaariot vaihtelevat sen suhteen, kuinka aktiivinen rooli kirjastoammattilaisille varataan. Tekoölyn kehittäjä on selvästi aktiivisempi toimija kuin palveluiden käyttäjä, puhumattakaan tilanteesta, jossa tekoöly korvaisi ammattilaiset. Myös vaadittavan teknologisen ymmärryksen taso vaihtelee eri rooleissa: palvelujen kehittäminen vaatii alan toimijoilta syvempää algoritmista osaamista kuin muut vaihtoehdot.

Tarkkailijana vai muutosagenttina?

Uuden teknologian työryhmän työpajassa tuli esiin erilaisia rooleja suhteessa muutokseen ja myös tässä asiassa kirjastoammattilaisten aktiivisuuden tasolle nähtiin eri vaihtoehtoja. Osaltaan pidettiin mahdollisena, että kirjastoammattilaiset tarkkailevat muutosta enemmänkin ulkopuolelta, mutta myös aktiivisempia rooleja muutosagenttina tai vähintäänkin muutokseen osallistujana pidettiin mahdollisena.

Kirjastoalan järjestöistä IFLA (2020, 2–3) on tehnyt suosituksia siitä, miten kirjastot voivat sopeuttaa palveluitaan tekoölyajan yhteiskunnallisiin tarpeisiin. IFLA pitää mahdollisena, että tekoöly voi merkittävästi vaikuttaa alan työmarkkinoihin ja toteaa, että erityyppiset kirjastot voivat kohdata vaikutukset eri tavoin. IFLA peräänkuuluttaa kirjastoalan järjestöjä tarjoamaan tukea kirjastoalan ammattilaisille uuden äärellä. IFLA näkee keskeisenä käyttäjien



tukemisen tekoälyn ymmärtämisessä elinikäisen oppimisen hengessä ja sen varmistamisen, että kirjastojen käytössä olevat tekoälysovellukset ovat toimintatavoiltaan eettisiä ja turvaavat käyttäjien oikeuksia.

Viime aikoina on käynnistynyt keskustelu kirjastojen roolista tekoälylukutaidon tai algoritmisen lukutaidon levittämisessä. Algoritmisen lukutaito pitää sisällään esimerkiksi perusymmärryksen algoritmeista, niiden käytön tunnistamisen, algoritmisten tekniikoiden soveltamisen sekä kyvyn arvioida algoritmien yhteiskunnallisia vaikutuksia ja vaikutusvaltaa. Nykyisellään esimerkiksi ACRL:n informaatiolukutaidon kehukset eivät riittävässä määrin sisällä näitä näkökulmia. (Ridley & Pawlick-Potts 2021, 4–8). Kanadassa ollaan tartuttu ajatukseen ja eri kirjastotahot ovat yhteistyössä käynnistämässä AI for all -projektia (aiforall.ca), jossa kehitetään yleisten kirjastojen kautta levitettävää algoritmilukutaidon koulutusta. Projektin etenemistä on kiinnostava seurata myös muista maisista ja muilta kirjastosektoreilta käsin.

Kokeilemalla oppii

Isoja pohjoisamerikkalaisia yleisiä kirjastoja edustava Urban Libraries -järjestö on muutama vuosi sitten kasannut johtajille suunnatun muistion tekoälystä. Muistiossa hahmotellaan kirjastoille tekoälyn suhteen roolia, jossa yhdistyy ajatusjohtajuus, kirjastojen luotettavuus tiedon tarjoajana sekä demokratian ja oikeudenmukaisuuden edistäminen. (Urban Libraries 2019, 4.) Muistion keinot tavoitteeseen pää-

semiseksi ovat osin melko yleisellä tasolla, mutta ne ovat hyvää ajateltavaa myös tieteellisille kirjastoille.

Kirjastoja suositellaan rakentamaan toimintaa vahvuuksiensa pohjalle ja keskittymään esimerkiksi digitaaliseen lukutaitoon ja osallistamiseen. Vastaavasti järjestö katsoo, että kirjastojen tulee puhtaasti teknologian sijasta keskittyä siihen, miten tekoäly vaikuttaa kansalaisten elämään. Kirjastojen ei kannata yrittää olla asiantuntijoita kaikilla aloilla vaan luoda kumppanuuksia – tekoälyn kehittämistä ei kuitenkaan pitäisi jättää ainoastaan teknologiayritysten tehtäväksi. Keskeistä on ottaa henkilökunta mukaan. Tämän voi aloittaa esimerkiksi tarjoamalla mahdollisuuksia oppia uutta (Urban Libraries 2019, 4).

Kirjastoalan sisäinen koulutustarjonta on lisääntynyt. Myös Suomessa on vuoden 2021 aikana järjestetty useita kirjastolaisille suunnattuja webinaareja tekoälystä ja uusista teknologioista. STKS:n Uuden teknologian työryhmän webinaarien lisäksi yleisten kirjastojen alueellisen kehittämisen puitteissa on järjestetty useampikin koulutustapahtuma ja myös Kirjasto verkkopäivät tarttui aiheeseen. Lisäksi internetissä on paljon oppimateriaaleja tekoälystä ja koneoppimisesta, joten itsenäiseen oppimiseen on hyvät mahdollisuudet, kunhan vain löytää itselle sopivat sisällöt.

Kansainvälisen AI4LAM-verkoston (Artificial intelligence for libraries, archives and museums) koulutus-työryhmä on kartoittanut verkosta löytyviä kulttuurilaitoksille sopivia



tekoölyyn liittyviä koulutusmateriaaleja. Kooste materiaaleista on tallennettu AI4LAMin verkkosivuille (sites.google.com/view/ai4lam/news/training-resources). Listalta löytyy niin Suomessa toteutettuja Elements of AI -verkkokursseja, webinaaritallenteita vaihtelevista aiheista ja Digital Humanities -tahojen tuottamia materiaaleja. Tarkoituksena oli tutkia materiaalien sopivuutta kulttuurilaitoksille ja saada selville, onko materiaalien sisällöissä aukkoja suhteessa GLAM-organisaatioiden tarpeisiin (AI4LAM 2021a, 2).

Ryhmän mukaan kirjastoissa, arkistoissa ja museoissa on tarpeen kehittää tekoölyyn liittyviä taitoja, sillä tämä parhaiten varmistaa sen, että teknologioita sovelletaan käyttöön sopivalla tavalla ja kriittisellä otteella sekä kulttuurialan kokoelmatyön haasteet huomioiden. Lisäksi ryhmä katsoo, että GLAM-sektorin asiantuntijoilla on laajemminkin yhteiskunnalle annettavaa esimerkiksi koneoppimisen eettisen käytön edistämisessä ja tekoölyn harjoitusdatan hallinnan tukemisessa. (AI4LAM 2021a, 2–4).

Työpajassamme keskustelimme siitä, miten parhaiten varmistaa, että omassa organisaatiossa on riittävästi osaamista uudenlaisten järjestelmien hankkimiseen tai kehittämiseen. Lisäksi pohdimme, miten osaamme tunnistaa tai opimme tunnistamaan prosesseja, joita voi tekoölysovellutusten avulla parantaa. Kirjastojen tekemät omat kokeilut ja niihin tutustuminen ovat tärkeitä myös tästä syystä.

Kirjastot avainasemassa uusien teknologioiden haltuunotossa

Tekoölystä käytävä keskustelu, kirjoitetut tutkimukset ja kannanotot – ja kieltämättä myös tämän artikkelin laatiminen – tuovat mieliimme edellisen digitalisaation tuoman ilmiön eli sen, miten internet ravisteli kirjastoammattilaisten arkea. Olemme uuden kirjastojen digiloikan lähtöviivalla, ja jotkut meistä jo lihaksiaan lämmittelemässä.

Kirjastojen asema tiedeyhteisöissä antaa oivallisen aseman, mutta myös velvollisuuden, tarttua tekoölyn tuomiin haasteisiin ja mahdollisuuksiin. Kirjastoammattilaisten vankka asiantuntemus tiedon saavutettavaksi tekemisestä, lähdekritiikistä, tiedon tuottamisen etiikasta ja nykyisistä digitaalisista välineistä luo hyvän pohjan tekoölyilmiön haltuunottoon.

Kirjastot ovat tiedeyhteisöissään sopivissa solmukohtissa: onhan kirjastoilla rajapinta niin käyttäjiin, opetukseen, tietohallintoon kuin yliopiston ulkopuolisiin tieteen ekosysteemien toimijoihin kuten kustantajiin. Kirjastoilla on hyppysissään myös tekoölyn kehittämiseen tarvittavaa lähtödataa (Kangas, 2020). Jotkut viittaavatkin tieteellisten kirjastojen merkitykseen tekoölyyn liittyvän tiedepolitiikan rakentajina (Henry, 2019).

Vuonna 2017 MIT:n kirjastonjohtaja Chris Bourg kannusti kirjastojen johtajia kohtaamaan pelottomasti tämän vaikealta tuntuvan uuden ilmiön. Viitaten moniin tekoölyä koskeviin riskeihin ja uhkiin, etenkin tieteellisen tiedon ja tutkimuksen näkökulmasta,

hän kannustaa kirjastoja ottamaan roolia ”demonien kesyttäjinä”.

Kansallisia ja kansainvälisiä ammattilaisten ryhmiä seuraamalla näkee, että yhä useampi kirjasto valmistautuu tavalla tai toisella uuteen digiloikkaan. Kuten työpajamme ja siellä esitelty kirjallisuuskatsaus osoittivat, meillä on monta vaihtoehtoista polkua ottaa tekoälyä haltuun. Jokainen voi tehdä jotain ja aloittaa voi vaikka osallistumalla STKS:n Uuden teknologian työryhmän webinaareihin tai liittymällä ryhmän jäseneksi. ❖

*Meidän kirjastossamme pitäisi olla strategia...
tai joku käsitys siitä, miten tekoälyä voi hyödyntää.*
(Työpajan osallistuja, 2021.)

Lähteet

- AI4LAM. 2021a. *AI training resources for GLAM - a snapshot - for public comment*. October 2021. Saatavilla: <https://docs.google.com/document/d/1FMoQVhfXPBj9yWYmkCLjrpvLnQzKKjAXX6Fjo6Cndk/edit#> (11. 11. 2021)
- BOURG, C., 2017. *What happens to libraries and librarians when machines can read all the books?* Saatavilla: <https://chrisbourg.wordpress.com/2017/03/16/what-happens-to-libraries-and-librarians-when-machines-can-read-all-the-books/> (20. 6. 2020)
- HENRY, G., 2019. Research Librarians as Guides and Navigators for AI Policies at Universities. *Research Library Issues*, 299, 47–66.
- IFLA 2020., IFLA Statement on Libraries and Artificial Intelligence. Saatavilla: https://www.ifla.org/wp-content/uploads/2019/05/assets/faife/ifla_statement_on_libraries_and_artificial_intelligence.pdf (11. 11. 2021)
- KANGAS, P., 2020. Mikä on tekoälyn seuraava siirto? *Kirjastolehti* 29. 6. 2020. Saatavilla: <https://suomenkirjastoseura.fi/kirjastolehti/mika-on-tekoalyn-seuraava-siirto/> (11. 11. 2021)
- GASPARINI, A. & KAUTONEN, H., [2021/2022]. Understanding Artificial Intelligence in Libraries. *Liber Quarterly* [vertaisarvioitavana].
- RIDLEY, M. & PAWLICK-POTTS, D., 2021. Algorithmic Literacy and the Role for Libraries. *Information technology and libraries*. June 2021. Saatavilla: <https://doi.org/10.6017/ital.v40i2.12963> (11. 11. 2021)
- Urban Libraries. 2019. AI Leadership Brief. Saatavilla: https://www.urbanlibraries.org/files/AI_Leadership-Brief.pdf (12. 11. 2021)

HELI KAUTONEN
Suomalaisen Kirjallisuuden Seuran kirjasto
heli.kautonen@finlit.fi

PIRJO KANGAS
Ellibs Oy / Lingsoft Oy
pirjo.kangas@gmail.com

Mikko Lappalainen, Juha Hulkkonen, Juho Inkinen,
Aleksi Kallio, Markus Koskela, Mona Lehtinen,
Mats Sjöberg, Osma Suominen ja Laxmana Yetukuri

Automaattisen sisällönkuvailun ohjelmiston rakentaminen – case Annif

Sisällönkuvailun automatisointiratkaisut ovat puhuttaneet kirjastomaailmassa viime vuosina, ja erilaisia kokeiluja on tehty niin Suomessa kuin maailmallakin. Kansalliskirjastossa kehitetty automaattisen sisällönkuvailun Annif-työkalu on herättänyt paljon mielenkiintoa monissa organisaatioissa ja kokemukset ensimmäisistä käyttöönotoista ovat olleet lupaavia. Mitä kehitysvalintoja Annifia rakennettaessa on tehty, ja minkälaisia haasteita kuvailun automatisointiin ylipäätään liittyy?

Tässä artikkelissa esitellään tekstiaineistojen automaattisen sisällönkuvailun ohjelmiston rakentamisen haasteita ja ratkaisuja erityisesti KAM-sektorin (kirjastot, arkistot ja museot) organisaatioiden näkökulmasta. Esimerkkinä käytetään Kansalliskirjastossa kehitettävää automaattisen sisällönkuvailun työkalua Annifia. Artikkelin pohjautuu selvitykseen, *Sisällönkuvailun automatisoinnin haasteita ja ratkaisuja kulttuuriperintöorganisaatioissa* (Hulkkonen ym. 2021), joka julkaistiin alkuvuodesta 2021 osana CSC:n, Kansalliskirjaston ja Kansallisarkiston yhteistä High-Performance Digitisation -hanketta.

1. Sisällönkuvailun automatisoinnin lähtökohdat

1.1 Manuaalinen, puoliautomaattinen ja automaattinen kuvailu
Sisällönkuvailun tarkoituksena on ilmaista kuvailtavan aineiston (esim.

kirja, artikkeli, verkkosivu jne.) aiheaineiston löytämisen helpottamiseksi. Kirjastoissa ja muissa kulttuuriperintöalan organisaatioissa sisällönkuvailua on tehty pitkään ihmistyönä valmiita asiansastoa ja/tai muuta kontrolloitua sanastoa hyödyntäen, ja sisällönkuvailijoista on muodostunut oma ammattiryhmänsä ainakin isoimmissa kirjastoissa. Verkkomaailmassa sisällönkuvailua tehdään usein sivutyönä toisen pääasiallisen työn ohessa, kuten esimerkiksi toimittaja sisällönkuvaillee kirjoittamansa artikkelin osana toimitustyötä. Tämä käyttäjäryhmien moninaisuus onkin hyvä huomioida automaattisen kuvailun ohjelmistoja ja palveluita suunniteltaessa, erityisesti silloin kun ajatuksena on tuottaa monissa erilaisissa ja erikokoisissa organisaatioissa hyödynnettäväksi soveltuva yleiskäyttöinen ohjelmisto.

