

論文のどこに、どんな形式で データが引用されているか？ — DIASでの実態調査

中原 陽子*1、北本 朝展*2、清水 敏之*3、島井 博行*1、吉川 正俊*1

*1 京都大学、*2 国立情報学研究所、*3 九州大学

背景

- オープンサイエンスを推進する上で必要
 - データ作成者の成果を見える化することが課題になっている
- データ引用サービスの活用
 - データDOIを活用して引用を調査するサービスの存在（DataCiteのAPI等）
- 問題： サービスでは把握できないデータの利用事例がある
 - 例： 10.20783/DIAS.501
 - DataCiteのAPI利用8件
 - 実態調査結果55件
- 検討： データDOIだけでは把握できない利用事例を実態調査し、リポジットリが対応すべきことを検討する

方法

- DIASデータの利用成果論文を収集し、実際のデータ引用が論文の中のどこにどんな形式で出現するかを調査
 - DIAS公開データ357件（データDOIあり146件、データDOIなし211件）
- DIASデータの利用成果の収集（203論文、242か所）
 - 検索して収集（181論文）
 - Google Scholarを利用（データDOI、データDOI以外（データセットID、URL等））
 - ユーザの報告から収集（22論文）
 - 成果報告フォーム、ダウンロードユーザアンケート
- どこに：References、謝辞、指定セクション、本文、その他に分類
- どんな形式で：データDOIを記載、データDOI以外を記載に分類

指定セクションとは？

- AGU : Open Research

- <https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Data-and-Software-for-Authors>

DATA & SOFTWARE CITATION

Please cite in your References/Bibliography section a formal citation to the data/software described in the Availability Statement. Doing so will provide a citation credit for the data/software. Additionally, please cite data and software created by others used in your research, also to ensure that data or software is described in a separate data or software paper, please include deposited data or software as separate citations.

論文誌はReferencesへの記載も求めている

- Nature : Data availability

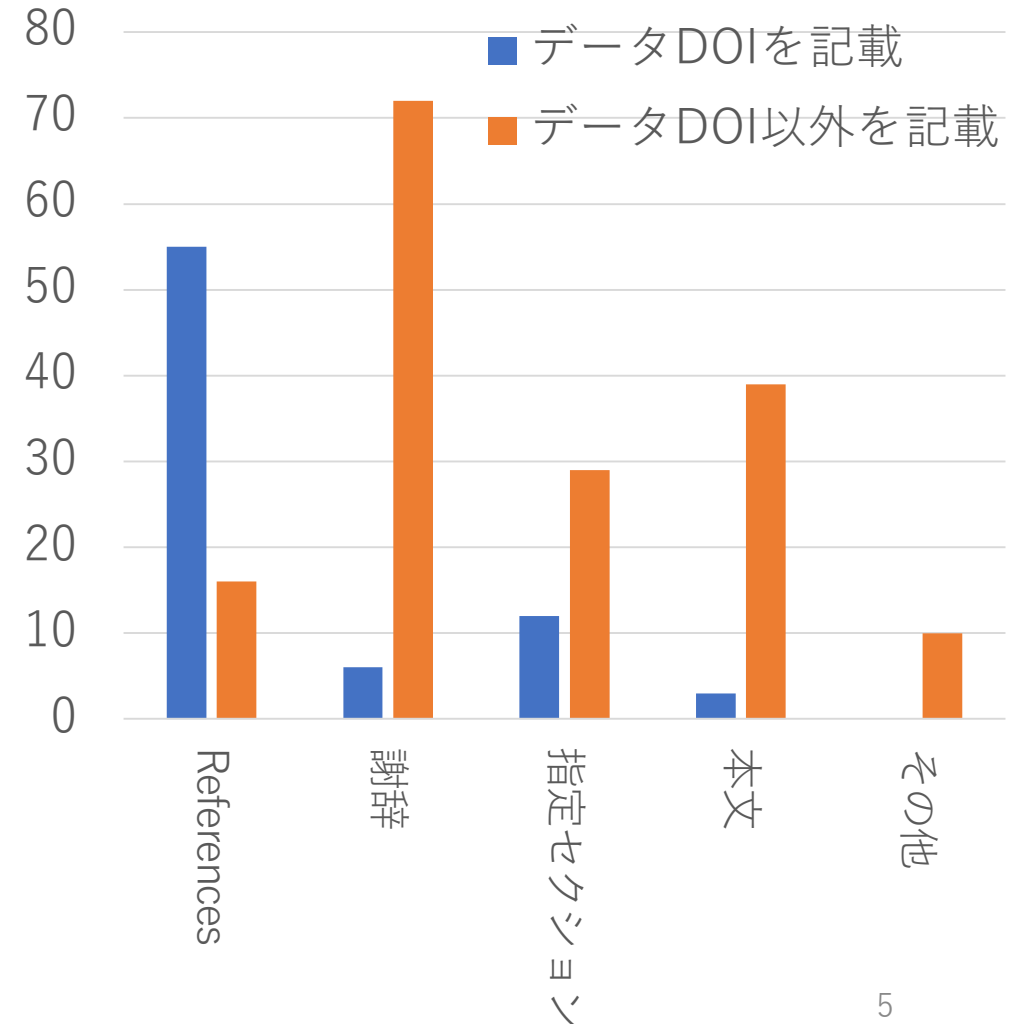
- <https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards>

Data citation: Datasets that have been deposited in repositories should be included as formal citations in the article reference list. This includes datasets generated during the study as well as existing datasets analyzed during the study. Citations of datasets should include the minimum information recommended by DataCite and follow Nature Portfolio style including: author(s), title, publisher (repository name), and identifier.

結果

- Referencesを利用する場合にはデータDOIが記載されることが多い
- References以外では謝辞が利用されることが多い
- データDOI以外の記載例：URL、データセット名、データセットID

データ引用記載箇所と記載形式



考察

- 理想：データDOIをReferencesに記載
- 現実：
 - ReferencesにはデータDOIが記載されることが多い
 - データ引用の記述場所としては謝辞の割合も大きい
 - データDOIではない形式で書かれている場合の検索は手間がかかる
 - ジャーナルのポリシーでは、指定セクションに加えてReferencesへの記載も求めているが、実際にはReferencesへの記載は少ない
 - 指定セクションでのデータ引用は41件、うちReferencesにもデータ引用があるものは3件

今後の課題

- データ引用を容易にする
 - リポジトリが引用文を例示することで引用形式を統一できる
- データ引用を調査しやすくする
 - 謝辞からReferencesにデータ引用の場所を移すことで、データDOIによる引用形式も普及することが期待できる
- データ引用を共有しやすくする
 - Mahalo Buttonを利用することでユーザが報告しやすくなる