

アーカイブの利用：天文学の場合

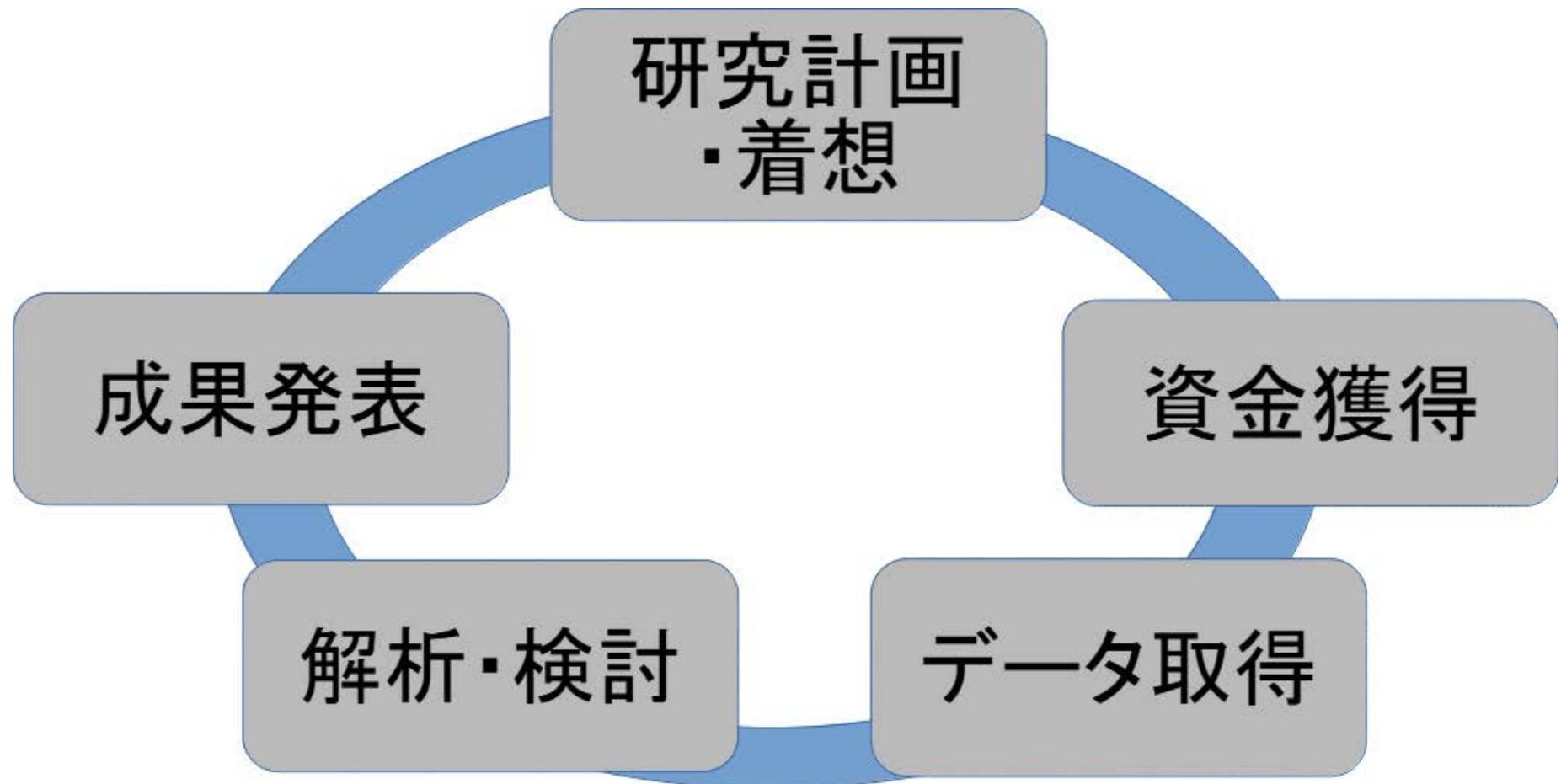
- 2020/10/16 デジタルアーカイブ学会第5回研究大会
サテライトワークショップ「自然史・理工系デジタルアーカイブの現状と課題」
- 玉澤春史(京都市立芸術大学/京都大学)
- (もち時間15分)