Puoliautomaattisessa sisällönkuvailussa jokin järjestelmä antaa halutulle tekstille asiansaehdotuksia, jotka käydään läpi ihmistyönä. Tarkoituksena on

sujuvoittaa kuvailutyötä näillä pohjaehdotuksilla; ihmiselle jää kuitenkin päävastuu kuvailusta.

Täysin automaattisessa kuvailussa ohjelmiston/algoritmin tuottama kuvailu etenee sellaisenaan ilman ihmistarkistusta järjestelmän tietokantaan. Tässä menetelmässä sudenkuoppana on se, ettei asiasanoituksen laatu välttämättä vastaa ihmistyönä tehtyä. Positiivisena puolena kuitenkin voidaan nähdä se, että näin saatetaan saada edes jollain tasolla kuvailtua aineistoa, joka muuten jäisi kokonaan kuvailematta. Kuvailutyötä automatisoivilla ratkaisuilla voidaan myös rikastaa olemassa olevaa metadataa.

Erilaisten dokumenttien sisältöä kuvaillaan monista eri lähtökohdista. Varsinainen asiasanoitus, koneellisesti tai ihmisvoimin, tehdään käyttötarkoitukseen sopivan kuvailusanaston perusteella, esimerkkinä vaikkapa Yleinen suomalainen ontologia yso tai sen edeltäjä Yleinen suomalainen asiasanasto ysa. Tällä pyritään varmistamaan kuvailun yhteismitallisuutta ja aineiston löydettävyyttä.

Aineistoa voidaan myös kuvailla vapaammin – ja astetta subjektiivisemmin – käyttämättä kontrolloitua sanastoa tai laventamalla sanaston käyttöön liittyviä ohjeistuksia. Tällöin kuitenkin puhutaan ennemmin vaikkapa avainsanojen määrittämisestä tai aineiston ”taggaamisesta”.

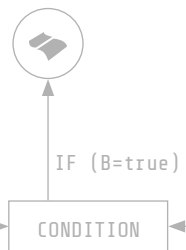
1.2 Sisällönkuvailun laadun mittaus

Sisällönkuvailun laadun mittaaminen ei ole suoraviivainen tehtävä. On vai-

keaa määrittää ehdottomat kriteerit laadukkaalle kuvailulle tai löytää jokin vertailukohta, joihin muiden samasta tekstistä tehtyjen kuvailujen oikeellisuutta voitaisiin verrata. Kuvailijoiden välillä, ja usein myös saman kuvailijan eri kerroilla tuottamien kuvailujen välillä, on harvoin kovin suurta yhdenmukaisuutta. Empiirisissä kokeissa eri ihmisten tuottamien kuvailujen yhdenmukaisuus on ollut 30–50 % (Medelyan 2009).

Laadun mittaamisessa ja arvioinnissa voidaan hyödyntää esimerkiksi kuvailijajoukon keskinäistä samanmielisyyttä. Mitä enemmän yksittäinen kuvailu sisältää samoja asiasanoja kuin muut samasta tekstistä tehdyt kuvailut, sitä laadukkaampana tätä kuvailua voidaan pitää (ks. Rolling 1981). Tässä kuvailijaksi voidaan laskea algoritmitkin. Voidaan esimerkiksi laskea, kuinka monta samaa asiasanaa algoritmit ovat ehdottaneet dokumentille ja mitä ihmisten valitsemia asiasanoja jäi puuttumaan. Näin voidaan laskea tiedonhaun arvioinnista tuttuihin saantiin ja tarkkuuteen perustuvia tunnuslukuja, esim. F1-tunnusluku on saannin ja tarkkuuden harmoninen keskiarvo. Joskus arvioinnissa on myös otettava huomioon ehdotettujen asiasanojen järjestys, etenkin tilanteessa jossa oletetaan, että ensimmäiset asiasanat ovat tärkeimpiä. Mittarina esim. NDCG (Normalized Discounted Cumulative Gain) huomioi tämän.

Toinen tapa lähestyä sisällönkuvailun laadun arviointia on kysyä ihmisten mielipidettä ehdotetusta kuvailusta (ks. esim. Golub ym. 2016). Asiantuntijat





(tai maallikot) voivat arvioida tekstile annettujen asiasanojen laatua ja hyödyllisyyttä jollain asteikolla (esim. 1–5 tai hyödyllinen/kelvollinen/väärä). Vastaaajilta voidaan kysyä lisäinformaatiota ja perustelua avovastauksissa tulosten tulkinnan avuksi.

1.3 Automaattisen sisällönkuvailun työkalu Annif

Sisällönkuvailun avuksi on kehitetty työkaluja, joista yksi on tässä artikkelissa käsiteltävä Annif (Suominen 2019). Annifin toiminta perustuu erilaisiin kieliteknologiaa ja koneoppimista hyödyntäviin työkaluihin ja algoritmeihin, joita voi käyttää erikseen tai yhdessä toistensa kanssa. Annif on avointa lähdekoodia, ja kenen tahansa hyödynnettävissä. Annif on myös muokattavissa tietyn käyttötapauksen tarpeita varten. Annifilla on rajapintapalvelu, joka mahdollistaa sen integroinnin käyttäjien omiin järjestelmiin.

Artikkelin kirjoittamishetkellä Annifia on hyödynnetty eniten suomen-, ruotsin- ja englanninkieliseen asiasanoitukseen Yleisen suomalaisen

ontologian yso:n perusteella. Annif on kuitenkin periaatteessa kieliriippumaton, sikäli kun taustatyökaluista löytyy halutun kielen tuki. Annif on ainakin toistaiseksi ollut käytössä pääasiassa puoliautomasoiduissa sisällönkuvailuprosesseissa. Annifista ja sen taustamenetelmistä voi lukea lisää englanniksi GitHubin wikisivulta (github.com/NatLibFi/Annif/wiki) tai osoitteesta annif.org. Annifilla on myös suomenkielinen asiakaswiki (www.kiwi.fi/display/Finto/Annif).

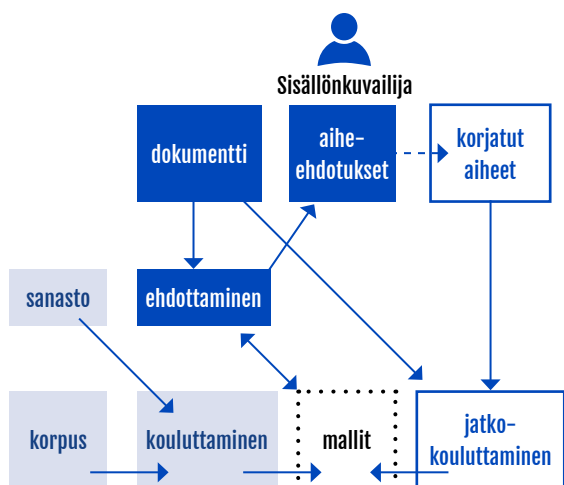
2. Automaattisen kuvailun ohjelmiston keskeiset toiminnallisuudet

Automaattisen kuvailun työkalut tarjoavat useita eri toimintoja. Tässä luvussa on keskitytty erityisesti Annif-työkalun tarjoamiin toiminnallisuuksiin. Muissa vastaavissa työkaluissa toiminnallisuudet voivat olla hiukan erilaisia. Automaattisen sisällönkuvailun ohjelmiston keskeiset toiminnallisuudet ja prosessit on esitetty kuvassa 1.

2.1 Aihe-ehdotusten tuottaminen

Sisällönkuvailijat ovat puoliautomaattisen sisällönkuvailun järjestelmän varsinaisia käyttäjiä. Sisällönkuvailijan kannalta järjestelmän keskeisin toiminnallisuus on järjestelmän sille syötetyn dokumentin pohjalta tuottamat asiasana- tai luokkaehdotukset.

Ehdottaminen voi tapahtua rajapintapalvelun kautta. Tällöin sisällönkuvailija käyttää omaa kuvailujärjestelmäänsä, joka on integroi-



Kuva 1. Automaattisen sisällönkuvailun ohjelmiston keskeiset toiminnallisuudet ja prosessit

tu automaattisen sisällönkuvailun järjestelmään. Toisessa vaihtoehdossa, joka ei vaadi järjestelmäintegraatiota, sisällönkuvailija syöttää tekstin suoraan automaattisen sisällönkuvailun järjestelmään esimerkiksi verkkolomakkeen kautta ja saa ehdotukset nähtäville kopioitavaksi varsinaisen kuvailujärjestelmään. Tässäkin tapauksessa verkkolomake tukeutuu automaattisen sisällönkuvailun järjestelmän rajapintapalveluun. Annifiin perustuva verkkolomake löytyy esimerkiksi Finto AI -palvelusta (kuva 2.).

Järjestelmän tuottamissa ehdotuksissa on toisinaan aiheita, jotka ovat dokumentin kannalta toissijaisia tai virheellisiä. Toisaalta jokin aivan keskeinen aihe voi puuttua ehdotuksesta. Tällöin sisällönkuvailija voi korjata kuvailuehdotusta poistamalla epärelevantteja aiheita ja lisäämällä uusia ennen kuin kuvailu tallennetaan kuvailujärjestelmään. Samalla kuvailujärjestelmä voi lähettää sekä alkuperäisen dokumentin että korjatun kuvailun automaattisen sisällönkuvailun järjestelmään, joka niiden perusteella päivittää mallia, korjaa toimintaansa ja kehittää ehdotuksia paremmiksi.

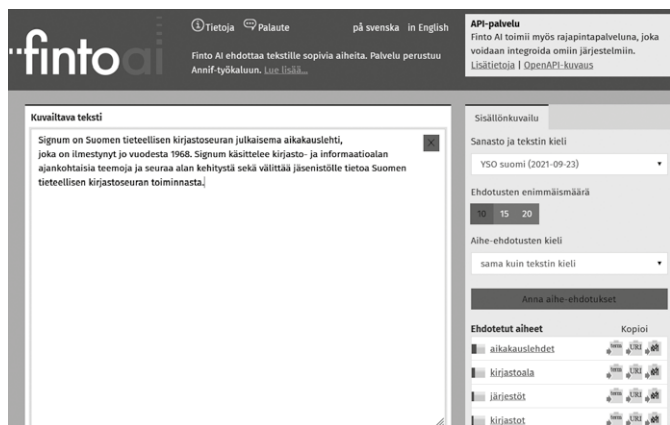
2.2 Koneoppimismallien kouluttaminen

Tekstin sisällönkuvailuun tarkoitettujen koneoppimismallit ovat käytännössä aina sidoksissa tiettyyn sanastoon sekä luonnolliseen kieleen. Käytettävän sanaston ja kielen valinta tapahtuu luomalla ”projekti”, jossa määritellään joukko asetuksia, kuten käytettävä sanasto, kieli, algo-

ritmi ja mahdolliset hyperparametrit eli koneoppimismallien rakenteeseen liittyvät asetukset.

Kun projekti on määritetty, se täytyy yleensä vielä kouluttaa esimerkkidokumenteilla. Kouluttamisessa luodaan projektin määritysten mukainen koneoppimismalli ja opetetaan se esimerkkidokumenttien avulla enustamaan aiheita dokumenteille. Kouluttamisen jälkeen projekti on käyttövalmis.

Kouluttaminen tapahtuu syöttämällä koneoppimismallille korpus, joka koostuu joukosta etukäteen kuvailtuja dokumentteja. Kouluttamisen lopputulokseen vaikuttavat myös käytettävät esikäsittelytavat (esim. sanojen perusmuotoistaminen eli lemmatisointi) sekä koneoppimismallin hyperparametrit. Kouluttaminen saattaa olla raskasta ja vaatia paljon laskentaresursseja ja/tai RAM-muistia. Neuroverkkomallien kouluttamista nopeutetaan usein grafiikkasuorittimilla (GPU, graphics processing unit) suoritettavalla laskennalla, mutta siitä huolimatta mallin koulutus saattaa viedä tunteja tai jopa useita päiviä.



Kuva 2. Finto AI -palvelun Annifiin perustuva verkkolomake, johon käyttäjä voi syöttää haluamansa tekstin automaattisesti sisällönkuvailtavaksi.

Projektien sisältämät koneoppimismallit voidaan kouluttaa uudelleen, jos siihen on riittävästi syitä. Muuttunut sanasto, parempi koulutusaineisto tai halu kokeilla uusia hyperparametreja voivat toimia perusteena. Sanastoa päivitettäessä haasteena on sanastoon lisättyjen uusien käsitteiden tunnistaminen. Useimmat koneoppimismallit pystyvät tunnistamaan vain niitä aiheita, jotka esiintyvät koulutusdatassa, joten ne eivät pysty ehdottamaan uudessa sanastoversiossa lisättyjä käsitteitä.

2.3 Automaattinen laadun arviointi

Kun projekti on koulutettu, sen tuotamien ehdotusten laatua voidaan arvioida tilastollisesti jo kuvailtuja dokumentteja vasten. Arvioinnin ajatuksena on tuottaa joukolle dokumentteja (testikorpus) aihe-ehdotukset ja verrata algoritmin antamia ehdotuksia aiemmin ihmistyönä tehtyihin kuvailuihin. On tärkeä varmistaa, että testikorpukseen sisältyviä dokumentteja ei ole käytetty projektin kouluttamisessa eli ne ovat algoritmin kannalta uusia, koska muuten algoritmi voi muistaa yksityiskohtia koulutusdokumenteista ja antaa siksi epärealistisen hyviä ehdotuksia.

Ehdotusten laatua arvioidaan tyyppillisesti samankaltaisuusmittareilla kuten saanti, tarkkuus, F1-score ja NDCG. Jos arvioitu laatu ei tyydytä, voi ylläpitäjä kokeilla vaihtaa projektin asetuksia (esimerkiksi hyperparametrit) ja uudelleenkoulutuksen jälkeen arvioida laatua uudelleen, kunnes saavutettu laatu on riittävän hyvä.

3. Käytettävien algoritmien valinta

3.1 Tekstin luokittelu sisällönkuvailussa

Koneoppimisen näkökulmasta automaattiseen sisällönkuvailuun on useita erilaisia lähestymistapoja ja siihen voidaan soveltaa monia eri algoritmeja. Yleisesti ottaen sisällönkuvailua voidaan pitää moniluokkaisena tekstin luokittelutehtävänä, jossa jokaiselle syötetekstille poimitaan sitä parhaiten kuvaavat asiasanat käytössä olevasta sanastosta.

