



STKS 90 v
15.11.2019

TEKEEKÖ TEKÖÄLY TULEVAISUUDEN METRIKKA-ANALYYSIT?

Eva Isaksson
Helsingin yliopiston kirjasto
Tutkimuksen palvelut / Metriikka

HELSINGIN YLIOPISTON KIRJASTO
HELSINGFORS UNIVERSITETS BIBLIOTEK
HELSINKI UNIVERSITY LIBRARY

HuLib palvelukuvaus:
**Julkaisumetriikka
ja tutkimuksen
näkyvyys**

Metriikkaryhmä

Esimies:

Johanna Lahikainen

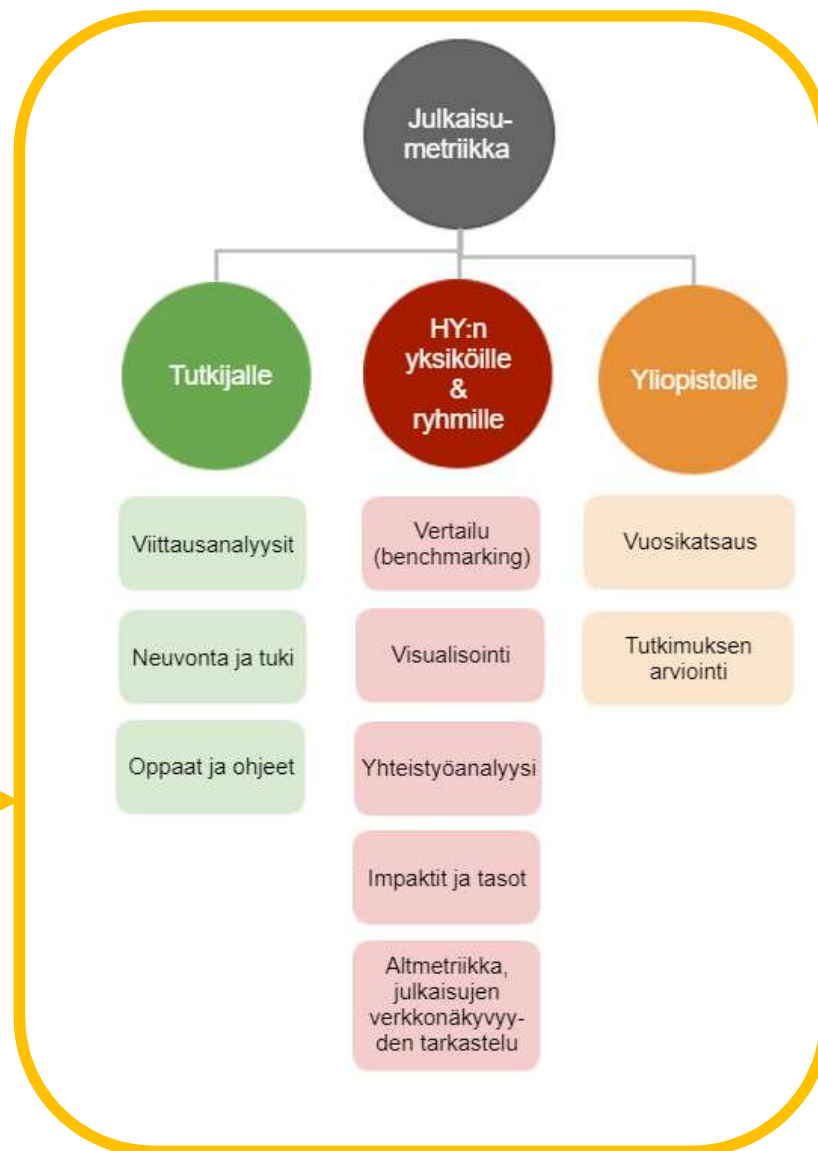
Vetäjä: Eva Isaksson

Osa Tutkimuksen
palveluita

Palvelujohtaja:

Päivi Kaiponen

HELSINGIN YLIOPISTON KIRJASTO
HELSINGFORS UNIVERSITETS BIBLIOTEK
HELSINKI UNIVERSITY LIBRARY



**Näkyvyys-
palvelut**

Altmetriikka

Neuvonta ja tuki

Kudos

PlumX

Altmetric Explorer

Tutkimusportaali
(Tuhat)