※本発表は分野を代表するものではありません。

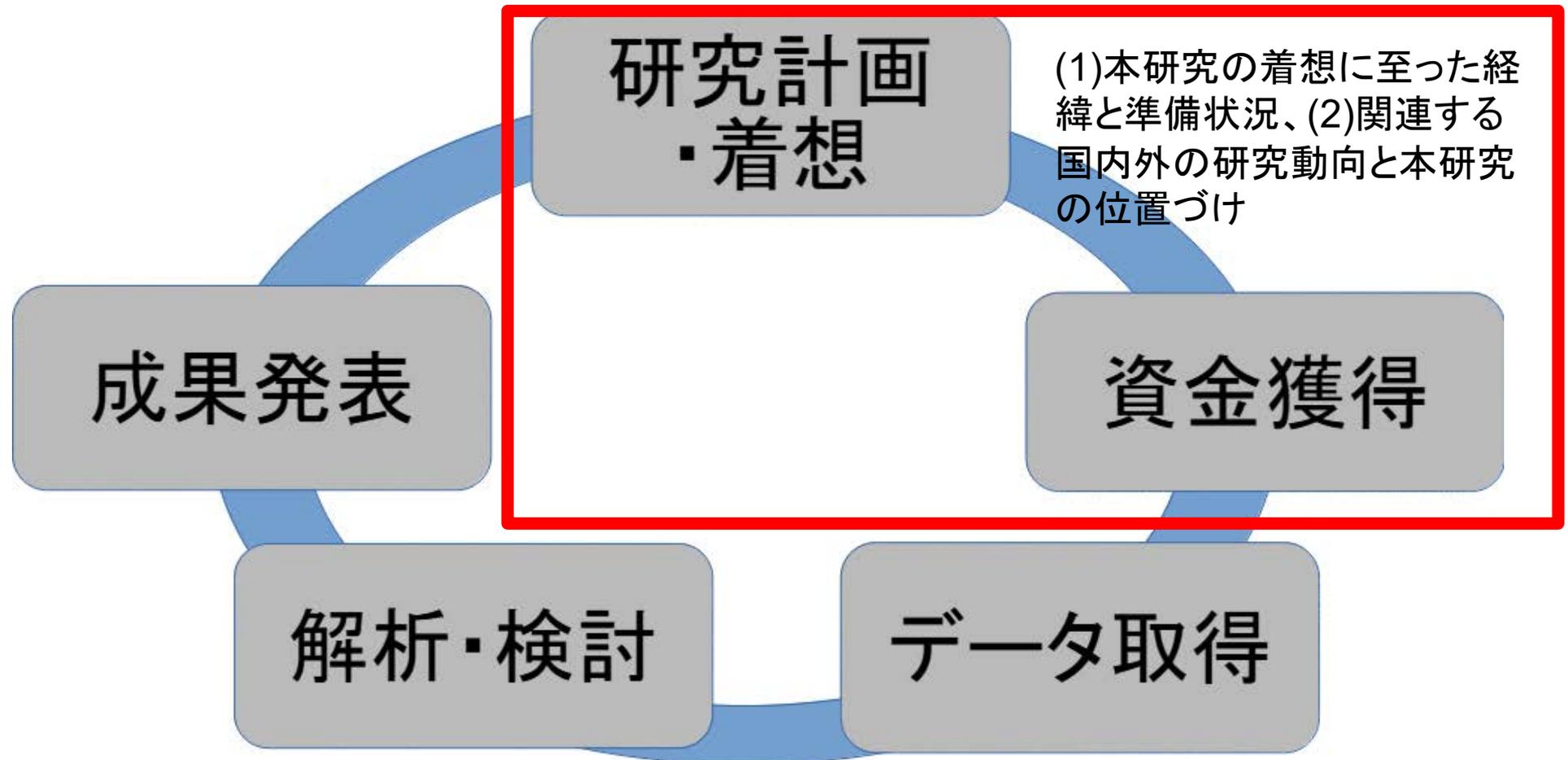
他の分野の方と話して感じる 天文学のアーカイブの特徴

- 公開性が高い
- データが大きい
- 検索が独特
- 新しいものの「アーカイブ」

研究のサイクルから アーカイブの使われ方を見る



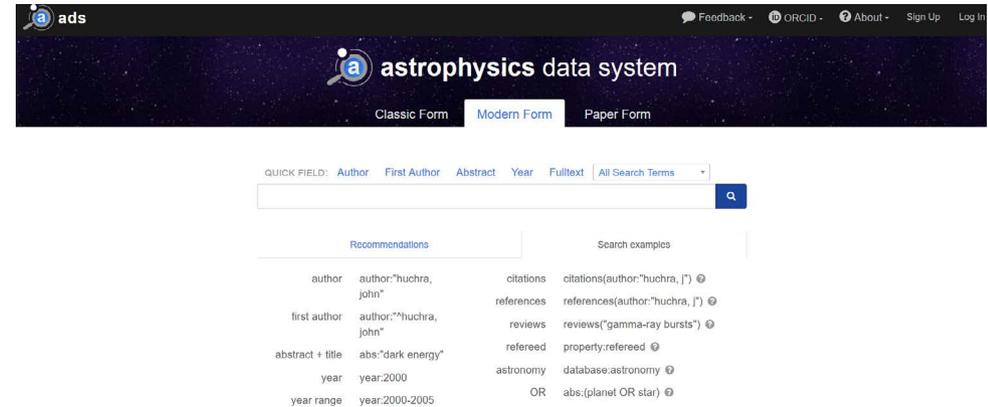
研究のサイクルから アーカイブの使われ方を見る



ADS

astrophysics data system

- 古いものはPDFスキャン, 最近のは雑誌へのリンクなど
- 引用・被引用論文をたどるのが容易
- 天体名などもカバー



使用例

The screenshot displays the ADS search interface. At the top, the 'ads' logo is on the left, and navigation links for 'Feedback', 'ORCID', 'About', 'Sign Up', and 'Log In' are on the right. Below the header, a search bar contains the query 'tamazawa,h' with a search button. A message indicates 'Your search returned 22 results'. On the right side of the results area, there are buttons for 'Date', 'Export', and 'Explore'. On the left, a sidebar lists authors and collections. The main results area shows four entries with checkboxes, document icons, and citation counts. A bar chart on the right shows the number of citations over time, with a legend for 'refereed' (blue) and 'non refereed' (green) citations.

