## Linking research data and articles

東京大学 空間情報科学研究センター 小野 雅史 maono@iis.u-tokyo.ac.jp

## 背景と概要

- 近年、「ビッグ・データ」や「データドリブン・サイエンス」というキーワードが注目を集めているように、研究開発分野においても、データを中心にした新たな試みが検討され始めている。
- 例えば、従来では研究業績と言えば査読論文が中心であったが、一部の分野では、 データ・ジャーナル等を通して、次世代の研究のために研究成果物であるデータを提供する行為を、重要な業績として積極的に評価するといった取り組みが行われている。
- 現状では国際的にも、論文に関する情報は基本的にScienceやNatureなどのオンライン・ジャーナルを通して提供され、データに関する情報はGEOSSなどの地球観測データの統合プロジェクト等を通して提供されており、論文とデータの二つの情報の間にはまだまだ距離があるのが実状である。
- しかしながら、論文とデータの双方にDOI(Digital Object identifier)のような共通フォーマットのIDを付加するような試みも一部では開始されており、今後は、こうした論文情報と研究データとの連携が重要なテーマとなると考えられる。

## 提案内容

- 異なる情報を相互に関連付けるための手法・アルゴリズムには様々なものが存在するが、現実問題としては、最新の論文や質の高いデータは有料だったり法・権利関係がややこしかったりするものも多く、時に、高度なアルゴリズムの勉強よりも、必要な材料を集めて分析できるだけの社会的な条件を整えることのほうが難問だったりする。(参考として、2013年に起こったAaron SwartzによるMITのデータ大量ダウンロード事件など。)
- しかし、論文に関して言えば、本文は有料コンテンツであってもタイトル・著者名といった書誌情報は公開されていたり、またデータに関して言えば、データ本体は厳しく保護されていてもメタデータは比較的自由な活用が認められるケースも多い。
- そこで、論文の書誌情報とデータのメタデータの情報を使って両者を連携するアプローチを提案する。



## アプリケーションの試作

- CiNii ArticlesのOpenSearch API(http://ci.nii.ac.jp/info/ja/api/a\_opensearch.html)と「データ統合・解析システム(DIAS)」 (http://www.editoria.u-tokyo.ac.jp/projects/dias/)が提供するメタデータの情報を使い、両者を関連付けるWebアプリケーションを 詩作中。
- 下図の左側のリストがDIASにアーカイブされているデータに関する情報で、右側がCiNiiから取得した関連論文のリストである。
- 現在、実験用サーバで試験中の段階で、公開できるまでの条件が整っていないため、今回は紹介に留める。