Helda
digitaalinen
arkisto

Tunnisteet

Neuvonta ja tuki

ORCID

DOI

**Metriikka-
palveluja
muualla
yliopistolla**

YPA
tutkimushallinnon
metriikka

Rapo - julkaisu-
tiedot

Mitä muita

METRIKKATILASTO 2019

Tilanne 8/2019

Metrikkatilastot

Syöttölomake

Poista tallennuksia

Valitse vuosi: --Valitse-- Näytä

Vuosi 2019

Analyyseja:	12
Käytetty aika:	38 p 3 h
Maksullisia analyyseja:	7 (58 %)
Maksullisiin analyyseihin käytetty aika:	11 p 2 h

Tilaajan rooli

Tiedekunta tai erillislaitos	7	58 %
Tutkimushallinto / yliopistopalvelut	2	17 %
Osasto / yksikkö	1	8 %
Tutkimusryhmä	1	8 %
Yksittäinen tutkija	1	8 %
Yhteensä	12	100 %

Analyysityypit

Yhdistelmä eri analyysejä / laaja selvitys usealla välineellä	6
Muu Lisätiedot	4
Altmetriikka	1
Viittausanalyysi	1

Tiedekunta tai yksikkö - kpl

Humanistinen tiedekunta	3	25 %
Lääketieteellinen tiedekunta	2	17 %
Tutkimuspalvelut / Yliopistohallinto	2	17 %
Bio- ja ympäristötieteellinen tiedekunta	1	8 %
HYKS	1	8 %
Kasvatustieteellinen tiedekunta	1	8 %
Matemaattis-luonnontieteellinen tiedekunta	1	8 %
Valtiotieteellinen tiedekunta	1	8 %
Yhteensä	12	100 %

Tarkoitus

Yksiköiden (tiedekunta yms.) julkaisutoiminnan arviointi ja kehittäminen	7	58 %
Rahoitushaku	2	17 %
Toimihaut	2	17 %
Vuosikatsaus	1	8 %
Yhteensä	12	100 %

Tiedekunta tai yksikkö - tunnit

Tutkimuspalvelut / Yliopistohallinto	180	64 %
Humanistinen tiedekunta	36	13 %
HYKS	25	9 %
Valtiotieteellinen tiedekunta	15	5 %
Lääketieteellinen tiedekunta	13	5 %
Matemaattis-luonnontieteellinen tiedekunta	9	3 %
Kasvatustieteellinen tiedekunta	3	1 %
Bio- ja ympäristötieteellinen tiedekunta	1	0 %
Yhteensä	282	100 %

KIRJASTON METRIKKAPALVELUT OVAT ASiantuntijatyötä

- Kirjaston metriikkaryhmä käyttää kaikkia HY:lle hankittuja välineitä ja tietokantoja
- Lisäksi poimimme tarvittaessa Raposta/TUHATista kohdennetusti
- Neuvomme soveltuvien välineiden valinnassa ja käytössä
- Teemme sopimuksen mukaan erityyppisiä toimeksiantoja
- Laadimme toimeksiannosta analyysiraportin
- Teemme käytettävissä olevan aineiston pohjalta mm. visualisointeja
- Seuraamme metriikan kehitystä
- Olemme yhteydessä muiden suomalaisten korkeakoulujen metriikka-asiantuntijoihin
- Olemme yhteydessä palveluntarjoajiin ja metriikan kansainvälisiin asiantuntijoihin

VÄLINEITÄ, ASIAANTUNTIJOITA, KÄYTTÄJIÄ

Metriikan välineet

- Kalliita
- Väsymättömiä
- "Powerful algorithm",
"State of art"
- Kilpailevat asiakkaista
- Luovat mielikuvia tuotteesta

Metriikka-asiantuntijat

- Kalliita
- Tarvitsevat lomia
- Asiantuntevia, mutta tarvitsevat välillä koulutusta
- Verkostoituvat
- Käyttävät välineitä ja vertailevat

Metriikan loppukäyttäjät

- Kalliita, etenkin dekaanit ja rehtorit
- Kärsivät aikapulasta
- Tarvitsevat koulutusta, eivät välttämättä ehdi
- Tarvitsevat tietoa päätöksentekoon
- Tavoittelevat hyvää tulosta

METRIIKAN ASEMA TUTKIMUKSEN ARVIOINNEISSA ON KASVANUT

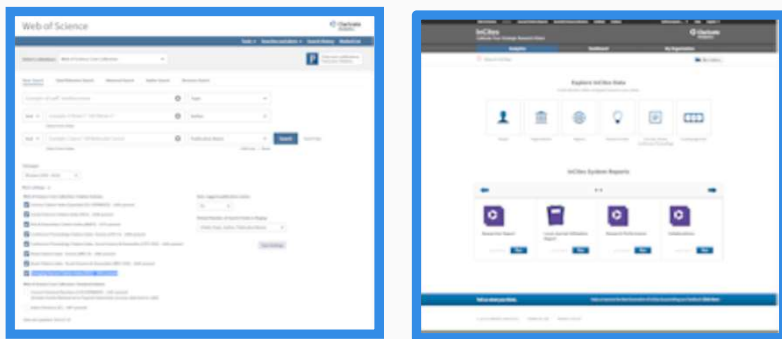
- Metriikka on enenevässä määrin korvannut muita arvioinnin menetelmiä

