Policy citation databases offer new ways to understand the impact of social sciences research

Tracking the policy impact of academic research is notoriously difficult, especially when academics are not directly involved in policymaking processes. However, the recent development of tools to index and organise online policy documents has cast new insights into how research, particularly in the social sciences is used in policy. Commenting on a recent study carried out using Overton.io, **Martin Szomszor***, highlights how social science research outperforms other fields of research in terms of policy impact and how by tracing policy impacts it is possible to produce more nuanced pictures of research impact outside of the dominant anglophone research landscape.*

Citation data have been a staple analytical tool in the physical and life sciences disciplines for more than 50 years and are utilised across the research lifecycle by various stakeholders. Although citation databases were originally developed to support researchers with search and discovery use cases, the development of citation indicators and their subsequent use in performance benchmarking have led to more widespread use of citation data in research evaluation.

However, this success has been largely focused on a subset of academic research where citations are typically concentrated in journals. This serves the hard sciences well, but within the social sciences, arts and humanities, these types of citation only represent a fraction of the attention given to research and undercount how it has been used more widely.

The central issue is that standard citation indicators reflect only the strength of influence within academia and are unable to measure impact beyond this domain. This gap has led many to investigate adjacent data sources for possible signals of impact, so-called altmetrics, searching for evidence that research engaged with society and led to wider socio-economic or cultural impact. Among these sources, policy documents (and more broadly grey literature) have been identified as an important venue where underpinning research is recognised as contributing to or influencing policymaking.

This view is supported by analysis of the REF2014 impact case study database which found that 20% of case studies were associated with the topic *Informing government policy*, most frequently in Panel C – social sciences. In these case studies, the organisations that utilised the underpinning research are listed in the *Sources to corroborate the impact* section, often with links to specific documents. Hence, it is possible to trace the impact of research through the published articles and their citations from grey literature sources, albeit with a relatively high analytical cost. The Dimensions database currently returns 37 publications that utilise the impact case study database to investigate policy influence.

Recently, Overton.io has emerged as a new data source that indexes these kinds of links and reduces the need for costly impact case study creation and manual data extraction. By crawling the publicly available documents from over 30,000 international sources, Overton is able to create a large database containing more than 5 million documents and over 14 million citations from reports, whitepapers, guidelines, parliamentary transcripts, legal documents and many other publication types. Our recent paper *Overton: A Bibliometric database of policy document citations* is a high-level survey of the bibliometric potential of the policy citation database, answering basic questions relating to the variety of documents and publication sources indexed (e.g. geography, language, year) and the network of citations that is extracted (e.g. volume, distribution, discipline, time-lag).

One key finding from the paper is that citations from documents indexed by Overton are not focussed on the hard sciences, as is found in traditional bibliometric databases, but are more likely to be made to journals in the social sciences disciplines. Table 1 shows the number of unique DOIs referenced and their total citations aggregated using Scopus ASJC research areas – a high-level aggregation across journal subjects.

ASJC Research Area	Number of Unique DOIs	Total Citations
Social Sciences and Humanities	839,011	6,036,912
Health Sciences	1,398,586	4,411,784

Date originally posted: 2022-03-23

Permalink: https://blogs.lse.ac.uk/impactofsocialsciences/2022/03/23/policy-citation-databases-offer-new-ways-to-understand-the-impact-of-social-sciences-research/

Blog homepage: https://blogs.lse.ac.uk/impactofsocialsciences/

ASJC Research Area	Number of Unique DOIs	Total Citations
Life Sciences	880,389	3,475,953
Physical Sciences and Engineering	407,796	1,495,634
Multidisciplinary	63,370	203,044

Table.1 Number of Overton references to Scopus ASJC research areas

By aggregating at a more granular level, namely research disciplines, it becomes clearer where references are concentrated, as shown in Figure 1. Although Medicine is the journal discipline cited most overall, it dominates the health sciences research area. In social sciences, citations are spread over a larger range of subjects and are more numerous on aggregate.

