Towards reliable data – counting the Finnish Open Access publications

CRIS 2016, St Andrews, June 10, 2016

Jyrki Ilva jyrki.ilva@helsinki.fi

Permanent address: http://urn.fi/URN:NBN:fi-fe2016061020120

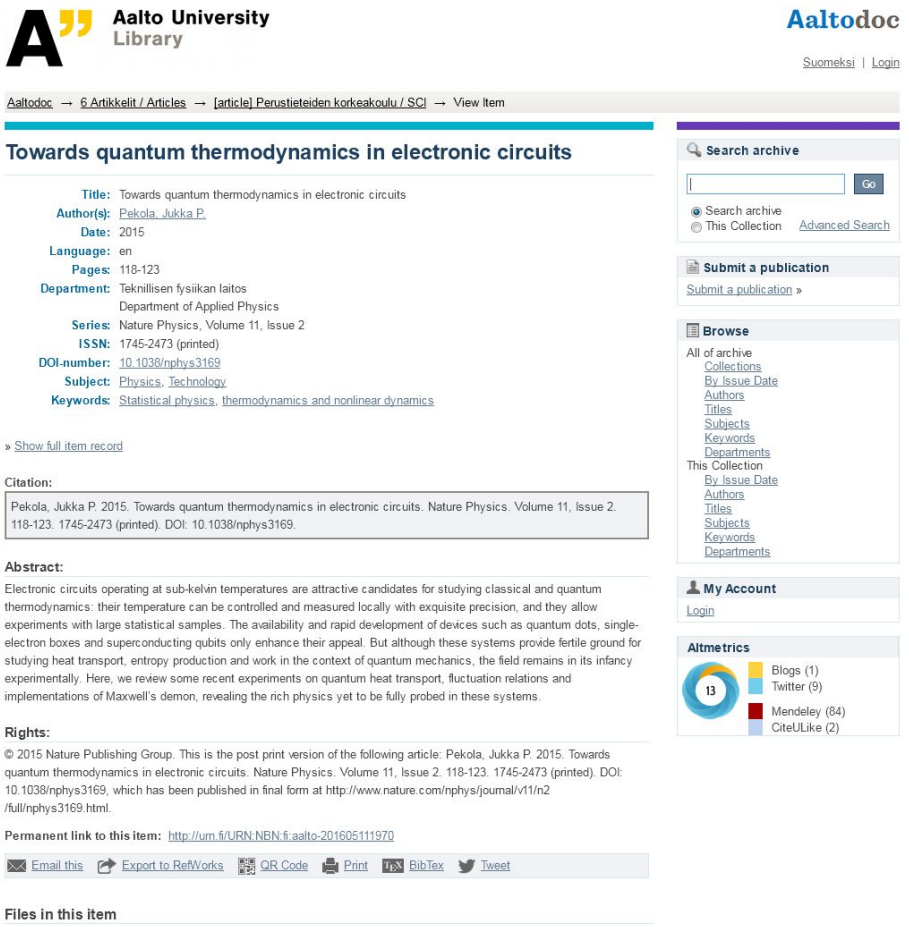
Open Access as a political goal

- Open access is moving forward on a European level
- In Finland the Ministry of Education and Culture has funded Open Science and Research, a big national project
  - Open Access, Open Research Data and Open Research Methods

OPEN SCIENCE AND RESEARCH

Supporting Open Access in Finland

- Quite a lot has been done already
  - Finland has a fairly well-developed network of institutional repositories
  - Many of the universities have self-archiving mandates
  - The Academy of Finland launched an Open Access mandate in 2015