A. Jappe (ISSI proceedings 2019):

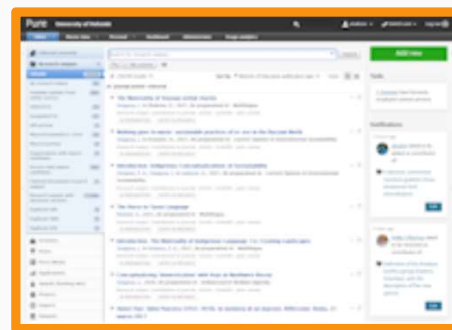
- Eurooppalaiset tutkimuksen arvoinnit 2005-2014: 82 avoimesti saatavaa raporttia
- Bibliometriikan käyttö oli erityisen intensiivistä seuraavissa: Alankomaat, Italia, Ruotsi, Norja, Suomi, Tanska (näiden osuus 60% bibliometrisista arvioinneista)
- Web of Science –pohjainen metriikka hallitsevana
 - 90% kaikista bibliometriikkaa käyttäneistä
 - 15% käytti Scopusta tai WoS+Scopus-yhdistelmää
- Tieteenalakohtaiset impaktit käytetyimmät indikaattorit

METRIIKAN VÄLINEARSENAAALIMME

Web of Science + InCites



Pure

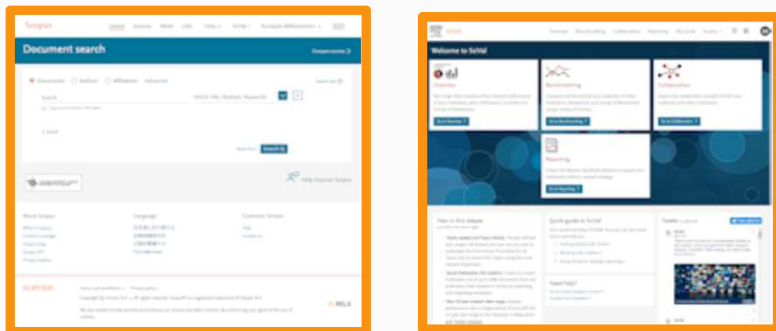


+ Rapo



Scopus

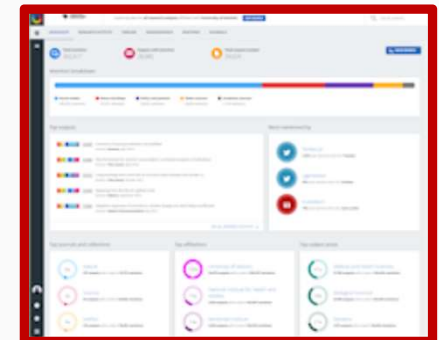
+ SciVal



Google Scholar



Altmetric Explorer



Clarivate



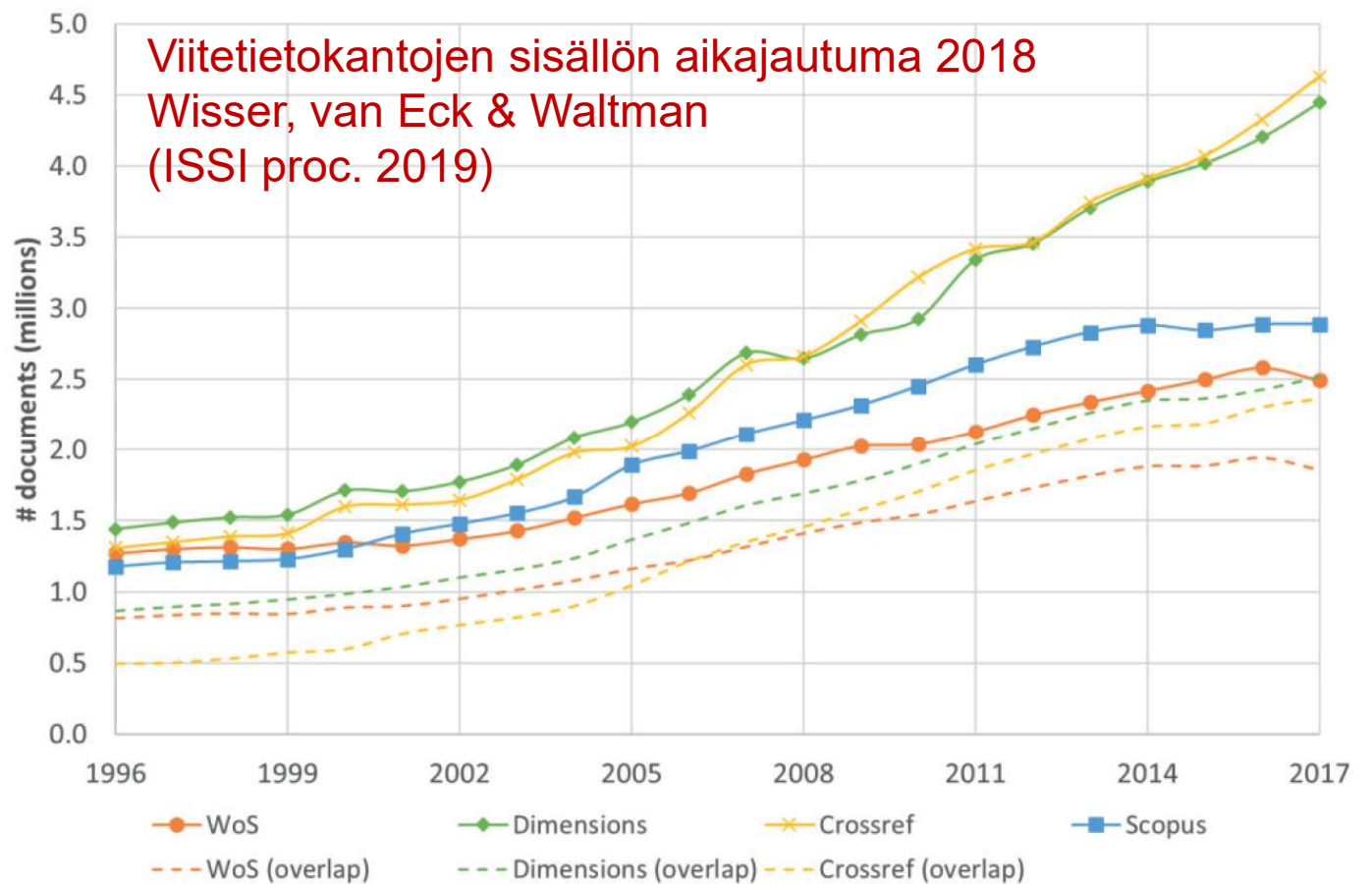
Elsevier



Digital Science

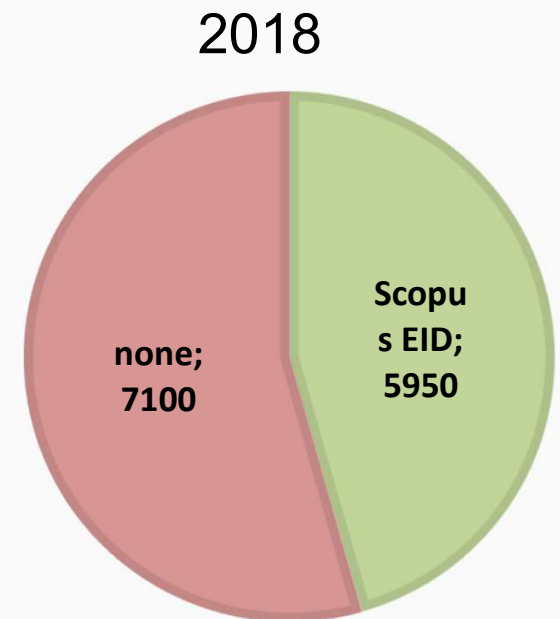
VIITETIETOKANNAT NOTKUVAT JULKAISUJA

