

2020JpGU Union 地球惑星科学の進むべき道 10
ビッグサイエンスとオープンサイエンス

チャット質疑応答討論

個人名は省略してあります。

- 01:38:27 座長：T:質疑応答は最後の総合討論でまとめて行いますので、質問などはチャットで随時お寄せください。
- 01:40:13 S:AGU Journal Specific Guidance is now available for four journal:
<https://www.agu.org/Publish-with-AGU/Publish>
- 01:43:46 座長のTさんビデオ消すの忘れてらっしゃる
- 01:46:15 座長マイクがなんか雑音拾ってるのかな……？

- 02:42:43 A:SCIENCE は自然を分析し統合しているが、原因と結果の関係をHOWとWHYで探求していると思います、オープンデータはHOWの結果であり、それを使ってWHYする理論の分野が遅れていると思います、原因を変える方法も含

- 02:44:34 R:質問：省庁等が責任を持って収集しているデータ以外について、各分野のデータのアーカイブ、維持と提供に責任を持つ機関・体制をどのように構築するか？予算の確保の方法も含めて検討して頂くと、将来の方向性を考える上で有用であると思います。

- 02:45:04 K:Hさんの発表のなかで触れられていた地震波形記録のデジタルイズについて、日曜日の高校生セッションでも発表がありましたので共有いたします <https://confit.atlas.jp/guide/event/jpgu2020/subject/O04-P02/advanced>

- 02:46:02 K:オープンサイエンス（オープンコラボレーション）のよい例かと思います

03:10:25 座長:N先生からの質問をご本人の許可を得て共有させていただきます。H先生へ質問です。

データ生成者へのcreditや評価を与えるために、データにDOIを付与することは非常に重要な「研究者の新しい常識」にすべきだと思います。ただ、現状では、DOIに基づいて研究者を評価しようというまでには至っていないと思います。今後、こうしたことを推進するためには、どのような努力や試みが必要でしょうか。研究論文数・被引用数を評価としていくためには、何が必要でしょうか？

03:10:39 M:長文すみません。この場で議論できることではないのかもしれませんが、一言書かせていただきます；

データの生成・整備もふくめて科学成果を出すための活動はなんらかの評価対象になるべきという国際動向と理解します（ボランティアの論文の査読者が集まらないと嘆くエディターの声を聴いたこともあります。そういう科学者が科学活動全体に対する奉仕活動をそのままにしておいてよいのかという問題です）データについては、先端的な観測実験では、測定理論、手法、データ取得方法、処理アルゴリズムなどが科学者の重要な研究活動・業績の一部になってきました。一方、どなたがやってもルーチン的にほぼ同質なデータ取得できるものと切り分けが必要です。前者については科学者の重要な業務であり、また最先端の研究現場にいつづけなければよいデータは生成・整備されません。その科学者活動としての評価をしなければならないのではないのでしょうか。私が以前から申し上げている持論は、大学教育において教育と研究半々を重要なエフォートとするものとするならば、今後は、データの生成・整備と研究とを主なエフォートとする、データマネジメント兼務研究職などの職業枠で若手研究者の育成ふくめて進めることも重要な考え方ではないかと考えていますが、いかが思われますでしょうか？

少子高齢化して教育への資源投資を減らしてよいなどという意見があったとしても、日本という国が科学技術の投資を減らして立ちゆくのか、質の高い科学研究を行う大学や研究機関は教育にとどまらず論文だけでない科学からの成果物を社会的存在価値としてどう位置付けていくか、という視点もふくめて、人類共有資産としてのデータその他科学の成果物の問題が議論されることが重要と思います。

- そこでは共有資産となったデータを生成した科学者が、例えば職を離れたり亡くなったりしても後世の研究者が利用可能なデータ資産となりように整備するコストが必要になります。従来の研究予算の一部でこれらをすべて負担することは、すでに先進諸国では無理があると国も考え始めて新たな施策が現れています。日本の科学はこのままでは社会全体の便益をむしろ損ねる危険がないのか、社会全体として科学への資源投資のありかたを考え直す議論があってよいのではないのでしょうか)