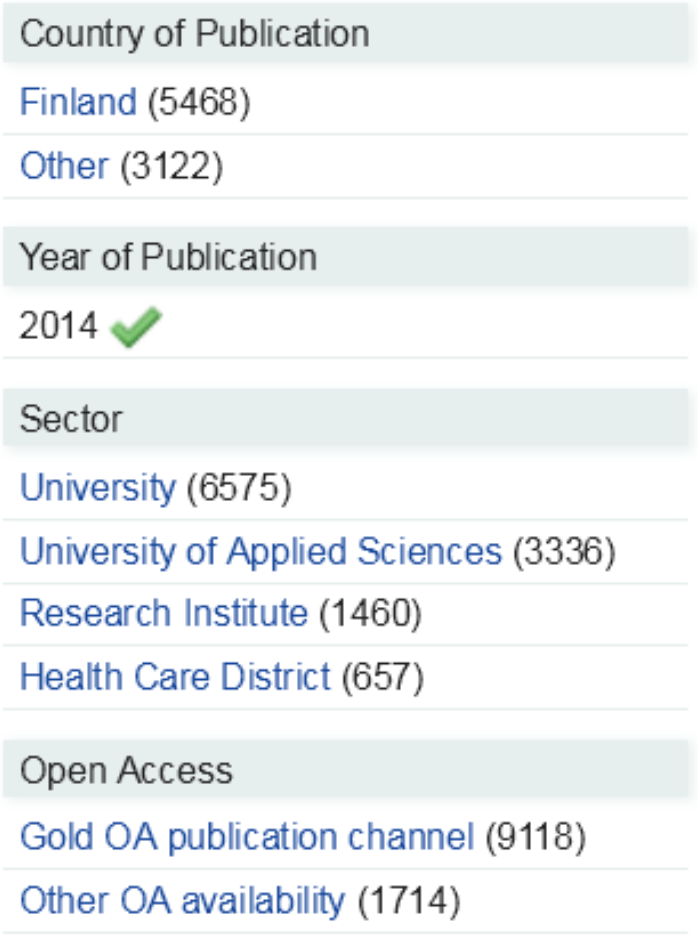


Open Access and state funding?

- Would it be possible to use the Open Access availability of research output as a factor in the funding model for the universities?
  - Currently 13 % of all state funding for the Finnish universities is distributed based on the number and quality of the publications
  - The committee responsible for the university funding model has been discussing this, but no final decision has been made
  - Possibly starting from 2019, as the funding is based on the publication data from the previous three years

The current amount of OA in Finland?

- According to the publication data collected by the Ministry of Education and Culture, the Finnish research organizations produced more than 40.000 publications in 2014
  - About 24.000 of all reported publications were peer-reviewed
- How many Open Access publications?
  - The 2014 data would indicate that about 26% of all reported publications were OA
  - 18% of all peer-reviewed publications were OA
  - 19% of the peer-reviewed journal articles were OA



But can we actually trust these numbers?

- Unfortunately, a closer look reveals that there are many problems with the data, some of them quite serious
  - The definitions that have been used were not clear enough
  - As long as the OA status of a publication doesn't have any financial effect for the organizations, the motivation for collecting and verifying the information remains low
- At the moment the OA data is not reliable enough be used as a basis for distributing funding to the universities
- The Ministry of Education and Culture has already taken steps to rectify the situation, starting from 2016

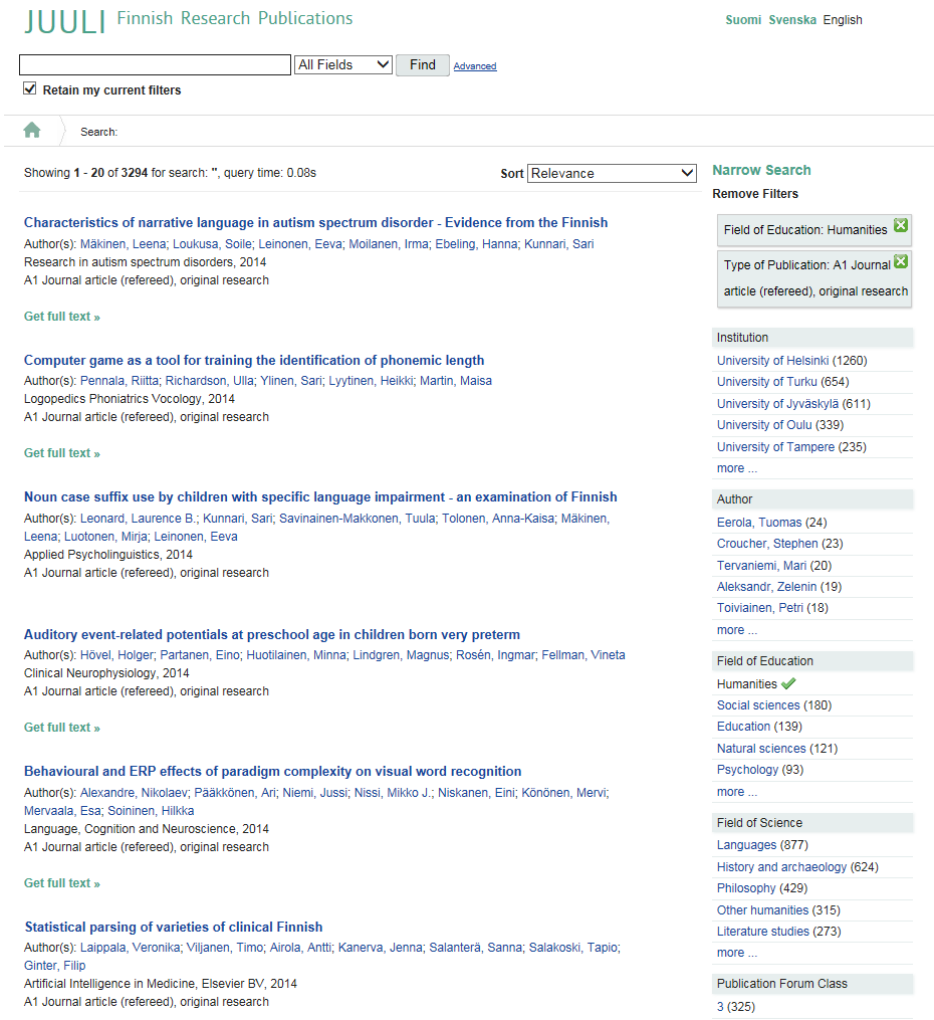
Collecting the publication data (1)