- Crossref lisännyt valtavasti DOI-volyymeja
- Paineita aineiston kasvattamiselle viitetietokannoissa
- Dimensions kilpailee volyymillä ja on jo ohittanut Scopus
- Robotit rouskuttavat kertyviä datamassoja



KUN ALGORITMI SAAPUU MÄTSÄÄMÄÄN JULKAISUJA

- Helsingin yliopiston Pure-pohjainen CRIS TUHAT
 - Elsevier osti Puren 2012
- 0,25 miljoonaa julkaisutietuetta
- Julkaisut voidaan tunnistaa mätsäysalgoritmillä, joka liittää niihin Scopuksen EID-tunnisteen
- **Ajatuksena tutkimushallinnon tee-se-itse –metriikka:**
 - Poimitaan yksikön julkaisut TUHATissa
 - Viedään sopivasti aika- ym. rajattu setti SciVal-analytiikkavälineeseen
 - Saadaan heti-valmis bibliometrinen tarkastelu



SCOPUS EID-MÄTSÄYSTÄ KÄYTÄNNÖSSÄ:

TUHAT

For Children Only ; Abridgement of crossover characteristics in the Finnish translation of L.M. Montgomery's Emily trilogy

≠

Scopus

Vaccination is not for children only, and is indispensable for patients with diabetes!