Sovellettaessa tekstin luokittelua sisällönkuvailuun, jokaista sanaston asiasanaa vastaa erillinen luokka eli asiasana-avaruudessa on yhtä monta ulottuvuutta kuin sanastossa on asiasanoja. Automaattinen sisällönkuvailu poikkeaa tavanomaisista koneoppimisen luokittelutehtävistä luokkien lukumäärän osalta. Koneoppimisessa on tyyppillistä, että luokkia on suhteellisen vähän, yleensä korkeintaan joitakin kymmeniä tai satoja. Sisällönkuvailussa käytetyissä sanastoissa asiasanojen lukumäärä voi sen sijaan olla kymmeniätuhansia tai jopa enemmän. Esimerkiksi Wikipediassa käytetään yli miljoonaa kategoriaa sivujen luokitteluun. Tekstin luokittelua suureen määrään eri luokkia on viime vuosina tutkittu suhteellisen paljon. Siihen on kehitetty myös useita uusia menetelmiä, sillä sille on löydetty useita taloudellisesti merkittäviä sovelluksia.

Useat käytössä olevat algoritmit palauttavat myös asiasanoja vastaavat hyvyysarvot eli arviot asiasanojen rele-

vanssista kyseessä olevalle syötetekstil-
le. Näitä hyvyysarvoja voidaan käyttää
asiasanojen rajoittamiseen esimerkiksi
siten, että ainoastaan tiettyä kynnysar-
voa korkeampia hyvyysarvoja vastaa-
vat asiasanat otetaan huomioon, tai
asiasanojen järjestämiseen relevanssin
mukaiseen järjestykseen.

3.2 Algoritmityyppit

Automaattisessa sisällönkuvailussa käy-
tettyjä algoritmityyppiejä ovat esimer-
kiksi:

- leksikaaliset algoritmit
- tiedonhakualgoritmit
- luokkakohdaiset luokittelualgoritmit
- puualgoritmit
- tiivistysalgoritmit
- hierarkkiset luokittelualgoritmit
- neuroverkot.

Eri algoritmityyppit lähestyvät teh-
tävää eri näkökulmasta. Esimerkiki-
si tiedonhakualgoritmien tapauksessa
sisällönkuvailutehtävä voidaan tulkita
tiedonhakuongelmana, jossa kuvail-
tava syöteteksti vastaa hakukyselyä ja
jokaista sanaston asiasanaa vastaa si-
tä varten laadittu asiasanaa kuvaileva
dokumentti tai joukko dokumentteja.
Puualgoritmit puolestaan perustuvat
hajota ja hallitse -periaatteeseen, jossa
alkuperäinen ongelma jaetaan yksin-
kertaisemmiksi osaongelmiksi ja joiden
tuloksista alkuperäisen ongelman rat-
kaisu lopuksi koostetaan.

Annif hyödyntää tekstin kirjoit-
tushetkellä pääasiassa leksikaalisia ja
puualgoritmeja, mutta kokeiluja eri
algoritmityyppien toimivuudesta teh-
dään jatkuvasti.

Algoritmityyppien rajat eivät ole
tarkasti määriteltyjä, ja sisällönkuvai-
lussa käytettävät menetelmät saatta-
vatkin sisältää osia tai ominaisuuksia
useista algoritmityypeistä. Algoritme-
ja voidaan myös yhdistää, ja parhaat
tulokset saavutetaankin usein erilais-
ten algoritmien tuloksia yhdistelevillä
(ensemble) malleilla. Annif tukee yk-
sittäisten algoritmien käytön lisäksi
yhdisteleviä malleja.

4. Yhteenveto

Sisällönkuvailua tehdään aineiston
löydettävyyden parantamiseksi, eli
tiedonhakihoita varten. Kuvailun au-
tomatisoinnilla pyritään helpottamaan
ja sujuvoittamaan kuvailijan työtä ja
saattamaan kuvailun piiriin aineistoja,
jotka muuten saattaisivat jäädä kuvai-
lematta. KAM-sektorilla kuvailutyötä
tekevät esimerkiksi organisaatioiden
tiedonhallinnan ammattilaiset, mutta
kuvailua tehdään paljon myös ns. oman
toimen ohella. Automaattisen kuvailun
järjestelmiä rakennettaessa yhtenä haas-
teena onkin järjestelmää hyödyntävien
kuvailijoiden heterogeeninen tausta, ja
sitä kautta erilaiset tarpeet.

Kuvailua automatisoitaessa on tär-
keää seurata ja mitata kuvailun laatua.
Sisällönkuvailun laadun mittausta ei ole
kuitenkaan suoraviivainen tehtävä, kos-
ka absoluuttisen oikeaa kuvailua on
usein mahdotonta määritellä. Useissa
tapauksissa voikin olla tarpeen hyö-
dyntää erilaisia laatumittareita saman-
aikaisesti.

Annifin kaltaisen automaattisen
sisällönkuvailun työkalun käyttöön-

otossa lähtökohtana on sopivan kuvailusanaston valinta. Käytettävien koneoppimismallien opettamiseen tarvitaan valitulla sanastolla ihmistyönä kuvailtuja esimerkkidokumentteja. Sisällönkuvailupalvelun suunnittelussa tulee lisäksi ottaa huomioon käytettävien mallien käyttömahdollisuudet erilaisille aineistoille, koulutusaineistojen koko ja ylläpidon tarve.

Automaattiseen sisällönkuvailun ohjelmisto voi hyödyntää monia erilaisia algoritmeja, joiden lähestymistavat poikkeavat toisistaan. Parhaiden tulosten saavuttamiseksi onkin usein tarpeen yhdistellä eri tyyppisiä algoritmeja. Lopulta Annifin kaltainen ohjelmisto tulee tarjota käyttäjien hyödynnettäväksi mahdollisimman helpolla tavalla, esimerkiksi rajapintaintegraationa olemassa olevaan kuvailujärjestelmään. ❖

Lähteet

GOLUB, K., SOERTEL D., BUCHANAN G., TUDHOPE D., LYKKE M. & HIOM D., 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.
<https://doi.org/10.1002/asi.23600>

HULKKONEN, J., INKINEN J., KALLIO A., KOSKELA M., LAPPALAINEN M., LEHTINEN M., SJÖBERG M., SUOMINEN O. & YETUKURI L., 2021. *Sisällönkuvailun automatisoinnin haasteita ja ratkaisuja kulttuuriperintöorganisaatioissa*. <https://urn.fi/URN:ISBN:978-951-51-7233-4>

MEDELYAN, O., 2009. *Human-competitive automatic topic indexing* (Doctoral dissertation, The University of Waikato).
<https://researchcommons.waikato.ac.nz/handle/10289/3513>

ROLLING, L., 1981. Indexing consistency, quality and efficiency. *Information Processing & Management*, 17(2), 69-76. [https://doi.org/10.1016/0306-4573\(81\)90028-5](https://doi.org/10.1016/0306-4573(81)90028-5)

SUOMINEN, O., 2019. Annif: DIY automated subject indexing using multiple algorithms. *LIBER Quarterly*, 29(1), pp.1-25. DOI: <http://doi.org/10.18352/lq.10285>

MIKKO LAPPALAINEN
 Kansalliskirjasto
mikko.lappalainen@helsinki.fi

JUHA HULKKONEN
 CSC – Tieteen tietotekniikan keskus
juha.hulkkonen@csc.fi

JUHO INKINEN
 Kansalliskirjasto
juho.inkinen@helsinki.fi

ALEKSI KALLIO
 CSC – Tieteen tietotekniikan keskus
aleksi.kallio@csc.fi

MONA LEHTINEN
 Kansalliskirjasto
mona.lehtinen@helsinki.fi

MARKUS KOSKELA
 CSC – Tieteen tietotekniikan keskus
markus.koskela@csc.fi

MATS SJÖBERG
 CSC – Tieteen tietotekniikan keskus
mats.sjoberg@csc.fi

OSMA SUOMINEN
 Kansalliskirjasto
osma.suominen@helsinki.fi

LAXMANA YETUKURI
 CSC – Tieteen tietotekniikan keskus
laxmana.yetukuri@csc.fi

Tekoälyohjelma Iris.AI:n alkutaival, käyttäjäkokemukset ja uudet kilpailijat

Tekoälyohjelma Iris.ai otettiin Helsingin yliopiston kirjastossa käyttöön syksyllä 2017 ja Itä-Suomen yliopiston kirjastossa huhtikuussa 2018. Palvelua markkinoitiin tutkijan assistenttina, jolla tieteellisen tiedon määrä saadaan nopeammin haltuun. Millaisia olivat käyttäjien kokemukset Iris.aista?

”Iris AI on keinoälypohjainen työkalu open access -tutkimuksen kartoittamiseen. Tee haku millä tahansa tutkimuspaperin verkko-osoitteella, Iris lukee sen abstraktin ja rakentaa lukemansa pohjalta tutkimuskartan etsimällä siihen keskeisimmät sisällöt 30 miljoonan avoimen tutkimuspaperin joukosta.”

Näin määritteli nuori startup-yritys Iris.ai palvelutuotteensa syksyllä 2016, kun sitä tarjottiin ensimmäisen kerran Helsingin yliopiston kirjastolle. Yritys oli perustettu edellisenä vuonna 2015. Ensimmäinen yhteistyön vuosi yrityksen ja kirjaston kesken käytettiin palvelun kehittämiseen ja erityisesti keskityttiin siihen, millaista tiedonhakupalvelua tieteelliset kirjastot ovat tottuneet vaatimaan.

Edelläkävijäasiakkaana toimiminen on tuonut Helsingin yliopiston kirjastolle myös huomattavaa näkyvyyttä tiedonhaun kentällä toimivien organisaatioiden parissa. Paitsi suomalaisista organisaatioista niin myös useista eurooppalaisista kirjastoista on otettu yhteyttä ja toivottu kertomaan palvelun käytöstä ja sen vaikutuksista tiedonhakuun. Näin ollen kirjaston on ollut mahdollista profiloitua uusiin

palveluihin positiivisesti ja rohkeasti suhtautuvana proaktiivisena toimijana, aivan kuten suuren tieteellisen kirjaston tuleekin toimia.

Premium-versio ja sen Focus-työkalu valmistuivat syksyllä 2017 ja tässä vaiheessa Iris.ai (the.iris.ai/) tuli laajemminkin tieteellisten kirjastojen tietoisuuteen. Helsingin yliopiston kirjasto päätyi hankkimaan käyttölisenssin huhtikuussa 2018, Itä-Suomen yliopiston (UEF) kirjasto lokakuussa 2018. Lisäksi Iris.ai on ollut pilottikäytössä useissa yliopisto- ja ammattikorkeakoulukirjastoissa, ja palvelua on kehitetty paljon kirjastolaisten ja kirjaston asiakkaiden kommenttien pohjalta.

Miksi tekoäly tuli tieteelliseen tiedonhakuun?

On arvioitu, että joka päivä julkaistaan noin 4 000 uutta tieteellistä artikkelia tai muuta julkaisua. Vuositasolla puhutaan siis puolestatoista miljoonasta tuotoksesta. Suuret monialaiset tietoaaineistot ja niiden luonnollisen kielen parempi ymmärtäminen haastavat perinteiset hakusanoilla ja tietokantakohtaisilla asiasanastoilla toimivat tietokannat. Iris.ai:n taustalla

UPDATE

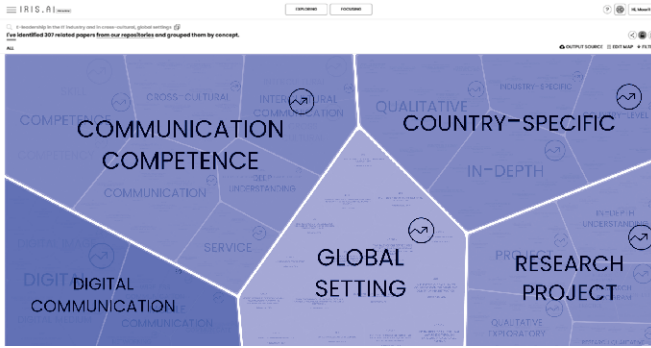
true

MORE

false

2: get issue detail

VALUE=1
+trackNo



Kuva 1. Iris.ai käsittekartta aiheeseen E-leadership in the IT industry and in cross-cultural global setting

olevassa CORE-tietokannassa on tällä hetkellä yli 200 miljoona dokumenttia, kun Iris.ai aloittaessa niitä oli noin 30 miljoonaa. PubMed-tietokantaan lisättiin 192 057 uutta tietuetta vuonna 1967, vuonna 2020 määrä oli jo yli 1,4 miljoonaa.

Ovatko suuret julkaisumäärät tutkijan hallittavissa? Tekoälyn on tarkoitus säästää tutkijan aikaa.

Julkaisumäärän lisäksi kasvavat vaatimukset tieteidenvälisestä tutkimuksesta ja sen avulla saatavista innovaatioista ovat lisänneet paineita etsiä uusia ideoita tieteellisen tiedon kokonaisuudesta, eikä ainoastaan yhden tieteenalan tietokannoista.

Samaan aikaan keskusteluun on noussut myös avoin tiede ja tutkimusjulkaisujen entistä avoimempi

saatavuus ja tiedekustantamisen ja -julkaisemisen murros. UEF-kirjastossa pohdittiin myös, olisiko Iris.ai mahdollisesti yksi toimiva väline avointen julkaisujen hakuun.

Iris.ai hakuvälineenä

Iris.ai ei markkinoinut itseään tietokantana kirjastoille, vaan tutkijan assistenttina, jolla tieteellisen tiedon määrä saadaan nopeammin haltuun. Se ikään kuin ”lukee” dokumentit ja seuroi käyttäjälleen niistä relevanteimmat.

Iris.ai mahdollistaa tiedonhaun kiinnostavan tutkimusartikkelin URL-tai DOI-linkkiin pohjautuen ja näyttää hakutulokset visuaalisena käsittekarttana (Kuva 1). Maksullinen organisaatiolisenssi mahdollistaa karttojen muokkaamiseen ja haun tarkentamisen Exploring- ja Focusing-työkaluilla. Innovatiivista oli yksittäisten hakusanojen ja Boolean logiikan sijaan hakumahdollisuus omin sanoin 300–500 sanan mittaisella tiivistelmällä ja tulosten visuaalinen esittämistapa.

Moni jättää haun käsittekarttavaiheeseen, vaikka Focusing-työkalulla on tarkoitus tarkentaa tulosta, ottaa mukaan ja sulkea pois käsitteitä ja käsiteryypäitä.