- The Ministry of Education and Culture has been collecting publication data from the universities since 2011
  - The universities of applied sciences joined the data collection in 2012
  - Data collection from the state research institutes and central hospital districts started in 2015
- CSC - IT Center for Science is responsible for the data collection
  - Close collaboration with the research organizations, the Ministry and the Federation of Learned Societies

Type of Publication
A1 Journal article (refereed), original research (56489)
A2 Review article, Literature review, Systematic review (3213)
A3 Book section, Chapters in research books (12464)
A4 Conference proceedings (17418)
B1 Non-refereed journal articles (9698)
B2 Book section (6535)
B3 Non-refereed conference proceedings (4191)
C1 Book (1640)
C2 Book (editor), chapters in research books, conference proceedings or special issue of a journal (2626)
D1 Article in a trade journal (9922)
D2 Article in a professional book (incl. an introduction by the editor) (6160)
D3 Professional conference proceedings (661)
D4 Published development or research report or study (5427)
D5 Textbook, professional manual or guide (1299)
D6 Edited professional publication (1)
E1 Popularised article, newspaper article (13015)

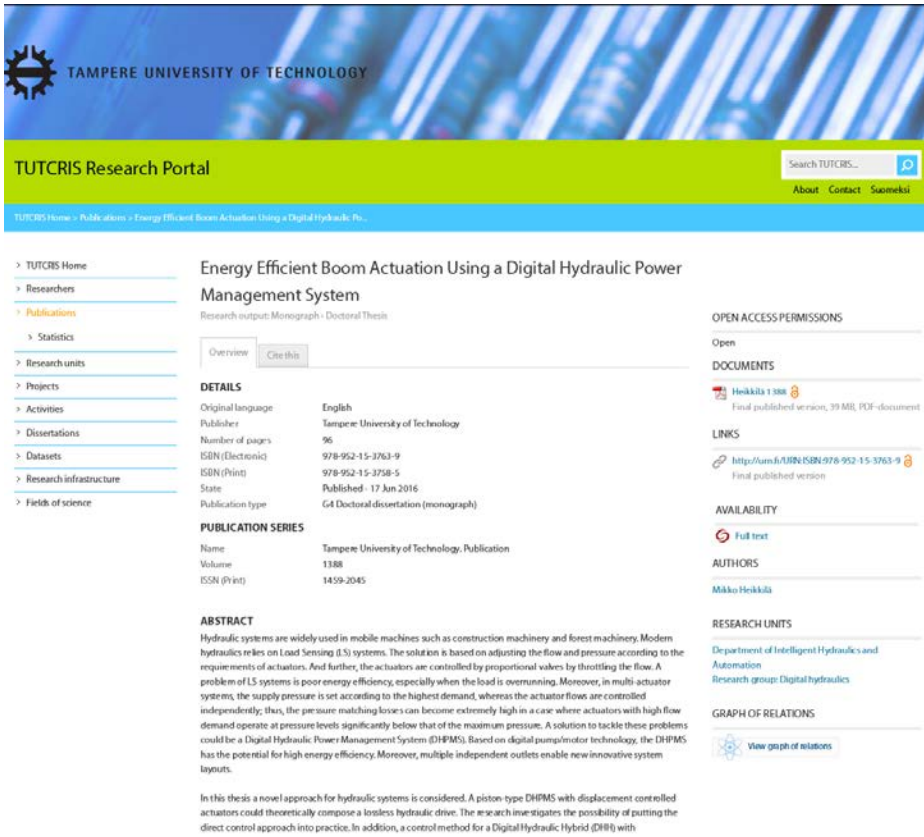
Collecting the publication data (2)