TUHAT

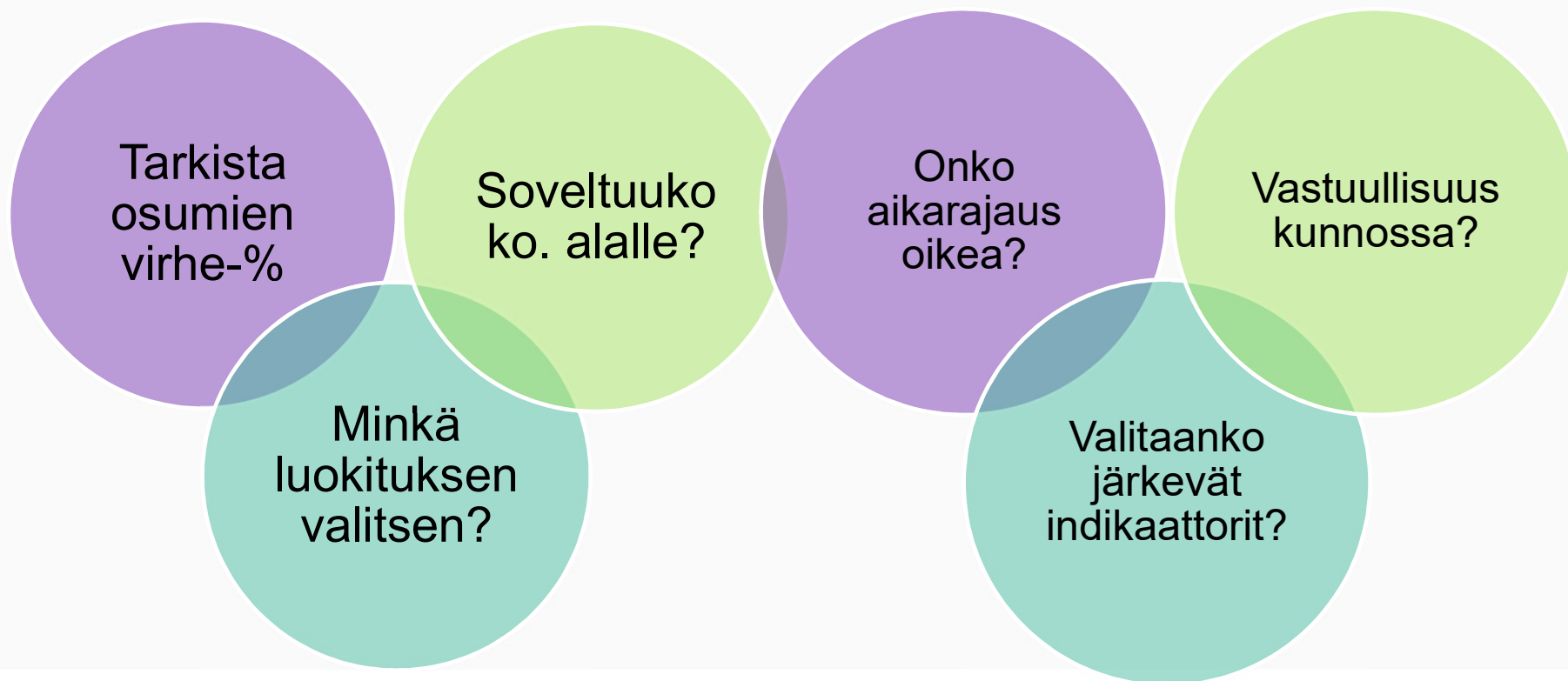
Crisis within a crisis? ; Foreigners in Athens and traces of transnational relations and separations