Search topic	Document count
g31: health, patient, care, patients, medical, RESULTS, METHODS, surgery, hospital, clinical	(142 docs)
g25: management, practice, project, knowledge, service, Implementation, professional, survey, Issue, progr...	(115 docs)
g72: policy, economic, market, country, sector, countries, price, Income, economy, International	(33 docs)
g27: political, law, article, legal, century, society, history, national, conflict, debate	(17 docs)
g34: physical, score, Intervention, mental, depression, questionnaire, disorder, satisfaction, positive, behavi...	(16 docs)
g79: would, even, way, still, like, much, possible, become, question, therefore	(13 docs)

Kuva 2. Iris.ai:n focusing-työkalun käsityrväs.

Käyttäjien kokemuksia – ruusuja

Itä-Suomen yliopiston kirjastossa on kerätty palautetta Iris.ain käytöstä opiskelijoilta. Seuraavassa esitetyt näkemykset ja sitaattit ovat Maarit Putouksen vuosina 2020–2021 vetämiltä tiedonhaun Moodle-kursseilta, joissa opiskelijoilla on ollut ylimääräisenä ja vapaaehtoisena tehtävänä kokeilla Iris. ai:ta. Kokeilijoita on ollut parikymmentä. Lisäksi palautetta on kerätty muutamilta tutkijoilta, joiden kanssa Iris.aita on käytetty heidän tutkimusongelmiinsa liittyvissä tiedonhauissa.

Näin positiivisesti Iris.aita kuvailee eräs käyttäjä:

”Auttaa löytämään sellaisia uusia yhtäläisyyksiä mitä ei välttämättä itse tajuaisi etsiä. Toki samaa arvoa saa Googlen algoritmin avulla mutta olen kokenut molempien ikään kuin täydentävän toisiaan ja molempia käyttämällä löytää mainiosti uutta relevanttia luettavaa. On se helppoa näin nykypäivänä!”

Muita Iris.ain myönteisiksi arvioituja ominaisuuksia ovat:

- nopeus
- luonnollisella kielellä haku
- sanojen yhteyksiä ja kategorisointeja/ryhmittelyjä
- vastaan voi tulla sellaisiakin käsitteitä ja yhteyksiä, joita ei ole edes osannut ajatella
- auttaa tutkimuksen alkuvaiheessa hahmottamaan aiheen taustaa

- hieman laajemman perspektiivin/ kontekstin saaminen
- esim. opiskelijalle, joka ei tunne vielä käsitteitä eikä osaa hakusanoilla hakuajan rajat

Visuaalisuudesta – käsitekartasta – tykkäävät kaikki!

”Käyttöliittymä on käytettävyydeltään virkistävän moderni, ja visuaalisesti pelkistetty ja kaunis verrattuna moneen muuhun käyttämäni hakutietokantaan. Hakutulokset oli melko pirstaleinen ja artikkelien tuoreus vaihteli melkoisesti. Vanhin oli vuodelta 1994 ja uusin vuodelta 2016. Oli mielenkiintoista kokeilla tämän tyyppistä käyttöliittymää.”

Käyttäjien kokemuksia – risuja

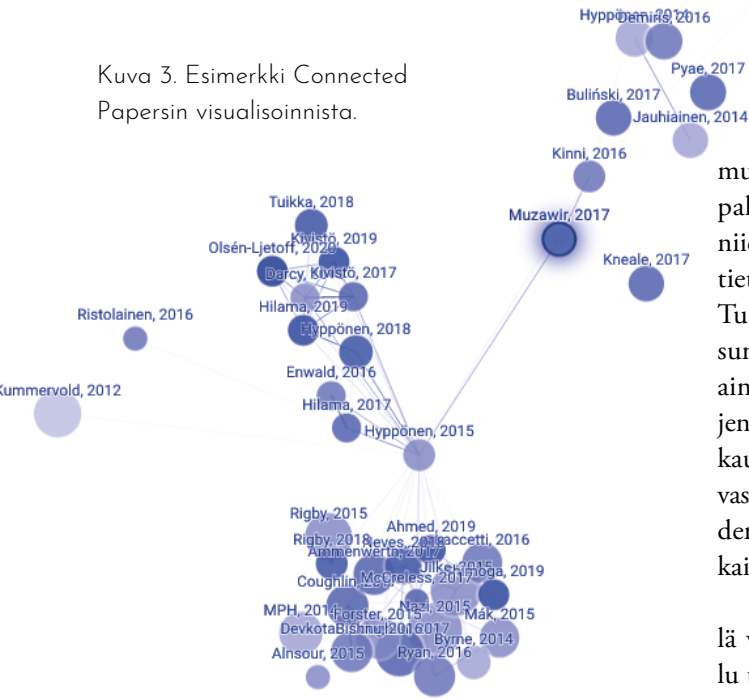
Iris.ain käyttöön liittyy myös kielteisiä kokemuksia.

”Eiväthän nämä ole oikeaa asiaa! Miten tähän voi luottaa? Tähän ei kannata tuhlaa aikaa! Google Scholar pystyy paljon parempaan.”

Ensimmäinen ongelma on, että käsitekartan sisältö on ”mitä sattuu”. Tulokset eivät näytä relevanteilta, tärkeimmät käsitteet puuttuvat kartalta eivätkä asiakokonaisuudet yhdisty keskenään. Toisaalta tutkimusongelmissa, joissa on tarkat käsitteet, myös tekoäly toimii paremmin.



Kuva 3. Esimerkki Connected Papersin visualisoinnista.



Tulosmäärät nähtiin joskus yllättävän pieninä verrattuna perinteisiin hakumenetelmiin. Mitä, miksi ja miten jättää tuloksista pois?

”Minulle jäi vielä epäselväksi, että miten saan tämän ohjelman avulla varmistettua, että tekstit ovat tieteellisiä? Onko sellaista rajausmahdollisuutta edes olemassa tässä ohjelmassa?”

Pohjadataan merkitys tieteellisessä tiedonhauissa on olennainen. CORE-tietokantaan haravoidaan eri organisaatioiden julkaisuarkistoja, ja ne pitävät sisällään myös gradu-, kandidaatintyö- ja amk-opinnäytetyötasoisia julkaisuja. Vertaisarviointirajausta tai opinnäytetöiden pois sulkua ei ole.

PubMed/Medline on integroituna Iriksen, mutta tämä ei riitä. Monet

muut open access -aineistoa hakevat palvelut ovat maksutta käytössä ja niiden vaikuttimena näyttäisi olevan tieteen avoimuus ja saavutettavuus. Tutkimustiedosta $\frac{2}{3}$ on vielä maksumuurien takana ja suuri osa tästä aineistosta on käytettävissä kirjastojen perinteisten väylien/tietokantojen kautta. Aineiston integrointi maksua vastaan Iris.ai:n haettavaksi ei ole muiden kuin Iris.ain näkökulmasta ratkaisu.

Kielen ymmärtämisessä on Iriksellä vielä oppimista. Fokusointityökalu tekee esim. tällaisia käsiteryypäitä: would, even, way, still, like, much, possible, become, question, therefore.

Harva palautetta kirjoittanut käyttäjä on edennyt Focusing-työkalulla. Karttavaiheessa käyttäjät innostuvat – tai turhautuvat. Focusing-työkalu on vaikeasti hahmotettava ja työläs. Jos viitetulos on pieni, hakutuloksen/viitelistan silmäilee nopeasti.

Uusia yrittäjiä tekoälypohjaisen tiedonhakuun

Tieteellisen tiedon valtavan kasvun myötä tiedonhaun markkinoille tulee pioneiryritysten jälkeen jatkuvasti toimijoita. Toisena syynä on tietenkin myös tekoälyn voimakas kehittyminen ja siihen kohdistuvat panostukset. Avoin julkaiseminen lisää massaa uusien tulokkaiden käyttämiin tietokantoihin (esim. Semantic Scholar), jolloin tekoälystä saadaan mahdollisimman paljon irti.

Tässä kolme esimerkkiä tekoälyä hyödyntävistä palveluista, joissa kaikissa tulokset esitetään visuaalisena karttana ja usein haun lähtökohtana on kiinnostava artikkeli (vrt. helmenkasvatus- tai lumipallostrategia):

- **Connected Papers:** Visuaalinen työkalu, jonka algoritmi pohjaa artikkelien samankaltaisuuteen, päällekkäisiin viittauksiin ja bibliografiseen kytkentään (Kuva 3.). Palvelu on maksuttomasti käytettävissä ja sitä hyödynnetään mm. Helsingin yliopiston Meilahden lääketieteellisen kampuksen uusien artikkelien twiittaamisessa. (www.connectedpapers.com/)

- **Open Knowledge Maps:** Algoritmi etsii tekstin samankaltaisuuksia ja muodostaa niistä ryhmiä, joissa julkaisut muodostavat yhtenäisen kokonaisuuden. (openknowledgemaps.org/index)

- **ResearchRabbit:** Tuore tulokas tekoälypohjaiseen tiedonhaussa. Tässäkin pohjalla ovat käyttäjälle tärkeät artikkelit, joiden perusteella palvelu suosittelee muita julkaisuja. Se nimittääkin itseään "Spotify for papers" -iskulauseella. Palvelu eroaa tietokantojen "Find related papers" -optiosta. Niissä on useimmiten käytetty yksinkertaisesti tietokannan metadataa ja indeksoituja termejä. ResearchRabbit hyödyntää tekoälyä ja algoritmi pohjautuu suosittelujärjestelmään. (www.researchrabbit.ai/)

B==C

true

ONLINE

Tietoympäristö muuttuu

Kirjastojen tietoasiantuntijoiden näkökulmasta tekoäly aiheuttaa myös ongelmia. Kuten tunnettua, kaikissa tekoälyä käyttävissä palveluissa algoritmien toiminta ei ole läpinäkyvää, johtuen erityisesti siitä, että ne ovat palvelujen ydintä ja liiketoiminnan perusta. Siten ei voida tietää miten hakutulos muodostuu ja esimerkiksi systemaattisissa kirjallisuuskatsauksissa niiden toistaminen on ongelmallista, sillä tekoäly ei kerro hakustrategiaansa.

Tieteellisissä kirjastoissa tarvitaan tietoisuutta ja ymmärrystä algoritmien toiminnasta tiedonhaussa. Tietoasiantuntijan on osattava valita ja ohjata tarkoituksenmukaisimpiin hakuvälineisiin. Osa asiantuntijuutta on verrata ja ottaa kantaa myös tekoälypohjaisiin järjestelmiin. ❖

KIRSI SALMI

Itä-Suomen yliopiston kirjasto
kirsi.salmi@uef.fi

JUKKA ENGLUND

Helsingin yliopiston kirjasto
jukka.englund@helsinki.fi

MAARIT PUTOUS

Itä-Suomen yliopiston kirjasto
maarit.putous@uef.fi

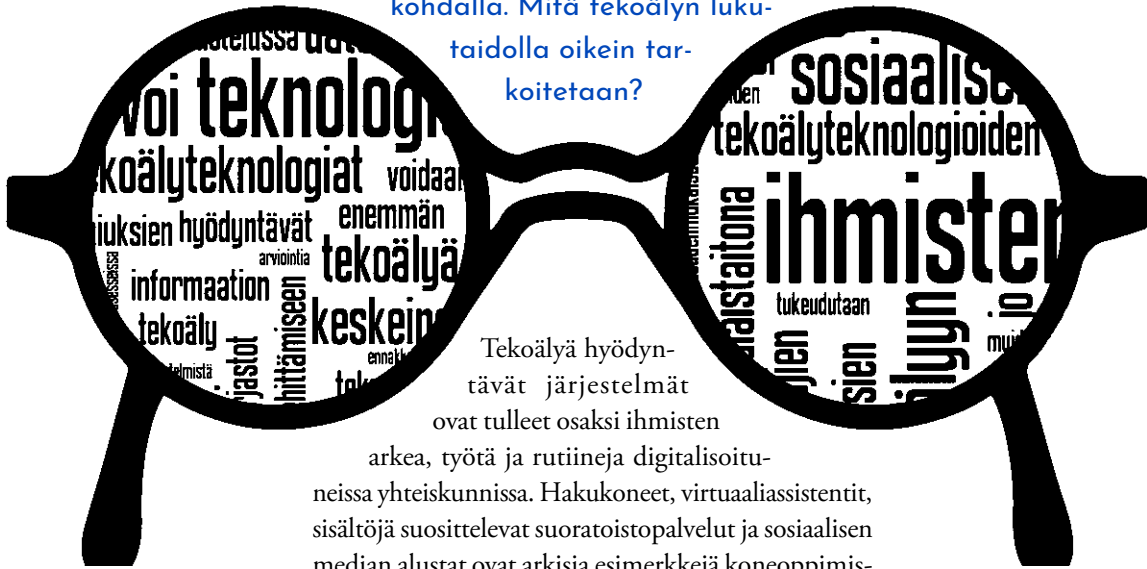
false

Heidi Enwald & Noora Hirvonen
TEKOÄLYN LUKUTAITO –

VALMIUKSIA, JOITA JO TARVITSEMME

Julkisessa keskustelussa keskitytään usein tekoälyn teknologiseen kehitykseen.. Tämän lisäksi huomiota tulisi kiinnittää uudenlaiseen osaamiseen, jota ihmiset tarvitsevat kyetäkseen hyödyntämään ja kriittisesti arvioimaan teknologioita ja niiden vaikutuksia. Tekoälyn lukutaidon (AI literacy) käsitettä käytetään viittaamaan tähän osaamiseen tekoälyteknologioiden kohdalla. Mitä tekoälyn luku-

taidolla oikein tarkoitetaan?



Tekoälyä hyödyntävät järjestelmät ovat tulleet osaksi ihmisten

arkea, työtä ja rutiineja digitalisoituneissa yhteiskunnissa. Hakukoneet, virtuaaliassistentit, sisältöjä suosittelevat suoratoistopalvelut ja sosiaalisen median alustat ovat arkisia esimerkkejä koneoppimista hyödyntävistä, dataan ”älykkäästi” reagoivista tekoälyteknologioista (Elliot, 2019). Juuri kyky oppia ja mukautua uuteen informaatioon tai ärsykkeeseen erottaa nämä järjestelmät useimmista digitaalisista tai algoritmipohjaisista järjestelmistä (Elliot, 2019).

Julkisessa keskustelussa keskitytään tavallisesti tekoälyn teknologiseen kehitykseen ja tulevaisuuden läpimurtoihin (Slotte, Dufva & Mertala, 2021). Elliotin (2019) mukaan niin kutsuttu tekoälyn vallankumous liittyy kuitenkin ensisijaisesti niihin valtaviin muutoksiin, joita nämä teknologiat tuovat, ja ovat jo tuoneet, ihmisten arkielämään. Kyse ei siis ole vain uusien teknologioiden kehittämisestä vaan niiden käyt-



IF (META = 1)
 ELSE (set DATA)



töönoton kiihtyvistä vauhdista useilla eri elämäalueilla (Sitra, 2020). Tekoälyteknologiat ovat yhä tiiviimmin osa sitä, mitä teemme ja muokkaavat sitä, keitä olemme (Elliot, 2019).