- The combined data is available for viewing at Juuli (www.juuli.fi), a national research publications portal developed by the National Library of Finland
- The processes of data collection are currently under development
- Starting from 2016, the new Virta publication information system will harvest data from local CRISes and provide the de-duplicated data via APIs



National data collection and the local CRIS

- At the university level the growing importance of national data collection has led to a widespread adoption of modern CRIS platforms
  - Most of the recently implemented systems have been based on Pure or Converis
- However, many of the smaller institutions are working with less sophisticated legacy systems



How to verify the OA status of a publication?

- One method is to check the status of each publication individually
  - May require quite a bit of resources
- The other method is to utilize outside data sources in an automated identification process
  - For example, it is possible to identify most of Gold OA journal articles by checking if the journal is listed in Directory of Open Access Journals (DOAJ)

- In many cases a combination of these both approaches is likely to produce the best results

Data sources for the Open Access status?

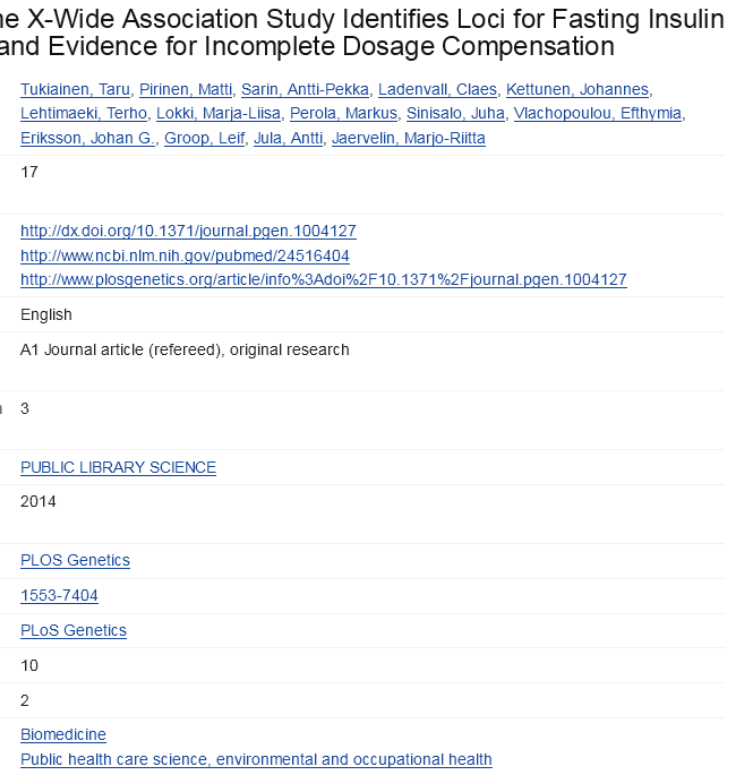
- DOAJ covers most of the OA journals actually used by researchers
  - However, currently there is no reliable method for articles published in Hybrid OA journals or articles that have appeared as book chapters or in conference proceedings
- It is relatively easy to check if the publication is available in the local institutional repository
  - If the publication is deposited into a non-local repository, there might not be an easy way to get notification of this
- International initiatives like OpenAire, SHARE or CHORUS may prove to be useful
  - There is still a lot of work ahead before they are comprehensive or reliable enough

OA status in the national data collection

- The fields used for indicating the OA status of the publication were updated starting from 2016
  - The aim was to keep the number of required fields small - the collection of new information costs money on the local level
- It was decided that there would be two separate fields for indicating the OA status of a publication
  - One field indicates whether the publication is OA via a Gold or a Hybrid OA channel
  - The second field is used for indicating whether it has been deposited into a repository
- There is quite a bit of overlap between Gold and Green OA
  - Many of the articles published in OA journals are also uploaded into a repository