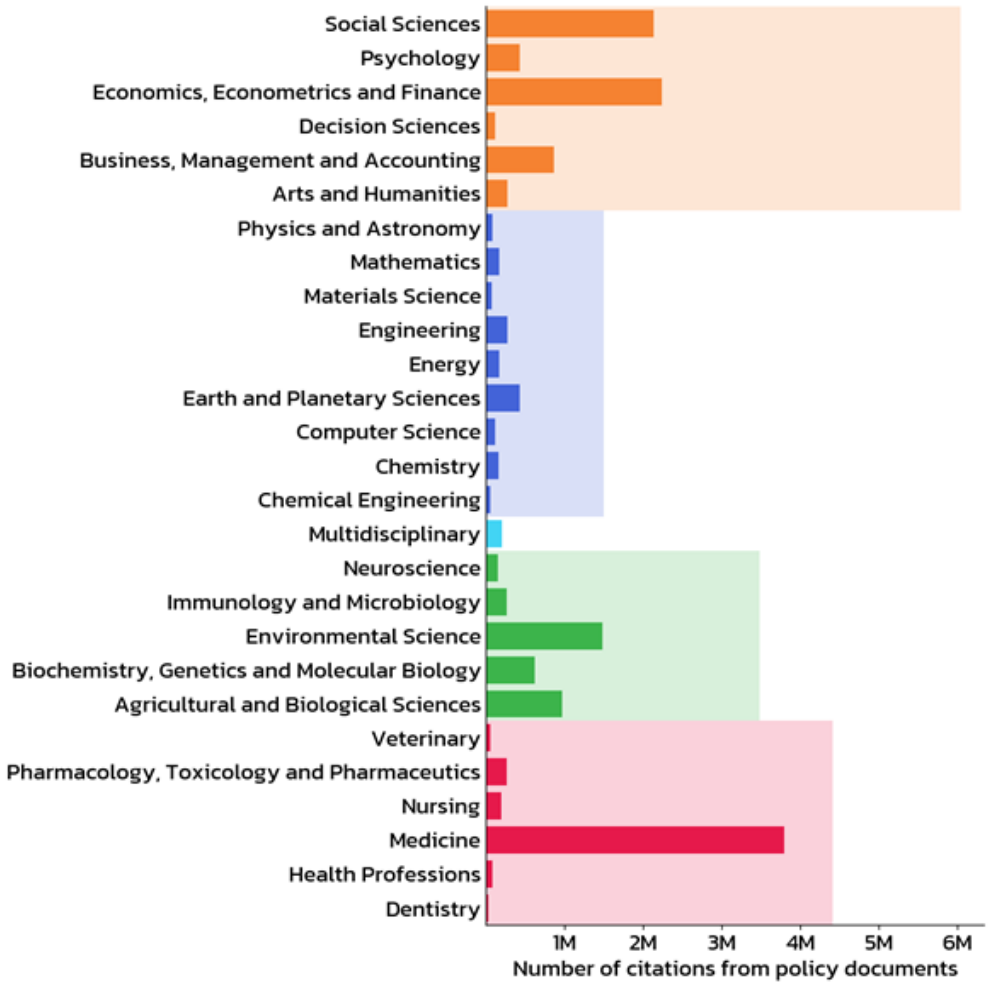

Fig.1 Citations from Overton policy document by ASJC Discipline

Date originally posted: 2022-03-23

Permalink: https://blogs.lse.ac.uk/impactofsocialsciences/2022/03/23/policy-citation-databases-offer-new-ways-to-understand-the-impact-of-social-sciences-research/

Blog homepage: https://blogs.lse.ac.uk/impactofsocialsciences/

Another striking difference is that of language representation. Those familiar with bibliographic databases will know they are dominated by English language publications, which severely limits the analytical potential in social sciences, arts and humanities, where research is often published in local languages. The documents indexed in Overton are far more representative, as shown in Figure 2. For each country listed, documents are classified as either published in English (orange), in a local language (purple), or another language (Other). Regions are also included for Intergovernmental Organisations (IGO), such as the World Health Organization and World Bank, and the European Union (EU). The only countries where local language documents are not the dominant content are countries where English is the native language, where there are colonial legacies (i.e. India and Singapore), or for the groups IGO and EU which publish for a global audience.

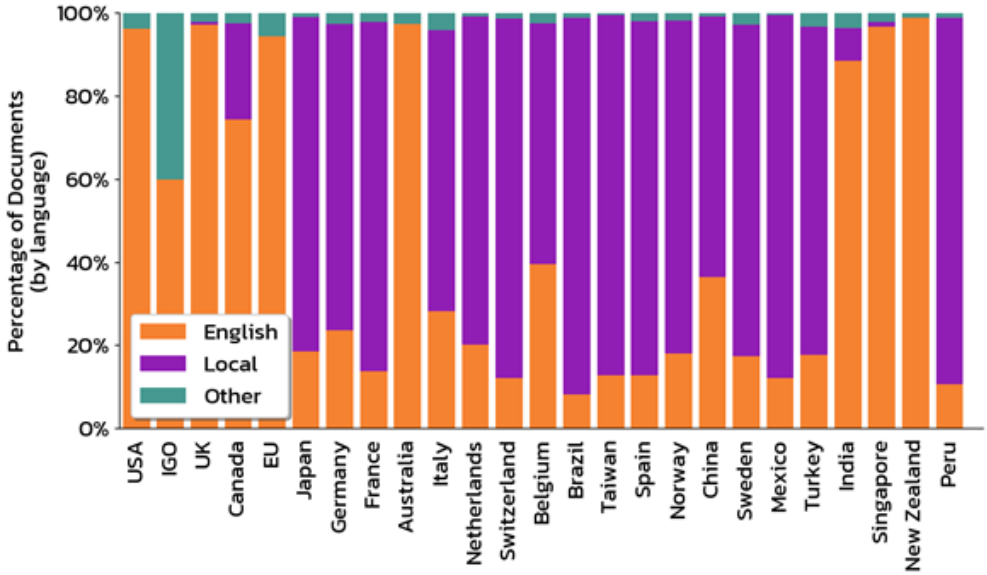

Fig.2 Language distribution of publications by publication source country

This is an exciting development for the research evaluation community, especially those with an interest in the social sciences. Not only does it show the potential for the survey and characterisation of large research portfolios, such as for institutions, funders, and publishers, it also provides new ways for individual researchers to understand how their own research is used outside academia, giving them additional resources to strengthen their professional profile.

The content generated on this blog is for information purposes only. This Article gives the views and opinions of the authors and does not reflect the views and opinions of the Impact of Social Science blog (the blog), nor of the London School of Economics and Political Science. Please review our comments policy if you have any concerns on posting a comment below.

Image Credit: Featured image Christa Dodoo via Unsplash, in text images reproduced with permission of the authors.

Date originally posted: 2022-03-23

Permalink: https://blogs.lse.ac.uk/impactofsocialsciences/2022/03/23/policy-citation-databases-offer-new-ways-to-understand-the-impact-of-social-sciences-research/

Blog homepage: https://blogs.lse.ac.uk/impactofsocialsciences/