Tekoäly on käsitteenä vaikeasti määriteltävä, mikä haastaa sen toimintaperiaatteiden ja vaikutusten hahmottamista. Keskeinen ongelma on myös tekoälyteknologioihin liittyvä black box -ilmiö eli se, että niiden toiminta ei ole läpinäkyvää (Adadi & Berrada 2018). Tekoälyteknologioiden integroitua arkielämäämme, tarvitsemme uudenlaisia valmiuksia niin tekoälyn ymmärtämiseen kuin sen arviointiin.

Tekoälyn lukutaito

Toistaiseksi tekoälyyn liittyvien ilmiöiden ja niiden vaikutusten tarkastelu on keskittynyt teknologioiden kehittämiseen: miten voidaan varmistaa, että tekoälysovelukset toimivat eettisesti ja reilusti? Teknologioiden kehittämiseen liittyvän keskustelun ja toimenpiteiden lisäksi huomiota tulisi kiinnittää kuitenkin myös uudenvuotiseen osaamiseen, jota ihmiset tarvitsevat kyetäkseen hyödyntämään ja kriittisesti arvioimaan teknologioita ja niiden vaikutuksia. Tekoälyn lukutaidon (AI literacy) käsitettä käytetään viittaamaan tähän osaamiseen tekoälyteknologioiden kohdalla.

Unicefin koostaman yhteenvedon (Penagos, Kassir & Vosloo, 2020) mukaan kansallisissa tekoälystrategioissa korostuvat erityisesti tulevaisuuden sukupolvien työelämässä tarvittavat taidot. Tekoälyn lukutaitoa tarvitaan

myös arjessa ja esimerkiksi Euroopan neuvoston (2019a) raportin mukaan mahdollisuuksia tekoälyyn liittyvään koulutukseen tulisi tarjota kaikille ja tekoäly tulisi ymmärtää tämän hetken todellisuuden, ei tulevaisuuden ilmiönä. Tekoälyn lukutaito mainitaan myös Euroopan neuvoston (2019b) suosituksessa, jossa EU:n jäsenvaltioita kehoitetaan panostamaan kansalaisten tekoälyyn liittyvän osaamisen vahvistamiseen tietoisuuden lisäämisen ja koulutuksen keinoin. Raportin mukaan näiden toimenpiteiden ei tulisi rajoittua tekoälyjärjestelmien toimintaan, vaan ulottua niiden vaikutuksiin.

Tieteellisessä kirjallisuudessa tekoälyn lukutaidon käsitteen käyttö on viimeisen seitsemän vuoden aikana selvästi lisääntynyt (Ng, Leung, Chu & Qiao, 2021), mikä kertoo tutkimuksen yleistymisestä ja aiheen tärkeydestä. Ng kollegoineen (2021) tarkasteli kirjallisuudessa käytettyjä tekoälyn lukutaidon määritelmiä ja niissä korostuivat neljä eri ulottuvuutta:

- ihmisten tietämys tekoälyteknologioista
- tekoälykäsitteiden soveltaminen erilaisissa ympäristöissä ja tilanteissa
- tekoälyn suunnitteluun ja arviointiin liittyvät näkökulmat
- eettisiin näkökulmiin liittyvät asiat.

Tekoäly muokkaa tapaamme toimia informaation kanssa

Usia lukutaitokäsitteitä esille nostettaessa tukeudutaan usein aiempiin lukutaitokäsitteisiin ja -määritelmiin (Ng ym. 2021). Kirjasto- ja



informaatioalalla tuttu informaatiolukutaidon käsite tarjoaa erityisesti tiedonhankintaa, arviointia ja käyttöä korostavan näkökulman tekoälyn lukutaitoon. Tekoälyteknologiat muovavat informaatiokäytäntöjä kuten sitä, miten informaatiota hankitaan, esitetään ja suodatetaan (Willson, 2017).

Tekoälyteknologiat ovat mahdollistaneet hakukoneiden muuntumisen kohti käyttäjien toiveita ennakoivia ja niihin mukautuvia ehdotuskoneita (suggest engine). Käyttäjän aktiivinen rooli tiedonhaussa voi silloin olla vähäisempi. Samalla hakukoneista on tullut vähemmän staattisia ja niiden tarjoamista sisällöistä vaikeammin ennakoitavia. (Haider & Sundin, 2019.) Samaan tapaan toimivat myös monet muut digitaaliset alustat, joissa sisältöjä mikro-kohdennetaan ja personoidaan.

Esimerkiksi sosiaalisen median TikTok-alustan suosittelujärjestelmää on luonnehdittu dataan hyper-reagoivaksi, sillä suosittelu ei perustu ainoastaan käyttäjän aktiiviseen toimintaan kuten tykkäykseen tai klikkauksiin vaan myös hienovaraisempaan dataan kuten siihen, kuinka pitkään käyttäjä on tiettyä videota katsonut (Zha, 2020). Lisäksi erityistä TikTokin kaltaisissa uusissa sosiaalisen median alustoissa on se, että sen keskiössä on ihmisten yhdistäminen heidän intresseihinsä vastaavan sisällön kanssa pikemminkin kuin muiden käyttäjien kanssa.

Näiden teknologioiden tarjoamat mahdollisuudet voivat muovata tapojamme kohdata ja käyttää informaatiota (Zha, 2020). Myös tapamme arvioida informaation autenttisuutta ja

luotettavuutta tulevat haastetuksi kun yhä suurempi osa digitaalisesta sisällöstä on tavalla tai toisella automaattisesti generoitua (Elliot, 2019).

Myös esimerkiksi Internetiin tuottamallamme tiedolla on merkitystä. Datan rooli on tekoälyjärjestelmistä puhuttaessa keskeinen, sillä järjestelmiä tyypillisesti opetetaan suurten datamassojen avulla, joko ihmisen valvonnassa tai ilman. Samalla data, jolla tekoälyä opetetaan voi vahvistaa vinoumia, stereotyyppioita ja ennakkokäsityksiä (Oxford Internet Institute 2021).

Vinoumia on havaittu esimerkiksi tekoälyavusteisissa päätöksentekijärjestelmissä, joita käytetään jo nyt terveydenhuollossa, pankki- ja vakuutustoiminnassa, rekrytointiprosesseissa ja jopa kaupunkien valvontajärjestelmissä. Vinoumat voivat liittyä esimerkiksi sukupuoleen, kansallisuuteen, ikään tai kulttuuristen stereotyyppien toistamiseen. (Ghallab, 2019.) Tekoälyteknologioiden oikeudenmukaisuudesta, läpinäkyvyydestä, turvallisuudesta ja vastuullisuudesta neuvotellaankin parhaillaan laajasti. Hyvä esimerkki tästä on Euroopan komission uusi tekoälyasetus (European Commission, 2021).

Tekoälyn lukutaito kansalaistaitona

Mitä moninaisempia tekoälyä hyödyntäviä ominaisuuksia sovelluksille ja datankeruulle kehitetään, sitä enemmän ne meihin vaikuttavat. Tekoäly ei kuitenkaan saisi aiheuttaa aiheetonta pelkoa ja siksi ymmärrys sen toimintaperiaatteista ja riskeistä on keskeistä. Millaista ”älyä”

on käyttämiemme ja elämiimme vaikuttavien teknologioiden taustalla?

Robinsonin (2020) mukaan pohjoismaisissa tekoälylausunnoissa korostuvat kansalaisten jaetut eettiset, vastuulliset ja ihmiskeskeiset perusteet. Tekoälyn lukutaito nähdään tulevaisuuden kansalaistaitona. Valmiudet toimia vastuullisena tekoälyn sovelluksia ymmärtävänä kansalaisena on nostettu esiin myös tutkimuskirjallisuudessa tekoälyn lukutaidon edistämisen lähtökohdaksi (Ng, Leung, Chu & Qiao, 2021).

Kirjastoilla on keskeinen rooli ihmisten valmiuksien edistämässä. Tuomi (2020) toteaa Kirjastolehdestä: ”Pohjois-Amerikassa kirjastot kantavat huolta siitä, että tekoälyn yleistyminen syventää digitaalista kuilua eri väestöryhmien välillä. Niinpä kirjastot haluavat tukea asiakkaitaan tekoälyn ymmärtämisessä.” Asiakkaiden lukutaitojen edistäminen on niin yleisten kuin tieteellisten kirjastojenkin kes-

keinen tehtävä.

Muun muassa tietokannat ja tiedonhankinnan tukijärjestelmät hyödyntävät yhä enemmän tekoälyä, ja niitä käyttävän voi olla tarpeen tunnistaa niiden toimintaperiaatteita. Erityisesti tieteellisissä kirjastoissa olennaiseksi voivat tulla myös tekoälyavusteisten tutkimusaineistojen keruun, käsittelyn ja säilyttämisen kysymykset (Cox, Pinfield & Rutter 2019). Tulevaisuudessa tekoäly voi olla kiinteä osa sekä opiskelijoiden että kirjastoammatilaisten työnkuvaa myös tavoilla, joita ei vielä kyetä ennustamaan. Tekoälyteknologioiden käytön ja vaikutusten tunnistaminen on tärkeää paitsi teknologioihin liittyvien riskien välttämiseksi, myös siksi, että voidaan tunnistaa näiden teknologioiden tarjoamia mahdollisuuksia. Tämän ymmärryksen kautta pystymme tukemaan kirjastojen asiakkaita tekoälyn lukutaidon valmiuksien karttumisessa. ❖

AI

END

Lähteet

- ADADI, A., & BERRADA, M., 2018. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (xAI). in *IEEE Access*, 6, 52138–52160. doi: 10.1109/ACCESS.2018.2870052.
- Cox, AM., Pinfield, S., & Rutter, S., 2019. The intelligent library: Thought leaders' views on the likely impact of artificial intelligence on academic libraries, in *Library Hi Tech*, 37(3), 418–435. <https://doi.org/10.1108/LHT-08-2018-0105>
- ELLIOT, A., 2019. *The Culture of AI: Everyday Life and the Digital Revolution*. Routledge.
- European Commission 2021. Proposal for a Regulation of the European Parliament and of the Council. Laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts. Explanatory memorandum. <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&curi=CELEX%3A52021PC0206>
- Euroopan neuvosto 2019a. Seminar report. Artificial Intelligence and its impact on young people. <https://rm.coe.int/ai-report-bil-final/16809f9a88>

Euroopan neuvosto 2019b. Unboxing Artificial Intelligence: 10 steps to protect Human Rights. Recommendation. <https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/168094664>

GHALLAB, M., 2019. Responsible AI: requirements and challenges. *AI Perspectives*, 1(1), 1-7.

HAIDER, J., & SUNDIN, O., 2019. Invisible Search and Online Search Engines: The Ubiquity of Search in Everyday Life. Routledge.

NG, DTK., LEUNG, JKL, CHU, KWS., & QIAO, MS., 2021. AI Literacy: Definition, Teaching, Evaluation and Ethical Issues. *Proceedings of the Association for Information Science and Technology*, 58(1), 504-509. <https://doi.org/10.1002/praz.487>

Oxford Internet Institute 2021. Pieni tekoälyopas. Vinoumat. <https://atozofai.withgoogle.com/intl/fi/bias/>

PENAGOS, M., KASSIR, S., & VOSLOO, S., 2020. National AI strategies and children. *Reviewing the landscape and identifying windows of opportunity*. UNICEF Policy brief. <https://www.unicef.org/globalinsight/media/1156/file>

ROBINSON, SC., 2020. Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI). *Technology in Society*, 63. <https://doi.org/10.1016/j.techsoc.2020.101421>

SLOTTE DUFVA, T., & MERTALA, P., 2021. Sähköä ja alkemiaa : tekoälydiskurssit Yleisradion verkkoartikkeleissa. *Media ja viestintä*, 44(1), 95-115. <https://doi.org/10.23983/mv.107302>

Sitra 2020. Megatrends 2020. <https://www.sitra.fi/en/topics/megatrends/>

TUOMI, P., 2020. Mikä on tekoälyn seuraava siirto? *Kirjastolehti*, 29.6.2020. <https://suomenkirjastoseura.fi/kirjastolehti/mika-on-tekoalyn-seuraava-siirto/>

WANG, Z., 2020. Exploring different notions of literacy: a literature review analysis of literacy research related to Artificial Intelligence and Big Data application. *IOP Conference Series: Materials Science and Engineering*, Volume 806, International Conference on AI and Big Data Application (AIBDA 2019) 20-22 December 2019, Guangzhou, China: <https://iopscience.iop.org/article/10.1088/1757-899X/806/1/012023/pdf>

WILLSON, M., 2017. Algorithms (and the) everyday. *Information, Communication & Society* 20(1), 137-150. <https://doi.org/10.1080/1369118X.2016.1200645>

ZHA, X., 2020. The unique power of TiTok's algorithm. The Interpreter. <https://www.lowyinstitute.org/the-interpreter/unique-power-tiktok-s-algorithm>

NOORA HIRVONEN

noora.hirvonen@oulu.fi

Informaatiotutkimus, Oulun yliopisto

HEIDI ENWALD

heidi.enwald@oulu.fi

Informaatiotutkimus, Oulun yliopisto

Maarit Cederberg, Tuula Hurskainen,
Elisa Mörsky, Sirpa Janhonen, Eija Kajava,
Juha Leppämäki ja Annamari Leppäniemi

Tietoa vahvistamassa ja saatavuutta parantamassa – NordILL 2021 -konferenssin satoa

The 14th Nordic Resource Sharing, Reference and Collection Management Conference (25.-26.10.2021) tarjosi innostavien puhujien lisäksi laajan katsauksen pohjoismaisten kirjastojen innovaatioihin sekä koronapandemian aiheuttamiin muutostarpeisiin kirjastojen palveluissa. Artikkelit luovat katsauksen konferenssin teemoihin: aineistojen digitointiin, e-lainaamiseen yli kirjasto- ja maarajojen, asiakastoimisuuden parantamiseen, avoimeen tieteeseen sekä kulttuuriperintöportaaleihin.

Perinteiset pohjoismaiset kauko- palvelupäivät järjestettiin loka- kuun lopulla koronapandemian takia vuosi alkuperäistä suunnitelmaa myöhemmin ja aivan uudella konseptilla: virtuaalisena. NordILL 2021 -konferenssin järjestelyistä päävastuun kantoivat Kansalliskirjasto ja Varastokirjasto, mukana järjestelyissä olivat myös Helsingin yliopisto, Helsingin kaupunginkirjasto ja Helsingin kaupunki. Konferenssissa kuultiin esitykset Pohjoismaiden lisäksi myös Virossa, Latviasta ja Venäjältä.

Koska tietoa täytyy jakaa!