URLs for the open access versions

- There should be a URL for each OA version of the publication so that the OA status can be verified
  - In the past there was only one field for this: the organizations usually chose to report a DOI leading to publisher's web site
  - There were also cases in which there was no URL provided for publications that were claimed to be OA



Chromosome X-Wide Association Study Identifies Loci for Fasting Insulin and Height and Evidence for Incomplete Dosage Compensation

Author(s):	Tukiainen, Taru, Pirinen, Matti, Sarin, Antti-Pekka, Ladenvall, Claes, Kettunen, Johannes, Lehtimaeki, Terho, Lokki, Marja-Liisa, Perola, Markus, Sinisalo, Juha, Vlachopoulou, Efthymia, Eriksson, Johan G., Groop, Leif, Jula, Antti, Jaervelin, Marjo-Riitta
Number of Contributors:	17
Online Access:	http://dx.doi.org/10.1371/journal.pgen.1004127 http://www.ncbi.nlm.nih.gov/pubmed/24516404 http://www.plosgenetics.org/article/info%3Adoi%2F10.1371%2Fjournal.pgen.1004127
Language:	English
Type of Publication:	A1 Journal article (refereed), original research
Publication Forum Class:	3
Publisher:	PUBLIC LIBRARY SCIENCE
Year of Publication:	2014
Journal / Series:	PLOS Genetics
ISSN:	1553-7404
Host publication:	PLoS Genetics
Volume	10
Issue	2
Field of Science:	Biomedicine Public health care science, environmental and occupational health

What qualifies as Open Access?

- It would be very nice to have a short and easy-to-understand definition for Open Access
- However, Open Access is actually a relatively complicated issue, with many different flavors

Gold OA?

Green OA?

Hybrid OA?

Gratis OA?

Libre OA?

Embargoes?

Versions?

OA definition for the data collection?

- The goal was to come up with a fairly strict definition of OA, which would still be permissive enough to be acceptable for all fields of research

OA definition for the data collection 1/4

"The publication can be read online in full and without restriction, printed out and copied at least for research use"

- The publication should be at least free to read ("gratis OA")
- Some of the Open Access advocates stress that all Open Access publications should be published under a CC BY license ("libre OA") so that their content can be re-used e.g. in text-mining
- However, there is no universal agreement on this

OA definition for the data collection 2/4

"The publication is available either a) directly from the publisher's service or b) no later than after the end of an embargo period set by the publisher, through archiving in a repository dedicated to a specific organization or field of science"

- Gold/Hybrid and Green Open Access are all accepted
- Institutional and subject-based repositories are both OK
- Embargoes are allowed for Green OA but not for Gold or Hybrid OA

OA definition for the data collection 3/4

"The publication is freely available from a service provided by either a publisher or research organization that 1) enables the harvesting of the metadata of publications and the indexing of them to other search services, and which 2) supports citations and links to publications preferably using URLs based on persistent identifiers (DOI, URN, Handle)"

- Papers posted to a personal or project web site don't qualify
- Although commercial networking services like Research Gate and Academia.edu are very popular among researchers, uploading a paper into them does not qualify as proper Open Access
- Persistent identifiers should be used whenever possible