≠

Scopus

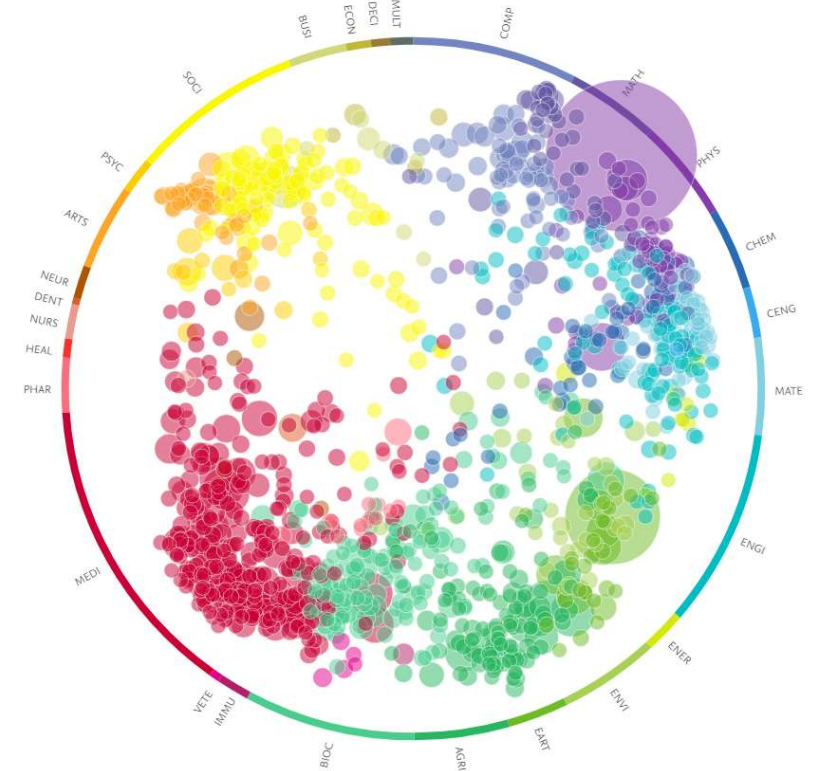
A Crisis Within a Crisis: The Extended Closure of an Opioid Treatment Program After Hurricane Sandy

MITÄ ASIAANTUNTIJA TEKISI?



TIETEENALALUOKITUKSET UUSIKSI?

- Perinteiset lehtisidonnaiset luokitukset eivät enää toimi erityisen hyvin
- Tiede elää, tieteenalojen kytkennät vahvistuvat
- Dimensions-luokittelualgoritmi: syö kokotekstiä, tuottaa FOR-luokiteltuja artikkeleita
 - Lastentauteja: Tähtitieteen lehdissä julkaistaan paljon psykologiaa?
- Scopus/SciVal: Topic Prominences (kuvassa HY)
- CWTS: Leiden Clustering Algorithm
 - HY:n ja Aallon tutkimuksen arvioinnit




```

158 int argIndex = 0;
159 while (argIndex < args.length - 1)
160 {
161     String arg = args[argIndex];
162     try
163     {
164         if (arg.equals("-q") || arg.equals("--quality-function"))
165         {
166             if (((argIndex + 1) >= args.length) || (!args[argIndex + 1].equals("CPM") && !args[argIndex + 1].equals("modularity")))
167                 throw new IllegalArgumentException("Value must be 'CPM' or 'modularity'.");
168             useModularity = args[argIndex + 1].equals("modularity");
169             argIndex += 2;
170         }
171         else if (arg.equals("-r") || arg.equals("--resolution"))
172         {
173             try
174             {
175                 if ((argIndex + 1) >= args.length)
176                     throw new NumberFormatException();
177                 resolution = Double.parseDouble(args[argIndex + 1]);
178                 if (resolution < 0)
179                     throw new NumberFormatException();
180             }
181             catch (NumberFormatException e)
182             {
183                 throw new IllegalArgumentException("Value must be a non-negative number.");
184             }
185             argIndex += 2;
186         }
187         else if (arg.equals("-a") || arg.equals("--algorithm"))
188         {
189             if (((argIndex + 1) >= args.length) || (!args[argIndex + 1].equals("Leiden") && !args[argIndex + 1].equals("Louvain")))
190                 throw new IllegalArgumentException("Value must be 'Leiden' or 'Louvain'.");
191             useLouvain = args[argIndex + 1].equals("Louvain");

```

Klusterointialgoritmin komentorivikoodia

HY:N METRIKKATIIMISSÄ TESTATTU CWTS:N MENETELMÄÄ

Assessing citation network clustering as indicator normalization tool

Riku Hakulinen¹ and Eva Isaksson²

¹ Riku.Hakulinen@helsinki.fi, ² Eva.Isaksson@helsinki.fi
University of Helsinki, Helsinki University Library, Research Services, PO Box 53,
00014 Helsinki (Finland)

Introduction

Attempts to use bibliometrics in assessing research performance requires a normalization procedure to cover publications from different fields. The classification behind this normalization has traditionally been broad categories based on scientific fields of journals. During recent years, due to the need for a more specific method, citation network clustering has been utilized to partition research output into (micro) research fields. Also, machine learning approaches with the aim of classifying publications based on their full text have

network is used as input in the clustering tool *RunNetworkClustering*, provided by CWTS through GitHub: [CWTSLeiden/networkanalysis](https://github.com/CWTSLeiden/networkanalysis). Further work on the data and clustering results is done using the statistical software R and the networks are visualized with VOSviewer.