- 03:26:10 M:Creative Commonsライセンスはどれくらい、オープンサイエンスにおいて求められるのでしょうか？利用申請を条件に提供するデータは、Creative Commonsにおける再利用可能の条件を満たすのが難しいように思います
- 03:32:04 M:再利用可能でなく、再配布可能の間違いでした。みません。
- 03:32:53 O:Creative Commons の件、データではない話ですが、私が先ほど紹介したGIS実習オープン教材では、コンテンツについて CC BY-SA 4.0 を明記しています。その方が、円満に再利用してもらえると考えたためです。共同研究者に、この種に詳しい人がいて、いろいろ教えてもらいましたが、この種を意識的に明記するという状況は、まだ地球惑星科学では相対的に少ないように思います。
- 03:33:38 K:どのライセンスか、というよりも、利用条件が明記されていることが重要なのではないのでしょうか？
- 03:34:42 M:利用申請が必要な理由によると思いますが、再配布を許さないのであればCCライセンスにはなじまないのではないのでしょうか。
- 03:50:17 Y:「データ整備を行うこと評価する」という点に於いて、DOIやCitationを増して定量化することも重要ですが、ポジションを用意するということが何よりも重要だと思います。データ整備が進まない背景には、データ整備ではポジションが得られない現状があります。
- 03:51:47 H:賛成です。評価して、ポストを用意することが、科学コミュニティとして重要です。
- 03:53:45 M:データ整備に従事して路頭に迷いかけているポストですが、とりあえずいますぐできることとして、データを利用する方は、データを適切に引用をしていただくことを希望します。
- 03:54:44 N:上記の議論、おおむね賛成です。ポストとまで一足飛びにはいかないにしても、研究者の評価軸の一つとして認識されるようになる必要があると思います。
-

- 03:54:54 A:私もY先生、H先生のご意見に賛成です。加えてい言うならば、長期的にデータの整備を行うには、位置づけと、それに基づく予算措置が重要だと考えます。
- 03:55:53 M:先日UNESCOが出す予定のオープンサイエンス提言書のための、国際アカデミーからの寄与文書を査読する機会を得ました。オープンサイエンス活動に関する人材評価はそこにも書き込まれていました。UNESCO提言が政府の施策にどこまで実効的に効果を発揮するかわかりませんが、科学者コミュニティからも主張していくことは重要と思います。
- 03:56:43 N:ただ、評価する方のポジションにおられる方の認識が変わる必要があります。現在は、データDOIやその雑誌での引用など、道具立ては揃いつつありますが、それを使った結果がどのように利用されていくか、ですね。
- 04:02:03 H:はい。そうでね、研究者の業績評価として、自著の論文（数と非引用数）だけでなく、生産されたデータとその引用数で評価する必要があります。データへのDOI付与は、これを進めることに貢献します。大学等の研究機関での研究者の評価の「項目」として、DOIのついた「データ論文」も取り入れることを提案します。
- 04:06:55 H:質問ではないですが、これだけの人数が接続していると、顔をstreamingで配信することでZoomの帯域を圧迫されそうで心配です。断続的に入っている「プツプツ音」はそれが原因かもしれません。
- 04:07:26 F:予算の出し方で、米国のように、データは公開しないといけないということに決めてしまえば、データ公開をする人を育成して雇わなくてはならないということになり、評価もせざるを得ない...

04:07:46 M:データ引用に関して、論文も同じといわれるとそうかもしれませんが、引用数がいまは少なくても、将来評価されて利用（引用）されるということもデータに関しては特にあるのではないかと思います。引用数が少ないからよくないと言われると、それは危ういのではないかと思います。引用するという土台は論文と同じ形式であっても、引用数などの指標は論文の評価関数とは別の評価関数でないとまずいと思います。

04:09:18 F:引用数にかんする村上さんに賛成です。データは蓄積と公開によって、将来役立つものなので、

04:10:35 K:ロングタームの長期にわたる過去データを収集するプラットフォームの構築は非常に重要だと考えます。データを蓄える場所というよりも高品質のデータを整理する場（プラットフォームの構築をNIIは現在目指しています。