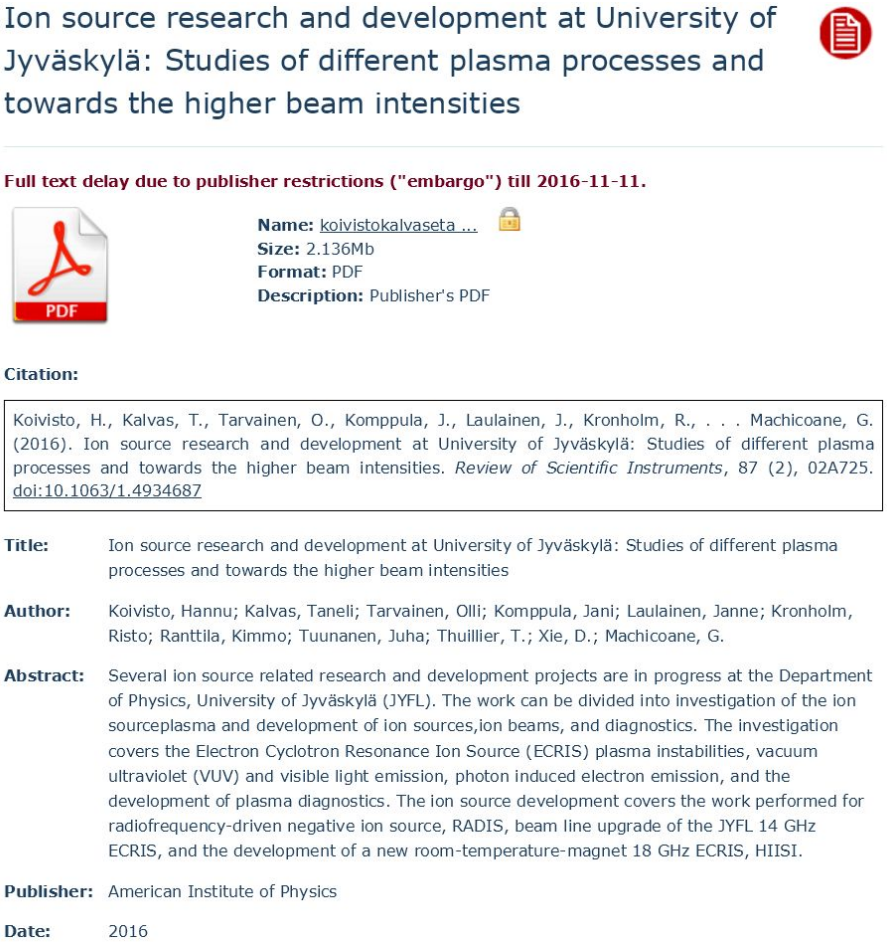
OA definition for the data collection 4/4

"The freely available version of the publication is, depending on the publication contract or publisher's policy, either the author's last self-archived (peer-reviewed) version or the final version published via the publisher's service"

- The OA version must be peer-reviewed to qualify
- OA pre-prints don't qualify – which means that in the case of e.g. ArXiV, the researchers must deposit a post-print version as soon as it becomes available to them

Further policy issues

- There are number issues which may require further discussion
- Should there be a maximum embargo – or no embargo at all - allowed for Green OA?
- Should all publications be deposited into a repository?
- Should there be a more strict policy on licensing (gratis OA vs. libre OA)



Ion source research and development at University of Jyväskylä: Studies of different plasma processes and towards the higher beam intensities

Full text delay due to publisher restrictions ("embargo") till 2016-11-11.

PDF

Name: koivistokalvaseta ...
Size: 2.136Mb
Format: PDF
Description: Publisher's PDF

Citation:

Koivisto, H., Kalvas, T., Tarvainen, O., Komppula, J., Laulainen, J., Kronholm, R., . . . Machicoane, G. (2016). Ion source research and development at University of Jyväskylä: Studies of different plasma processes and towards the higher beam intensities. *Review of Scientific Instruments*, 87 (2), 02A725. doi:10.1063/1.4934687

Title: Ion source research and development at University of Jyväskylä: Studies of different plasma processes and towards the higher beam intensities

Author: Koivisto, Hannu; Kalvas, Taneli; Tarvainen, Olli; Komppula, Jani; Laulainen, Janne; Kronholm, Risto; Ranttila, Kimmo; Tuunanen, Juha; Thuillier, T.; Xie, D.; Machicoane, G.

Abstract: Several ion source related research and development projects are in progress at the Department of Physics, University of Jyväskylä (JYFL). The work can be divided into investigation of the ion sourceplasma and development of ion sources,ion beams, and diagnostics. The investigation covers the Electron Cyclotron Resonance Ion Source (ECRIS) plasma instabilities, vacuum ultraviolet (VUV) and visible light emission, photon induced electron emission, and the development of plasma diagnostics. The ion source development covers the work performed for radiofrequency-driven negative ion source, RADIS, beam line upgrade of the JYFL 14 GHz ECRIS, and the development of a new room-temperature-magnet 18 GHz ECRIS, HIISI.

