The role of CRIS systems in measuring Open Access publications

Andrea Schamhorst, Ely Dijk, Peter Doorn and Mamin van Berchum
Data Archiving and Networked Services – DANS, Royal Netherlands Academy of Arts and Sciences, The Hague, The Netherlands

Abstract

• Open Access to scientific results – from publications to data – is high on the agenda of science policy drivers. Discussing there is needed that making your research available online is the key to gain visibility and impact (Lawrence 2001). The Open Access movement started in 1991, with calls for self-archiving (Green OA), and led to new technological solutions for scholarly communication. Most recently, the Dutch Green OA mandate (2019) and the European Gold OA model (2002) set the stage for research information, and commercial publishers (such as Elsevier and CONVERIS) have in principle agreed to make their systems CERIF compliant.
• The paper discusses different definitions of Open Access as implemented in technical systems, and the consequences of applying those different measures to determine the percentage of OA publications of a country. We compare the results of applying different definitions of Open Access publications based on the output side (Gold OA, different Gold OA registries) with the input side (research communication) and protocols (CERIF, OAI_PMH).
• We analyse the sources of ambiguity: what type of OA? Only publications in OA registered journals (DOAJ registry) or also those agreeing to Green OA self-archiving (Sherpa Romeo registry)? What type of OA mandate? Only OA registered journals (DOAJ registry) or also self-archiving? Where to harvest (CERIF) and which repositories (DOAJ, Sherpa)?

Stakeholders - OA debate

Definition of OA

Open Access (OA) means free online access to peer-reviewed research journals and articles.
• Green OA = author self-archiving (as preprint, or as a supplement in an institutional repository)
• Gold OA = to publish in a journal that makes its own articles OA (with or without Article Processing Charge)

Measuring OA

• Sources of ambiguity:
  • What type of publications? (all, peer reviewed articles, …)
  • What type of OA? Only publications in OA registered journals (DOAJ registry) or also those agreeing to Green OA self-archiving (Sherpa Romeo registry)?
  • In which RIS DB of publisher, international DB, local institutional RIS. Each might apply another internal counting method.

Visualizing OA in NARCIS

In May 2005, 16 major Dutch universities cooperatively launched DAREnet, the Digital Academic Repositories, making over 47,000 research papers available to anyone with internet access. From 1 January 2007, all the completion of the DARE programme, NARCIS Research Information has taken over responsibility for the DAREnet portal. On 2 June 2008, DAREnet has been incorporated into the scholarly portal NARCIS. In 2016 NARCIS provides access to over 438,000 open access publications from all Dutch universities, KNAW, NWO and a number of scientific institutes.

NARCIS – Research Information System

RIS collect and curate information about scientific knowledge production from the input side (staff, expenditures, projects) to the output side (publications, PhD students, patents). We see a trend to all encompassing RIS and to automatic coupling different RIS based on standards and protocols (CERIF, OAI_PMH). How to monitor?

NARCIS allows comparison and combination with other metadata.

Conclusions

• Different disciplines have different scholarly communication cultures, including sharing information and OA.
• Computer-based systems for Research Information have different ways OA can be tagged to output. The fact that they are often user-content driven can increase ambiguity.
• Interoperability of RIS, and standards as CERIF are the way of the future to gather and analyse data.
• Central harvesters – such as NARCIS – allow to visualize OA in Context. These visualisation also inform about the quality of OA measurement.

References