04:11:13 A:JpGUで各種のデータや情報の扱いをオープンデータとして議論するのは良いと思いますが、様々な分野の様々なデータ情報があり、扱いも様々なので、まずはJpGUで対象とするデータや情報の扱いに関するマッピングをすることがまずは重要かと。

04:12:40 H:引用数がすくないデータに価値がないとは、私も、思いません。第一歩として、まず、評価するという意味です。論文でも、10後に評価されることがあります。

04:12:51 A:データを取得した本にではなく、第三者がデータを維持する体制を作らない限り、本人が引退するとデータが失われることになると思います

04:14:17 S:引用数に関するコメントがありますが、データがどのような研究に使われたかも大事な情報です。その情報がPI、データベース整備にかかわった人、データベースのプロモーターに共有されるようなシステム構築・設計が必要です。

04:15:19 H:はい。引用数だけが、重要ではないというのは、賛成です。まず、評価する仕組みが必要です。

04:15:33 N:H先生、M先生に賛成です。データ被引用数だけに注目する必要はないと思います。

04:16:12 H:データ作成のシステムを作った人だけでなく、それを維持している研究者・技術者への（高い）評価が必要です。

04:16:26 N:DOIでデータ出版数(データ生成寄与の大きさ)は評価できますので。

04:16:57 Y:データ整備の専門性・高度化が強まるに従い、研究者のサイドワークの範疇を超えてきています。大学院重点化の名残の研究者の多い現代と比較して、将来にも同じ品質でデータ整備がなされるには、ノウハウのシステム化が必要不可欠だと思います。

04:17:45 A:1次データもしくはサンプルの意義とは、改竄されていないこと、ではないかと考えます。

利用後の2次データも論文の根拠として重要かと思いますが、2次データはその時代のバイアスがかかっているものと考えられますから、常に最新もしくは原点に立ち戻って研究するため資料としての1次データ保管は重要と思いますし、保管と公開は世界標準とすべきだと思います。1次データの重要性について理解を得られるよう声をあげ続けていくことが必要だと思います。

04:18:07 M:Aさんのご提案にあるように、第三者がデータを維持するのは重要だと思います。大学の理学に所属する立場では部局ではなんとも動きそうにないので、図書館がやはりある程度はその役割を担っていただくのがやりやすいと思います。まずは、あるサイズ以下のデータで開始して、行ければと思います。

04:18:33 F:これだけのお金を付けたからこれだけの成果が出た、ということが言いたいために「評価」したいんだと思いますが、データによる成果を客観的に測るのはもう無理なので、そうとうに主観的な評価でないとまずいと思います。

04:18:34 Y:評価システムのフレームワークを構築すると共に、同じ品質のデータ整備をより少ないリソースで整備できなければ、将来のデータは低品質の上に成り立つ荒いサイエンスで我慢する必要があります。

04:18:35 S:過去データについては、特にリモートセンシングについてはデータ処理がたびたび変更されており、生データからの処理をやり直すということが相当大仕事になります。

- 04:18:40 M:Uk、EUで特に、metrics、つまり評価手法の研究が盛んな印象です。論文評価尺度はbibliometrics計量書誌学でしたが、データ計量書誌学が今後研究される必要があります。
- 04:19:12 K:国所有や個人所有の建物の建物内で観測した地震観測データのように私有財産or機密と関わるデータの場合、オープンデータ化が難しい場合もある分野への考慮も同時に行われていって欲しいと考えています。オープンデータでなければジャーナルで発表できないという動きにはならないといいなと感じます。
- 04:19:45 K:データをどう残すのか？ は コンピュータサイエンスでも永遠の課題です。一緒にご検討させてください。
- 04:19:51 H:デジタルデータは少なくとも源データを全部残す必要があります。
- 04:23:27 F:>> コンビナーの皆様 このチャット欄に書かれたことをも、まとめに含めて頂きたいです。参加者の方ではスクロールしてしまっとうまく保存できません。

以上