Preliminary Tests

Algorithms

At least three clustering methods are tested on our data: Leiden algorithm with Modularity and

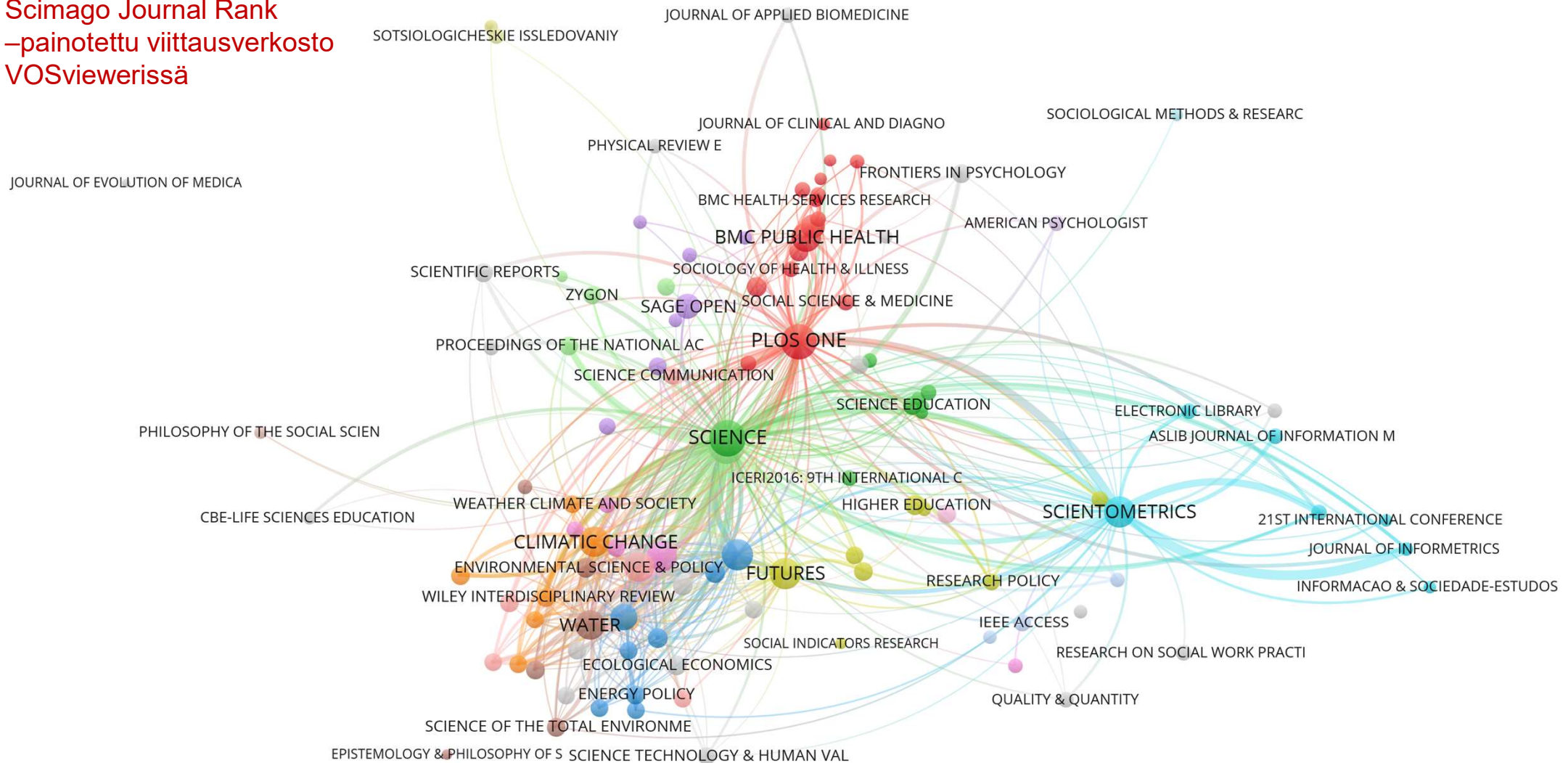
- Testattu CWTS:n algoritmia omiin datasetteihin
- Tuloksia testeistämme esitimme ISSIn konferenssissa syyskuun alussa (poster)