Konferenssin aloituspäivän pääpuhujiksi oli kutsuttu sekä IFLA-organisaation tuore presidentti Barbara Lison Saksasta että IFLA:n Document Delivery and Resource Sharing -komitean puheenjohtaja Peter Collins, joka toimii OCLC:n Resource Sharing -johtajana USA:ssa. Toisen konferenssipäivän teemojen tunnelmaan viritti Norjan



Making of NordILL 2021. Studiassa ja ruudussa pohjoismaisia kauko- palvelupäiviä toteuttamassa Liisa Savolainen, Johanna Vesterinen ja Päivi Jokitalo – teknistä toteutusta valvomassa Panu Kari. Kuva: Jussi Saari

kansalliskirjaston johtaja Aslak Sira Myhre.

Barbara Lison kertoi avauksessaan Euroopan aineistonvälityksen tulevaisuuden trendeistä avaamalla IFLA:n strategisen johtamisen teemoja. IFLA:n tavoitteena on mahdollistaa e-aineistojen kaukolainaaminen ja tarjota kaikille pääsy sähköisiin resursseihin. Lisäksi parhaillaan ollaan muuttamassa kansainvälisen kaukopalvelun maksuvälinettä IFLA-vuocheria sähköiseksi.

Peter Collins korosti IFLA:n roolia sähköisen aineiston lainaamisen organisoinnissa. Collins muistutti, että

monia kansainvälisen tason juridisia kysymyksiä ja aineiston saatavuuden rajoitteita on edelleen ratkaisematta. Hän nosti keskusteluun painetun aineiston tulevaisuuden ja eri formaateissa tarjot-

tavan aineiston tarpeen kirjastoissa. Sekä yleisten että tieteellisten kirjastojen kaukolainatilastot laskivat dramaattisesti vuonna 2020 pandemian vaikutuksesta sysäten liikkeelle digitointiprojekteja.

Aslak Sira Myhre näki tärkeänä digitointityön. Myhre korosti, että kirjastojen täytyy herätä keskustelemaan tulevaisuuden tiedonhakatavoista ja -välineistä. Digitoituna vanha kulttuuriperintö herää uudelleen henkiin. Tämä edellyttää kirjastoilta pääsyn luomista ja ohjaamista eri formaateissa olevaan aineistoon. Asian ytimessä ovat

kirjastoammattilaisten asiantuntemus ja luetteloinnin hallinta.

E-kirjojen lainaus Pohjoismaissa

Oman esityksensä konferenssissa saivat yleisissä kirjastoissa käynnissä olevat kansalliset e-kirjastoprojektit, joita esiteltiin Suomesta, Tanskasta ja Norjasta. Kansallisten projektien keskeinen tavoite on, että e-aineistoihin pääsy ja valikoima ei riippuisi asuinpaikasta. Tavoite on, että rinnakkaiset sähköiset aineistopalvelut saataisiin yhdelle alustalle.

Eri maiden e-kirjastoprojektit ovat hyvin eri vaiheissa. Mikkel Christoffersen esitteli Tanskan eReolen, joka on kymmenen vuotta käytössä ollut yleisten kirjastojen yhteinen e-kirjojen ja e-äänikirjojen portaali. Trygve Kikut puhui Norjan projektista, jossa alustan luominen on vasta suunnitteluvaiheessa; Virva Nousiainen-Hiiri ja Suvi Sivulainen kertoivat Suomen Digimedia-hankkeesta, jonka pilotointi on tarkoitus aloittaa vuonna 2022. Esitellyissä e-kirjaprojekteissa vaiheineen ja haasteineen voi havaita paljon samankaltaisuutta.

Kaikille maille yhteinen pohdinnan aihe on palvelun rahoitusmallin luominen kirjastojen ja niitä rahoittavien kuntien kesken. On myös olennaista rakentaa luottamusta kirjastojen ja erilaisten muiden toimijoiden, kuten kustantajien, jakelijoiden ja tekijöiden välille. Kirjaston tulisi passiivisen ostajan sijasta ottaa rooli aktiivisena toi-

Kirjastojen täytyy herätä keskustelemaan tulevaisuuden tiedonhakatavoista ja -välineistä.

minnan kehittäjänä ja sopijana niin, että kaikki osapuolet hyötyisivät mukana olostapalvelussa.

Tanskassa kirjastot ovat onnistuneet luomaan hyvää yhteistyötä kustantajien kanssa. Kirjaston palvelu on todettu erityisen hyväksi sekä vanhemman kirjallisuuden että lasten- ja nuorten kirjallisuuden esiin nostamiseen. Isojen kustantajien uusia nimekkeitä on esillä erityisesti siksi, että ne houkuttelevat uusia käyttäjiä palveluun, vaikka varausjonot kasvavat välillä pitkiksi.

Koronasulkujen aikana Tanskan eReolenin käyttö lisääntyi ja palvelusta tuli niin suosittu, että sen toimintamallia joudutaan kehittämään korkeiden kustannusten vuoksi. Sopivan palvelutason määrittäminen on haasteellista, mutta tavoitteena on edelleen kiinnostavan palvelun tarjoaminen asiakkaille.

Uusia kehityssuuntia kaukopalvelussa ja aineistojen yhteiskäytössä

Riin Olonen Viron kansalliskirjastosta esitteli uutta valtakunnallista lainausjärjestelmää Books on the Move, joka otetaan käyttöön ensi vuoden aikana. Sen sijaan, että kirjastoa pyydetään tilaamaan kirjat toisesta kirjastosta, tarjotaan asiakkaille mahdollisuus tilata kirjoja verkossa ja noutaa ne haluamastaan kirjastosta. Järjestelmä ohjaa pyynnön oikeaan kirjastoon. Palvelua kehitetään edelleen, sillä Virossa on 800 kirjastoa ja erilaisia kirjastojärjestelmiä käytössä. Lisäksi Virossa on kehitteillä Ystävältä

ystävälle -palvelu, jossa kirjallina voidaan siirtää lainaajan tililtä toiselle ja samalla vastuu siirtyy uudelle lainaajalle. Molempien lainaajien täytyy olla rekisteröityneitä käyttäjiä eikä kirjaan saa olla varauksia.

Lovisa Karlsson Kjellin ja Tobias Willstedt Ruotsin kansalliskirjastosta kertoivat Bläddrasta, joka on e-kirjasovellus mobiililaitteille ja tableteille. Bläddralla on ilmainen pääsy Ruotsin kansallisten vähemmistökielien kirjoihin. Näitä kieliä ovat suomi, jiddish, meänkieli, romani ja saame. Kansalliskirjasto aloitti palvelun vuonna 2019 kahdella vähemmistökielillä, saamalla ja romanilla.

Frida Leijon Lundqvist esitteli Tukholman yliopiston asiakaslähtöistä kaukopalvelu-, aineistonhankinta- ja open access -mallia. Kirjastossa on kolme ryhmää: hankinta, kaukopalvelu ja lisenssi/open access -ryhmä, jotka toimivat yhteistyössä. Asiakkaiden hankintapyynnöt otetaan vastaan yhdellä lomakkeella. Sen jälkeen eri tapaukset ohjataan jollekin kolmesta osastosta sen perusteella, mikä niistä on paras hankintatapa kyseiselle aineistolle. Koska kaukopalvelu on painottunut vanhemman aineiston hankintaan, uudemman aineiston open access -kehitys ei ole juurikaan vaikuttanut kaukopalvelutilausten määrään.

Jyväskylän kaupunginkirjaston Shokkiuutinen-hankkeen projekti-

Tanskassa kirjastot ovat onnistuneet luomaan hyvää yhteistyötä kustantajien kanssa.

updating postcovid close work
 inspiration hope resource sharing
 infodemics insights knowledge future ideas
 tips new ideas e-lending
 collaboration vision insight greater understanding
 news

Konferenssissa tietyt teemat ovat toistuneet vuosien varrella,
 vaikka uuttakin oppii. Kuva: Jussi Saari

suunnittelija Anne Koivisto kertoi hankkeen kolmesta osahankkeesta: Jyväskylän kaupunginkirjaston varastokokoelman läpikäynti auttaa uuden kirjastotilan vähenevissä kokoelmatiloissa, kun aineistoa siirretään Varastokirjastoon. VATI-projekti parantaa asiakaslähtöisyyttä, Varastokirjaston kokoelmat näkyvät KESKI-Finnaan kirjautuneelle asiakkaalle ja hän voi myös itse tehdä tilauksen Varastokirjaston aineistoon. ATTE-projektissa tehtiin esiselvitys kirjastojärjestelmäriippumattoman kaukopalvelun mahdollisuuksista. Asiakaslähtöisen, järjestelmäriippumattoman ja standardin mukaisen yhteisen kaukopalveluratkaisun etsintää jatketaan Varastokirjaston AIRUT-projektissa.

Elina Kähö ja Juha Leppämäki esittelivät Helsingin yliopiston kirjaston Alma-kehittämishanketta. Kaukopalvelun useimmat erilliset järjestelmät korvattiin Alma-kirjastojärjestelmän kaukopalveluominaisuuksilla. Tässä hyödynnetään sekä asiakasliittymä Primon kaukopalvelulomaketta että Alman rajapintoja tilausten vastaanottamiseen asiakkailta ja muilta kirjastoilta. Asiakaslähtöisyys lisääntyi niin, että nyt kaukolainat näkyvät Helkan asiakastiedoissa muiden lainojen ja varausten tapaan.

Avoimet tiedejulkaisut

Avointen tiedejulkaisujen (Open Access Journals) kansallisia alustoja esiteltiin neljästä maasta. Palvelut tarjoavat erilaisille toimijoille teknisen alustan tieteellisten lehtien julkaisemiseen.

Jesper Boserup Thestrup edusti tanskalaista Tidsskrift.dk -palvelua, joka on ollut käytössä vuodesta 2007. Avoimen julkaisualustan käyttö on pelastanut tieteellisiä julkaisuja lakkauttamiselta, ja alustalla saavat näkyvyyttä myös historialliset tiedejulkaisut.

Suomesta Antti-Jussi Nygård nosti esiin Journal.fi-alustan, jota ylläpitää Tieteellisten seurain valtuuskunta. Tavoitteena ollut edullinen alusta open access -julkaisemiseen toteutui. Alustan julkaisuvuodesta 2017 lähtien tarjolla olevien julkaisujen määrä on kasvanut jatkuvasti.

Jeroen Sondervan Alankomaista esitteli tänä vuonna julkaistun palvelun Openjournals.nl. Julkaisijoille asetetaan palvelussa joitakin vaatimuksia, kuten avointen Creative Commons -lisenssien käyttö. Myös toimitusprosessille asetetaan ehtoja, esimerkiksi vertaisarviointi.

Ruotsin kansalliskirjaston ylläpitämästä Publicera-alustasta kertoi Sofie Wennström. Alustan käyttö on julkaisijoille ilmaista, mutta esimerkiksi julkaisuformaattia rajoitetaan. Publicera



vaatii myös CC-lisenssiä ja vertaisarviointiprosessia.

Kaikkien esiteltyjen palveluiden yhteinen visio on, että alustoilla julkaistut avoimet tiedejulkaisut otettaisiin vakavasti tieteellisinä lähteinä. Tähän pyritään esimerkiksi Ruotsin ja Alankomaiden vaatimuksilla, joiden tarkoituksena on saada toimijat automaattisesti noudattamaan tieteellisen julkaisemisen standardeja. Yhteiset alustat nostavat avointen tiedejulkaisujen näkyvyyttä ja helpottavat niiden löytämistä. Alustojen tavoitteena onkin erikielisten julkaisujen helpompi kansainvälinen saatavuus.

Tieteelliset kirjastot tukevat avointa tiedettä

Arja Tuuliniemi Kansalliskirjastosta kertoi FinElib-konsortioista, joka neuvottelee vuosittain suomalaisten jäsenorganisaatioidensa puolesta e-aineistojen hankinnasta ja edistää avointa julkaisemista.

Yksi askel avoimen julkaisemisen (open access) tukemisessa ovat transformatiiviset sopimukset, jotka sisältävät sekä avoimen lukuoikeuden että avoimen julkaisuoikeuden kyseisen kustantajan aineistoihin. Sopimuksissa on myös ehtona välitön avoin julkaisu ilman, että aineistoa ensin pidetään maksumuurin takana. Transformatiiviset sopimukset ovat väliaikaisratkaisu, jolla kannustetaan ja rahoitetaan siirtymää kokonaan avoimeen julkaisemiseen. Avointa julkaisemista tukemalla on päästy siihen, että julkaisijat ovat

kehittäneet omia käytäntöjään ja bisnessmallejaan. Myös muiden tahojen tietoisuus avoimesta julkaisemisesta on jatkuvasti lisääntynyt ja yhä suurempi osa artikkeleista julkaistaan avoimina.

Michael Svendsen ja Lise Baltzer puolestaan kertoivat, kuinka Tanskan kansalliskirjastossa ja erityisesti Kööpenhaminan yliopiston kirjastossa tuetaan avointa tiedettä. Yliopiston kanssa on systemaattisesti kehitetty alustoja, työkaluja ja palveluita tukemaan avointa tiedettä. Kyse on hankintapolitiikan lisäksi myös siitä, että tutkijoita ja kirjoittajia tuetaan konkreettisesti avoimessa julkaisemisessa. Tanskan kansallisen vision tavoite on, että ajan myötä kaikki Tanskassa julkaistava tutkimus olisi avoimesti saatavissa. Tänä vuonna luku on jo 60 %.

Kulttuuriperintöportaalit Suomessa ja Latviassa

Riitta Peltonen Kansalliskirjastosta kertoi Finna-portaalista, joka tarjoaa yhdellä haulla pääsyn kaikkeen kansalliseen materiaaliin. Finnan kautta pääsee suoraan digitaaliseen aineistoon kuten kuviin, äänitallenteisiin, videoklippeihin, elokuviin, karttoihin ja alkuperäisiin tekstiaineistoihin. Mukana ovat lähes kaikki Suomen yleiset ja tieteelliset kirjastot, museot ja arkistot.

Digi.kansalliskirjasto.fi-palvelua esitteli Tuula Pääkkönen Kansalliskirjastosta. Tällä hetkellä palvelussa on digitoituna noin 14 miljoonaa sanomalehtisivua ja seitsemän miljoonaa

aikakauslehtisivua. Palvelussa on mukana myös muuta aineistoa, kuten väitöskirjoja, käsikirjoituksia, kirjoja ja pienpainatteita.