Publisher: American Institute of Physics

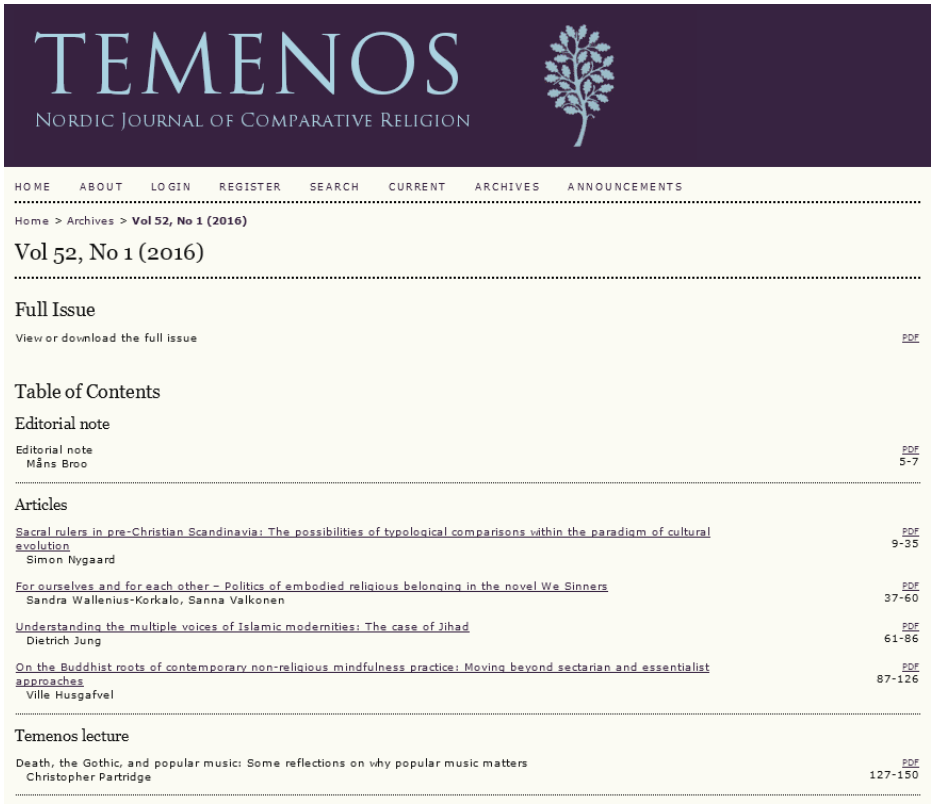
Date: 2016

Towards a seamless flow of information

- In an ideal world the OA status information should be transferred automatically from the publishers' services and repositories to other services
  - Linked open data, integration via APIs
  - Would benefit both the local, national and international OA monitoring efforts
- We can work on this on a national level, but there is a strong need for international co-operation as well

Working with the domestic publishers

- A project funded by the Ministry of Education and Culture is providing upgrades to an Open Journal Systems based centralized infrastructure
- Adoption of DOIs, ORCIDs, providing metadata for national and international aggregators



TEMENOS

NORDIC JOURNAL OF COMPARATIVE RELIGION

HOME ABOUT LOGIN REGISTER SEARCH CURRENT ARCHIVES ANNOUNCEMENTS

Home > Archives > **Vol 52, No 1 (2016)**

Vol 52, No 1 (2016)

Full Issue

View or download the full issue PDF

Table of Contents

Editorial note

Editorial note PDF 5-7
Måns Broo

Articles

Sacral rulers in pre-Christian Scandinavia: The possibilities of typological comparisons within the paradigm of cultural evolution PDF 9-35
Simon Nygaard

For ourselves and for each other – Politics of embodied religious belonging in the novel We Sinners PDF 37-60
Sandra Wallenius-Korkalo, Sanna Valkonen

Understanding the multiple voices of Islamic modernities: The case of Jihad PDF 61-86
Dietrich Jung

On the Buddhist roots of contemporary non-religious mindfulness practice: Moving beyond sectarian and essentialist approaches PDF 87-126
Ville Husgafvel

Temenos lecture

Death, the Gothic, and popular music: Some reflections on why popular music matters PDF 127-150
Christopher Partridge

Keeping track of the OA costs

- A national survey on the management of Article Processing Charges (APCs) at Finnish universities conducted in 2015
  - Result: we can make educated guesses on how much we pay, but we don’t really know
  - There are countries like the UK which are much further in this
- The monitoring of APCs has been out of scope for CRISes
  - Would it make more sense to integrate all of this information into one system or build separate systems which share information between each other?
  - Will this information be a part of the national data collection at some point?

Offsetting deals and the cost of OA

- OA monitoring information also essential for the negotiation of large-scale offsetting deals with publishers
  - Licensing costs paid from the library budgets: not part of the research management processes served by CRIS
  - However, if we move away from the subscription model, there will be a growing need to look at the total costs related to the publication of research outputs

Thank you!