17th INTERNATIONAL CONFERENCE
ON SCIENTOMETRICS & INFORMETRICS

September 2-5, 2019
Sapienza University, ROME | Italy



Testisetin (112 julkaisusarjaa)
 Scimago Journal Rank
 –painotettu viittausverkosto
 VOSviewerissä



YHTEENVETONA

- Tietokantojen kasvu kiihtyy
- Tunnisteiden lisääntynyt käyttö
- Kokotekstien parantunut saatavuus
- Tehokkaammat palvelimet
- Algoritmien kehitystyö

- Metriikka voimakkaassa kehitysvaiheessa
- Tulevaisuudessa paljon enemmän kuin indikaattoreita tuottava musta laatikko
- Uusia algoritmeja tuotetaan ja testataan
- Uusia löytöretkiä tekstiavaruuteen

”Computer generated hypotheses must still be tested. And that requires human researchers.”

Extance A.:
How AI technology can tame the scientific literature.
Nature. 2018 Sep;
561(7722):273-274. doi:
[10.1038/d41586-018-06617-5](https://doi.org/10.1038/d41586-018-06617-5)

<i>Waltman, Ludo</i>	561, 1288, 1301, 1339, 2121, 2358, 2370	<i>Wu, Jianben</i>	1662, 2507
<i>Wang, Chang</i>	782	<i>Wu, Jiang</i>	2688, 2736
<i>Wang, Chaoqun</i>	2782	<i>Wu, Jinshan</i>	2598
<i>Wang, Chun-Chieh</i>	2332	<i>Wu, Xia</i>	363
<i>Wang, Dongbo</i>	2499, 2501, 2710, 2712, 2732, 2740, 2742	<i>Wu, Xiaoyan</i>	2215
<i>Wang, Feifei</i>	2528	<i>Wu, Yishan</i>	1138
<i>Wang, Haiyan</i>	1128, 1859	<i>Xavier, Paula</i>	2680
<i>Wang, Hongyu</i>	2752	<i>Xiang, Bin</i>	2744
<i>Wang, Hui</i>	2620	<i>Xiao, Shaoyun</i>	2748
<i>Wang, Li</i>	2479	<i>Xiao, Tingting</i>	477
<i>Wang, Lili</i>	202, 1813, 2215, 2489	<i>Xiaoguang, Wang</i>	2752
<i>Wang, Panting</i>	2708	<i>Xie, Juan</i>	2760
<i>Wang, Peiling</i>	387	<i>Xie, Li</i>	2698
<i>Wang, Qian</i>	2742	<i>Xiong, Wenjing</i>	1537
<i>Wang, Qianfei</i>	2526	<i>Xu, Haiyun</i>	2518, 2526, 2744
<i>Wang, Shenghui</i>	1038	<i>Xu, Jing</i>	399
<i>Wang, Xianwen</i>	1156, 1166	<i>Xu, Li</i>	2628
<i>Wang, Xiaomei</i>	990, 1570	<i>Xu, Ping</i>	2628
<i>Wang, Xing</i>	2594	<i>Xu, Shuo</i>	2756
<i>Wang, Xu</i>	2477	<i>Xu, Xiaoke</i>	1166
<i>Wang, Yue</i>	2602, 2628	<i>Xu, Xinyue</i>	1044
<i>Wang, Yuefen</i>	596, 1677, 2632, 2718	<i>Xuan, Zhaohui</i>	573
<i>Wang, Yueqian</i>	1652	<i>Yajie, Wang</i>	1178
<i>Wang, Yuzhuo</i>	2702, 2726	<i>Yakimovitch, Nikita</i>	631
<i>Wang, Zhiqi</i>	459	<i>Yamashita, Yasuhiro</i>	2145
<i>Weber, Matthias</i>	1146	<i>Yan, Duanwu</i>	2602
<i>Wei, Fangfang</i>	1662, 2507	<i>Yan, Erjia</i>	87
		<i>Yan, Xiaoran</i>	2442
		<i>Yang, Liying</i>	1358, 2493, 2598

KUKA TEKEE TULEVAISUUDEN METRIIKKA- ANALYYSIT?

Ote ISSI-konferenssin 2019
proceedingsin tekijäindeksistä



KIITOS!

#HULIB

www.helsinki.fi/kirjasto

Eva Isaksson

eva.isaksson@helsinki.fi