Latvian kansalliskirjaston johtaja Arturs Zogla kertoi kirjojen ja lehtien digitoinnista. Kansalliskirjaston kautta on digitaalisessa muodossa saatavilla noin 13 000 kirjaa ja 1 500 lehtinimekettä. Latviassa koronavirustilanne eskaloitui nopeasti vuoden 2020 maaliskuussa, jolloin NLL (National Library of Latvia) suljettiin kokonaan yleisöltä. Kun samaa aikaa pääsyä useimpiin yleisiin kirjastoihin rajoitettiin, tuli tarve siirtää painetusta aineistosta digitaaliseen. Tekijänoikeudellisista haasteista huolimatta lehdet avattiin osapuolten yhteisellä sopimuksella ilmaiseen käyttöön.

Yhteenveto

Kahden vuoden välein järjestettävä pohjoismainen konferenssi on tarpeellinen ja antoisa tapa kokoontua yhteisten asioiden äärelle. Aina oppii uutta, vaikka tuttua ja vanhaakin on mukana. Esimerkiksi tietyt teemat ovat toistuneet vuosien varrella, ja tulevat varmasti toistumaan edelleen, kuten

avoin saatavuus ja e-lainaus. Niihin liittyy haasteita, joita yritetään yhdessä ratkoa.

Konferenssissa esiin nousseet yhteiset päämäärät ja toiveikkuus loivat kansainvälistä yhteisöllisyyttä, vaikka kollegojen kohtaaminen, vapaa keskustelu ja yhteiset ruokailuhetket puutuivat – niitä toki toivomme jatkossa. Paikka, jossa kohdataan kahden vuoden kuluttua on vielä arvoitus, mutta perinteisesti tapaamiset on järjestetty vuorotellen Suomessa, Norjassa, Ruotsissa, Tanskassa ja Islannissa. Yleensä järjestäjävastuu menee järjestyksessä niin, että se jolla on pisin aika konferenssin järjestämisestä on seuraavana vuorossa. Odotamme innolla!

Kiitokset innovatiivisesta ja ajatuksia herättävästä NordILL 2021 -konferenssista erityisesti Varastokirjaston kirjastonjohtaja Johanna Vesteriselle ja Kansalliskirjaston apulaiskirjastonjohtaja Liisa Savolaiselle sekä tapahtuman etenemisestä konferenssin aikana vastanneelle moderaattori Päivi Jokitalolle! Järjestelyt ja yhteydet toimivat erinomaisesti. 💎

Kirjoittajat

MAARIT CEDERBERG

Tampereen yliopiston kirjasto
maarit.cederberg@tuni.fi

TUULA HURSKAINEN

Suomen ympäristökeskuksen tietopalvelu
tuula.hurskainen@syke.fi

SIRPA JANHONEN

Varastokirjasto
sirpa.janhonen@nrl.fi

ELISA MÖRSKY

Kansalliskirjasto
elisa.morsky@helsinki.fi

EIJA KAJAVA

Eduskunnan kirjasto
eija.kajava@eduskunta.fi

JUHA LEPPÄMÄKI

Helsingin yliopiston kirjasto
juha.leppamaki@helsinki.fi

ANNAMARI LEPPÄNIEMI

Turun yliopiston kirjasto
annamari.leppaniemi@utu.fi

Avoimuus luo edellytykset osuvampaan metriikkaan – Eva Isakssonin haastattelu

Tieteellisen julkaisemisen viittausdatan avoin saatavuus, rikas metadata ja avoimet julkaisut luovat tulevaisuudessa palveluita, joista ”ei ole vielä hajuakaan”, arvioi Helsingin yliopiston kirjaston bibliometriikkaan erikoistunut tietoasiantuntija Eva Isaksson. Tässä haastattelussa eläkkeelle jäänyt Isaksson kertoo työuransa käännteistä, yhteistyöverkostojen merkityksestä ja hahmottelee julkaisumetriikan kehitystä, avoimen tieteen mahdollisuuksia sekä tasa-arvonäkökulman ja metriikan yhteyksiä.

Helsingin yliopiston kirjaston tietoasiantuntija Eva Isaksson tunnetaan kansallisesti ja kansainvälisesti työstään erityisesti tähtitieteen kirjastotoiminnan, julkaisumetriikan ja tasa-arvoasioiden parissa. Elokuun lopussa Isaksson jäi eläkkeelle, ja tässä haastattelussa hän kertoo työuransa vaiheista ja ajatuksesta erityisesti bibliometriikkapalveluiden kehityksestä.

Signum-haastattelua valmistellessa Isaksson oli jo ehtinyt totutella uuteen elämänvaiheeseen. Miten ensimmäiset pari kuukautta ovat sujuneet eläkkeellä?

”Ohjelmassa on ollut oman juhlatuennon tallenteen tekstittämistä ja sadonkorjuuta – olen innokas vihanpespuutarhuri. Mielessä on ollut jopa kolme eri kirjaprojektia, ja olen näin ensi alkuun odotellut, mikä niistä imaisee mukaansa”, Isaksson kertoo.

Fysiikan opinnoista tähtitieteen kirjastonhoitajaksi

Eva Isakssonin tausta on luonnontieteissä, fysiikassa, joka valikoitui omaksi alaksi osittain kuulovamman takia – ”matematiikassa kaavat eivät jutelleet”. Tie johti kirjastoon, kun jatko-opinnot piti rahoittaa tähtitieteen kirjaston-

Digitaaliset kirjat monipuolistavat käyttöä, mutta Isaksson toivoo painettujen kokoelmien säilyvän. ”Urani parhaita asioita on ollut se, että olen saanut olla tekemisissä painettujen kirjojen kanssa. Selailusta tuleva elämys on ollut yksi kirjastojen vahvimpia asioita.”



hoitajana 1980-luvun alussa.

”Sain 1981 teoreettisen fysiikan graduni valmiiksi. Graduohjaajani Raimo Keskinen, joka oli todella mahtava ihminen, etsi aktiivisesti ohjattavilleen työpaikkoja. Ensin ei tärpännyt – ja siksi en tehnyt kirjastoalan työuraani Otaniemessä. Sen sijaan Helsingin yliopiston tähtitieteen laitoksen kirjastossa tärppäsi. Tähtitieteen professori taisi peräti tulla opiskelija-asuntolaan minua värväämään.”

”Aivan ensiksi jouduin käymään lävitse laitoksen massiivisen laitos-sarjakokoelman, joka on monin tavoin erityinen esimerkki avoimesta tieteestä: kaukana toisistaan sijaitsevat observatoriot lähettivät toisilleen tutkimustuloksiaan ja muita julkaisujaan ilmaiseksi. Niitä kävin aluksi läpi kortiston avulla toista vuotta. Siitä se lähti.”

Rankingeista kohti vastuullista metriikkaa

Bibliometriikka astui Isakssonin työnkuvaan voimakkaasti 2000-luvun alussa, kun tutkijat alkoivat kiinnostua uusista julkaisudatan analytiikkavälineistä, kuten Web of Sciencesta.

”Jo aikaisemmin olimme tähtitieteessä panostaneet tähtitieteen julkaisujen keräämiseen. Merkittävänä apuna oli tähtitieteen oma bibliografinen viitetietokanta, Astrophysics Data System. ADS oli ensimmäinen avoin tietokanta, jossa viittaukset olivat avoimesti saatavilla. Se aloitti jo

1990-luvulla, ja sen käyttäjä orientoitui alusta pitäen näkemään julkaisut osana viittausten ja tutkimusdatan verkkoa.”

Bibliometriikka on ytimeltään julkaisemista koskevaa tietoa, jota oli totuttu käyttämään myös erilaisten vertailujen ja rankinglistausten tekemiseen. Etenkin lääketieteessä mittamisella oli pitkät perinteet.

”Lääketieteessä julkaistaan valtavasti, ja siellä on paljon lehtiä. Lääketiede on myös ehkä voimakkaammin kilpailtu kuin esimerkiksi fysiikka. Lääketieteessä kerättiinkin kaikkea viitetausdataa ja lehtiä arvotettiin hyväksi ja huonoiksi impact factoreilla (IF). Siellä on jo usean vuosikymmenen ajan sitouduttu siihen, että tutkijan on saatava paljon dataa julkaisemisesta, mikä taas on ohjannut metriikkaa voimakkaasti.”

Samalla kun julkaisumetriikan rooli on korostunut muillakin tieteenaloilla, on hämärtynyt se, mitä impaktifaktorit tai Julkaisufoorumi-luokitukset (Jufo) mittaavat: ne on tarkoitettu julkaisukanavien vertailuun, ei tutkijoiden saati yksittäisten tutkimusten arviointiin.

”Metriikassa on monta vuotta työskennelty sen eteen, että tutkijalistaukset saataisiin pois näkyvistä. Jos ne perustuvat lehtien impact factoreihin tai Jufo-tasoihin, kyse ei ole vastuullisesta metriikasta, vaikka tutkijat itse suostuisivat tähän.”

Metriikkapalvelu syntyi keskellä leikkauksia

Julkaisumetriikan ja tutkimuksen arvioinnin merkityksen kasvaminen on

vaikuttanut Isakssonin työuran aikana myös palveluihin. Helsingin yliopiston kirjaston keskitetty metriikkapalvelu aloitti vuonna 2015 kirjastoalan vaikuttajan ja tutkimuspalveluiden uranuurtajan Maria Forsmanin aloitteesta.

”Maria oli luonut pohjan metriikkatyölle, mutta hän jäi eläkkeelle 2015. Samassa yhteydessä kirjastorakennetta uudistettiin, ja tehtiin päätös keskitetyistä tutkimuksen palveluista, joiden osana olisi metriikka. Päädyin vetämään metriikkatiimiä”, Isaksson kertoo.

Samaan aikaan osuivat kuitenkin Helsingin yliopiston leikkaukset, jotka laittoivat metriikan käytön vastuullisuuden toden teolla koetukselle. Yliopistolaisista kerättiin tietoa, ja myös Jufo-tarkasteluista oltiin kiinnostuneita.

”Totesimme, että Jufo-luokitusta ei voi käyttää yksittäisten tutkijoiden tarkasteluun. Tarjosimme korvikkeeksi kullekin tutkimusalalle soveltuvasta tietokannasta poimittuja viittauslukuja. On vaikea sanoa, miten nämä tiedot lopulta vaikuttivat, mutta se oli järkyttävää ja hävettävää, oman urani aallonpohja”, Isaksson toteaa.

Vaikean syntäjäkohdan jälkeen metriikkapalvelu on vakiinnuttanut paikkansa kirjaston palveluvalikoimassa, ja kirjastolaisia – Eva Isakssonin johdolla – on osallistunut myös kansalliseen työhön, muun muassa vastuullisen metriikan edistämiseksi.

”Yliopistot ovat kiitettävästi huomioineet vastuullisen metriikan periaatteita ja sitoutuneet DORA-julistukseen.

Elämme kuitenkin vielä välivaihetta, jossa takaiskut rahoituksessa ja resursseissa voivat johtaa siihen, että metriikalta toivotaan apua, kun joudutaan päättämään, mistä leikataan. Viittauksista saatavia indikaattoreita on käytettävä varovasti ja ajatuksella.”

Tutkimuksen arvioinnissa kulttuurillinen murros

Yhtenä työuransa kohokohdista Isaksson mainitsee osallistumisen tutkimuksen arvioinnin kehittämiseen.

”Helsingin yliopistossa on panostettu todella paljon siihen, että prosessi toimisi ja että saisimme kaikki meille kuuluvat julkaisut kirjattua. Kun tiedämme, mitä julkaisudatassa on, olemme voineet tehdä tarkasteluja tieteenalakohtaisesti. Ei se aina helppoa ole ollut, mutta olen tyytyväinen siihen, että viimeisin arviointi 2018–2019 meni kuin Strömsössä.”

Arvioinnissa on tapahtunut myös tietty kulttuurillinen murros, kun on alettu entistä paremmin ymmärtää, mitä luvut voivat ja eivät voi kertoa. Tämä on vahvistanut yliopistokirjaston roolia asiantuntijatahona, mutta myös metriikkaa käyttäviltä yliopistolaisilta vaaditaan parempaa ymmärrystä.

”Bibliometriikassa on yhä voimakkaampana trendi, jossa ei etsitä vertailuja, vaan yhteyksiä, eli julkaisemisen sisäistä dynamiikkaa. Tutkimuksen arvioinnissakaan tärkeintä ei ole se, että arvioitavat saavat lukuja itsestään, vaan tärkeintä on se, että he saavat tietoa sii-



Eva Isaksson

- Toiminut Helsingin yliopiston kirjaston metriikaryhmän vetäjänä, julkaisumetriikan kansallisten ja kansainvälisten verkostojen asiantuntijana sekä Helsingin yliopiston tasa-arvo- ja yhdenvertaisuustoimikunnan pitkäaikainen jäsenenä.
- Vastuullisen julkaisumetriikan ja tutkimuksen arvioinnin edelläkävijä ja puolestapuhuja. Isaksson on toiminut pitkään arXiv Member of Advisory Board -jäsenenä ja LISA (Library and Information Services in Astronomy) -konferenssin järjestäjätoimikunnassa.
- Tietokirjailija, julkaissut teoksia mm. tasa-arvon ja rauhanaatteen teemoista (Rauhan laboratoriot, 1985 ja Nainen ja maailmankaikkeus, 1987). Moni teos avoimesti saatavilla Helsingin yliopiston Helda Open Books -kokoelmassa. Suunnitelmassa on jatkoteos Rauhan laboratorioille: ”Haluaisin kirjoittaa tutkijoista, jotka ovat tehneet sotaan ja aseiden kehittämiseen liittyvää tutkimustyötä. Mikä on ollut se konteksti, joka on ajanut heitä tähän tutkimustyöhön?”
- Eva Isakssonille on myönnetty Suomen Valkoisen Ruusun ansioristi vuonna 2017, Helsingin yliopiston Maikki Friberg -tasa-arvopalkinto vuonna 1999 ja Tiedonjulkistamisen valtionpalkinto vuonna 1986.

tä, missä kontekstissa heidän tutkimuksensa on tapahtunut. Viimeisimmässä arvoinnissa kävi ilmi, että yksiköissä on paljon vanhentuneita näkemyksiä siitä, mikä metriikan tuottamissa tuloksissa on tärkeää. Siitä sai alkunsa Bibliometriikkaa yliopistolaisille -kurssit, joiden tarkoitus on antaa metriikkaa hyödyntäville ihmisille perustiedot bibliometriikasta”, Isaksson sanoo.

Metriikan vallankumous vasta edessä

Merkittävimpänä metriikan välineenä Isaksson mainitsee 2010-luvulla ilmestyneet analytiikkatietokannat, joissa tietokantojen koko on mahdollistanut monenlaiset vertailut, trendianalysit ja klusteroinnit.

Viimeisimpänä trendinä julkaisumetriikassa on ollut viittausdatan avautuminen, mikä Isakssonin mukaan tulee mullistamaan julkaisutoiminnan tarkastelun.

