

# What can institutional repository download statistics tell us?

Harry R. Sidhunata<sup>1</sup>, Tanya Holm<sup>1</sup>, Maude Frances<sup>1</sup>

<sup>1</sup>The University of New South Wales, Sydney, Australia, [h.sidhunata@unsw.edu.au](mailto:h.sidhunata@unsw.edu.au)

<sup>1</sup>The University of New South Wales, Sydney, Australia, [t.holm@unsw.edu.au](mailto:t.holm@unsw.edu.au)

<sup>1</sup> The University of New South Wales, Sydney, Australia, [m.frances@unsw.edu.au](mailto:m.frances@unsw.edu.au)

## POSTER SUMMARY

This poster focuses on download statistics from UNSW institutional repository, UNSWorks [1]. UNSW Library has been collecting download statistics for many years, but recently we have taken a deeper dive into the statistics. Using download statistics for the period January to June 2019, we attempt to answer: Who is downloading content? What is being downloaded? And where is the content accessed from? The analysis of the answers to these questions may provide insights into the diversity of users, content and source as it relates to downloads.

## WHO IS DOWNLOADING CONTENT?

Users of a repository do not sign in and they predominantly access publication files without even visiting the repository. Therefore, an analysis of the IP address of the remote host accessing the record or file is required to understand who our users are. It can tell us whether the user is a UNSW staff member or student, the country/region of the user, and whether the user is accessing the file via another organisation, e.g. another university, a government department or a business.

## WHAT IS BEING DOWNLOADED?

Our analysis of downloaded content is based on resource type and author affiliation. At UNSW almost three quarters of downloads from our institutional repository are of theses. This is not surprising, since theses are generally not published elsewhere. The analysis provides insights about possibilities for increasing usage of other resources in the repository – especially in the light of their open access value.

Analysis of downloads by faculty demonstrates that at UNSW the *Faculty of Arts & Social Sciences* not only has the most highly-downloaded theses, but is also among the faculties with the most downloads. This could indicate that a repository is a “popular” discovery point for content in the arts and social sciences, possibly because some content with this subject matter is not readily found elsewhere or because stakeholders in these disciplines may not have access to conventional scholarly communication sources. This provides opportunities for engagement with researchers in this field, to promote the value of institutional repositories for exposure to diverse audiences and research impact.

## WHERE IS CONTENT ACCESSED FROM?

Our analysis shows that users find and download publication files predominantly from outside the repository – and they do this mostly via Google. A considerable amount of referrals are also from Trove (National Library of Australia) [2] and Open Access Theses and Dissertations (OATD) [3]. There is a correlation between highly-downloaded theses and Google Scholar, with almost two-thirds of highly-downloaded theses indexed in Google Scholar. The majority of downloads are direct file downloads - bypassing the repository landing pages. Research partners, stakeholders and the public may not be aware of the existence of our repositories. It is therefore important that we focus our attention on dissemination and discovery of repository content from external sources rather than focusing on the look and feel of the repository.

## REFERENCES

1. UNSWorks, The University of New South Wales institutional repository. <https://www.unsworks.unsw.edu.au>
2. Trove, an online service created by the National Library of Australia. <https://trove.nla.gov.au/>
3. Open Access Theses and Dissertations (OATD), a resource for finding open access graduate theses and dissertations published around the world. <https://oatd.org/>