

# 研究活動におけるオープンソース・データの利用に関する簡易調査

文部科学省 科学技術・学術政策研究所 データ解析政策研究室

## 要旨

本稿では、研究活動における新たな分析手法・指標の開発を念頭に、研究活動におけるオープンソース・データの利用状況の調査を目的として、物理・情報系分野におけるメジャーなプレプリントサーバである arXiv を対象に、プレプリント（原稿）中のオープンソース・オープンデータ言及回数を調査した。

ここでは、オープンソースとして github, オープンデータに Zenodo, figshare を取り上げて調査した。また、比較のための基礎データとして DOI も取り上げて調査した。本文中に記載されたメールアドレスを手がかりとして、各原稿には（割り当て可能なものについては）国籍を割り付け、初版発行の年月ベースで整理した。

結果、Zenodo, figshare ベースでみると arXiv 上でのオープンデータ言及はほとんど進んでいないことが分かった。他方で、github ベースでみた arXiv 上でのオープンソース言及は緩やかながら順調な伸びを見せており、2022 年以降全体投稿の 2 割以上の原稿において言及されていることが分かった。

## Brief survey on the use of open source / data in research activities

Research-Unit for Data Application, National Institute of Science and Technology Policy (NISTEP), MEXT

### ABSTRACT

The purpose of this paper is to investigate the use of open source data in research activities, with a view to developing new analytical methods and indices for research activities. We surveyed the number of open source and open data mentions in preprints (manuscripts) on arXiv, a major preprint server in the field of physics and information systems.

In this study, github was selected as open source, and Zenodo and figshare were selected as open data. DOI was also included as basic data for comparison.

We assigned nationalities to each manuscript (if assignable) using the e-mail addresses listed in the text as clues, and organized them on the basis of the date of first publication. The results were organized on the basis of the year and month of the first publication.

As a result, it was found that there has been little progress in mentioning open data on arXiv in terms of Zenodo and figshare. On the other hand, the github-based open source mentions on arXiv have shown a slow but steady growth, and more than 20% of all manuscripts have been mentioned since 2022.