”Kun ennen voitiin katsoa, mitä viittausluettelossa on, tällä hetkellä voidaan tekoälyn avulla katsoa, missä kontekstissa viittaaminen tapahtuu. Rikas metadata ja avoin saatavuus tarkoittavat sitä, että joku jossain kehittää jo tällä hetkellä palvelua, joka tulee mullistamaan näkemyksemme metriikasta. Uskon, että meillä on muutaman vuoden sisällä käytössä palveluita, joista ei ole mitään hajua tällä hetkellä.”

Palveluvalikoiman monipuolisuus tarkoittaa sitä, että metriikan asiantuntemusta tarvitaan entistä

« Monia kollegoita kiinnostaa, mistä Eva Isaksson saa uskomattoman määrän energiaa, joka näkyy hänen työssään ja persoonassaan. ”Kai se on uteliaisuus, joka ei ole hellittänyt vuosien varrella.”



enemmän. Myös julkaisemisen tavat digitaalisessa ympäristössä tulevat muuntautumaan.

”Tästä on tulossa iso viidakko. Meille tullaan tarjoamaan palveluja, lisäosia ja lisää metriikkaa. Tutkijat voivat viehättyä uudesta palvelusta, ja haluavat sen käyttöön. Metriikassa täytyy pystyä tunnistamaan, onko se relevantti, kertooko se oikeita asioita, onko se vakaata dataa – ja ennen kaikkea, kannattaako siitä maksaa.”

”Vaaranpaikka on ollut jo pitkään se luulo, että mitä nopeammin tietoa saadaan, sen parempi. Tämähän on altmetriikan sisäänrakennettu ajatus. Silloin ohitetaan se, mikä merkitys julkaisulla oikeasti on, ja missä ajassa sen voi saada selville. Jopa altmetriikassa voi hitaudella olla roolinsa. Todellisen vaikuttavuuden näkee niistä yhteyksistä, joita ei voi saada selville heti.”

Verkostot työn tukena

Isaksson on toiminut aktiivisesti sekä kansallisissa että kansainvälisissä verkostoissa. Henkilökohtaisesti tärkeimpänä hän pitää tähtitieteen kirjastoalan kansainvälistä verkostoa, erityisesti LISA-konferensseja (Library and Information Services in Astronomy), jotka tekivät hänestä kirjastolaisen. Kotimaisista keskeisimpänä Isaksson pitää kansallista julkaisumetriikkaverkostoa, joka toimii osana korkeakoulujen ja tutkimuslaitosten laaja-alaista Finn-ARMA-yhteistyöverkostoa.

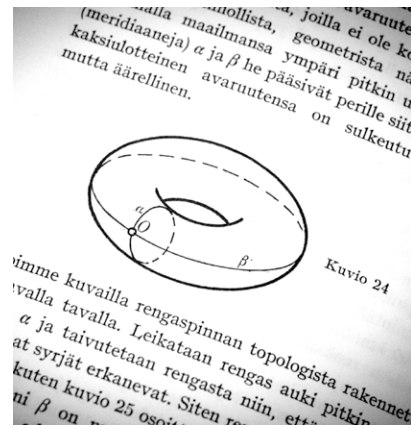
”Jaamme julkaisumetriikkaverkostossa yliopistojen kesken näkemyksiä ja

kokemuksia siitä, mitä bibliometriikassa on meneillään, mitä eri yliopistoissa tehdään ja miten siellä resursoidaan palveluita. Iso asia, joka tänä vuonna on laitettu liikkeelle, on kansallisen bibliometriikkaoppaan toteuttaminen kansallisena yhteistyönä.”

Verkostoissa tehty työ vaikuttaa metriikan kansallisen kehitykseen. Nykyään metriikka kytkeytyy voimakkaasti vastuulliseen tieteeseen ja avoimeen tieteeseen. Isaksson on työllään edistänyt molempia, mutta suhtautuu kriittisesti ”sahnelinään”.

”Saatetaan puhua vastuullisuudesta, ja nopeasti luetella, että tehdään näin ja näin, jolloin jo aistii, että ollaan tulosvetoisia. Ajatellaan myös, että kun yhden sijasta käytetään viittä indikaattoria,

se on sitä vastuullisuutta. Itse näkisin metriikan enemmän avoimen tieteen puolella, johon kuuluu avointen julkaisujen lisäksi avoimet abstraktit, avoimet viittaukset ja avoin data. Jos käytettävissä on mahdollisimman paljon avointa tietoa julkaisuista, voidaan tehdä osuvampaa metriikkaa. Meiltä on vuosien varrella kysytty, voiko tällaisen ja tällaisen tarkastelun tehdä, ja valitettavan usein olemme joutuneet vastaamaan, että ei voida, koska tietokannasta ei saada sitä tietoa irti. Jos data olisi avoimesti käytössä, tilanne muuttuisi voimakkaasti.”



Kuvio 24

≈ Rolf Nevanlinnan *Suhteellisuusteorian periaatteet* (1963) vei Eva Isakssonin kohti fysiikkaa. ”Nuoren ihmisen mieli räjähti, kun luin, että maailmankaikkeus saattaakin olla autonrenkaan muotoinen.”

Tasa-arvotyötä 1980-luvulta tähän päivään

Keskeinen teema Isakssonille on ollut tasa-arvon ja yhdenvertaisuuden edistäminen yliopistomaailmassa. Siinä onkin edistytty.

”Silloin, kun aloitin opiskelut, kuulovammaisilla opiskelijoilla ei ollut mitään tukirakenteita, esimerkiksi nykyistä tulkkaussysteemiä. Piti vain pärjätä. Mitä tulee tasa-arvoon, pitkään oli niin, että asioista päättivät miesprofessorit, jotka eivät nähneet tutkijantyötä naisten työurana. Naistutkijoiden näkyvyys ja naisten tutkimusuran houkuttelevuus on Suomessakin vähitellen lisääntynyt, ja monet sinnikkäät ja näkyvät tiedenäiset ovat toimineet tämän kehityksen kärkinä. Kuitenkin asenteet ja monet rakenteelliset syyt vaikeuttavat edelleen naisten tieteellistä urapolkua.”

Isakssonin omassa julkaisutuotannossa näkyvät tasa-arvon ja rauhanaatteen teemat muun muassa sellaisissa hiljattain uudelleen avoimesti julkaistuissa teoksissa kuin *Rauhan laboratoriot* (1985) ja *Nainen ja maailmankaikkeus* (1987), joka antoi lopullisen sysäyksen tasa-arvotyöhön.

Isaksson on 2000-luvun aikana toiminut mukana muun muassa yliopiston tasa-arvotoimikunnassa sekä esteettömyystyöryhmässä, ja hän on ollut valmistelemaan tasa-arvo- ja yhdenvertaisuussuunnitelmia.

”Asioihin on kiinnitetty lainsäädännössä huomiota, ja yliopistossa on etsitty sen toteuttamiseen toimenpiteitä.

Aiemmin tehtiin enemmän selvityksiä ja suunnitelmia, nykyään tasa-arvotyö on enemmän konkreettisia tekoja.”

Tasa-arvo- ja yhdenvertaisuustyö sointuu hyvin yhteen myös vastuullisen julkaisumetriikan edistämisen kanssa. Kysymykseen, mitä kaikkien yliopistolaisten tulisi tietää bibliometriikasta, Eva Isaksson vastaa näin:

”Kaikkien pitäisi tietää, että se, missä yliopistossa tutkija on töissä tai missä lehdissä hän julkaisee, ei välttämättä kerro kyseisestä tutkijasta. Se, mitä hän pystyy kertomaan tutkimustavoitteistaan, ja miten hän toteuttaa niitä, kertoo tutkijasta enemmän. Nyt, kun vallalla ovat yliopistorankingit, korostetaan liikaa yliopistokontekstia. Mutta oli ihminen kotoisin mistä maasta tahansa, aloitti hän uransa missä tahansa yliopistossa, oli hän vammainen tai jonkin vähemmistön edustaja – se ei täysin määrittele häntä. Lahjakkaita ihmisiä on kaikkialla. Miten metriikka rajoittaa sitä, mitä ajattelemme ihmisen mahdollisuuksista? Jos etsimme valmiita huippuja, helposti nousevia tähtiä, joilla on hyvät edellytykset päästä nopeasti uran alkuun, menetämmekö jotain? Tätä voisi metriikassa ajatella.”

*Isaksson kertoi monipuolisen uransa vaiheista myös elokuisessa juhluennossa *Maailmankaikkeus ja kirjasto, joka löytyy Unitube-videopalvelusta: bit.ly/evaisakssoninjuhluento2021*.*

STKS kannattaa muutoksia tekijänoikeuslakiin

Suomen tieteellisen kirjastoseuran (STKS) kirjastojuridiikka-ryhmän syksy on ollut kiireinen. Syyskuun lopussa lausuntokierrokselle tuli opetus- ja kulttuuriministeriössä valmisteltu luonnos hallituksen esitykseksi (HE) tekijänoikeuslain muuttamisesta. HE:n muutosehdotuksilla saatetaan kansallisesti voimaan DSM-direktiivi sekä päivitetään muita ajantasaistamista vaativia pykälä. STKS:n lausunnossa kannatettiin ehdotettuja muutoksia tekijänoikeuslakiin, erityisesti rinnakaistallennuspoikkeuksen sekä tekstin- ja tiedonlouhinnan osalta. Lausuntoja muutosehdotuksista annettiin yhteensä 223, ja lopullinen HE on tarkoitus saada eduskunnan käsittelyyn joulukuussa. Muutokset ovat kuitenkin herättäneet paljon vastustusta ja ehdotuksen uudelleenvalmistelua on vaadittu, joten aikataulut voivat muuttua. Kirjastojuridiikka-ryhmä järjestää seminaarin uudistetusta lainsäädännöstä, kunhan muutokset astuvat voimaan – toivottavasti vuoden 2022 aikana. Hallituksen esitykseen ja lausuntoihin voi tutustua Lausuntopalvelu.fi:ssä: www.lausuntopalvelu.fi. (Soile Manninen)

Musiikinluettelointi.fi-sivusto avattiin marraskuussa

Musiikin kirjastoluetteloinnin tietolähteeksi perustettu Musiikinluettelointi.fi-sivusto avattiin marraskuussa. Ajantasaista luettelointitietoa tarjoavan sivuston lähtökohdiana on suomalaisissa kirjastoissa noudatettavat kuvailusäännöt ja kuvailua koskeva ohjeistus. Sivuston aineistot ovat vapaasti kaikkien käytettävissä: ne on lisensoitu tekijänoikeuksista vapaalla CC0-lisenssillä. Ensimmäisenä sisältökokonaisuutena julkaistiin "Yhtenäistettyjen musiikkinimekkeiden ohjeluetelot", joka on tietokanta kirjastoluetteloinnissa käytettävistä säveltäjien teosten yhtenäistetyistä nimekkeistä. Musiikinluettelointi.fi-sivustoa ylläpitää Ylen arkiston suunnittelija Jaska Järvillehto, jolla on kokemusta musiikkikirjastotyöstä eri kirjastosektoreilta. Musiikkikirjastoalan Intervalli-lehdessä (2/2021) Järvillehto kuvaa ensimmäisen tietokannan syntyprosessia. (Musiikinluettelointi.fi)

Raportti: Korona-ajan kokemuksiin kohti uutta kirjastoa

Koronapandemia on vajaan parin vuoden aikana vaikuttanut kaikenlaisiin kirjastoihin ympäri maailmaa, ja vaatinut kirjastoja reagoimaan hyvin nopeasti muuttuvaan tilanteeseen ja kehysorganisaation muuttuviin tarpeisiin. Kirjastoalalla toimiva voittoa tavoittelematon OCLC Research haastatteli 29 kirjaston johtajaa 11 maasta kartoittaakseen heidän kokemuksiaan pandemiasta ja sen jälkeisestä ajasta. Haastatelluista 21 edusti tieteellisiä kirjastoja. Lokakuussa julkaistu New Model Library: Pandemic Effects and Library Directions -raportti on yhteenveto haastatteluiden havainnoista. Raportti käsittelee muun muassa työskentelyn joustavuuteen ja yhteistyömalleihin, virtuaalisiin ja fyysisiin kokoelmiin sekä työhön sitoutumiseen liittyviä kokemuksia, ja visioi samalla myös tulevaa: jokainen osio sisältää suosituksia kirjastoille. New Model Library -raportti on avoimesti luettavissa Ocl.org-sivustolla. (Ocl.org)



Ei tästä mitään tule, jos ei jo
jostain jotain löydy,
ja kohta ei näe, että mitä.

-Sirkka Turkka

n
HELSINGFORS.

FINSKA LITTERATUR-SÄLLSKAPETS TRYCKERI,

I. Sähkövirta ja sen vaikutukset.	Sivu 1
II. Sähköyksiköt.	1879. " 20
III. Sähkön mittaus-koneet.	" 25
IV. Dynamosähkökoneita.	" 37
V. Sähköpottoreita.	" 26
VI. Kaari- ja henkilampuista.	" 72
VII. Valon jakaminen ja jostamisesta.	" 80
Suomen Vakuutus-Tariffi-Yhdistyksen määräykset sähkövalaistuksen ja sähkövirran avulla tapahtuvan työn osaksi.	" 99

Muistiinpanot

- Nilsson, Skandinaavien miehen, Lund 1858.
- Holmgren, Skandinaavien miehen, Stockholm 1867.
- Westerlund, Skandinaavien miehen,
- Wright, Finlands historia,
- Malmberg (Mela), Finlands historia,

Ensimmäinen T

hallitukseen, s. o. poliittisiin puoluetaitel-
moksen) välillä. Näitä riitoja hyväksel-
laiset kutsuivat tyran-
on näitä useissa valtioissa
destaan ja sivistyskulttuurin
Korintissa, Polykrates Samos
Kreikkalaiset eivät kuitenkaan
vaan tyrannit karkotetaan ja
doksi tulee tavallista Ateenan
kansanvaltaisen luonnetta
valtaisia aineita.

*) Seittiäytvi *)jai, *)
nen **) huonosti **) liitutti



LOPISTO
 tyhönä, vakuutus, että tuo
 maan keskeisenä, ty-
 väämpöllisistä kate-
 koo-
 tija on tuomanveistäjä
 tohittaanut. Por-
 tilla pöytäkirjaan työvoiman
 uuten, Helsingin tea-
 H. J. H.
 N mukaan.
 kuvaa.
 2 mk. 50 penniä.
 Vähästä paljo tulee, ei,
 Ripinistä tuli loppu.
 flere utmärkta käl-
 kasom vid undervis-
 NEHALL
 nyttig och behölig,
 tta, skall vara väl-
 turen.
 , 1867.