QUICK FIELD: Author First Author Abstract Year Fulltext All Search Terms

← Start New Search tamazawa,h

Your search returned 22 results

Export Explore

AUTHORS

- Tamazawa, H 20
- Hayakawa, H 18
- Isobe, H 16
- Kawamura, A 14
- Ebihara, Y 11

more

COLLECTIONS

- astronomy 19
- physics 16
- general 1

1 2019ApJ...887....7I 2019/12 cited: 1
Intense Geomagnetic Storm during Maunder Minimum Possibly by a Quiescent Filament Eruption
Isobe, Hiroaki; Ebihara, Yusuke; Kawamura, Akito D. *and 2 more*

2 2018A&A...616A.177H 2018/09 cited: 13
A great space weather event in February 1730
Hayakawa, Hisashi; Ebihara, Yusuke; Vaquero, José M. *and 9 more*

3 2018PASJ...70...63H 2018/08 cited: 7
Sunspot drawings by Japanese official astronomers in 1749-1750
Hayakawa, Hisashi; Iwahashi, Kiyomi; Fujiyama, Masashi *and 7 more*

4 2018SoPh...293...8H 2018/01 cited: 11
Iwahashi Zenbei's Sunspot Drawings in 1793 in Japan

Years Citations Reads

■ refereed ■ non refereed

| Year | Refereed Citations | Non Refereed Citations |
|----------|--------------------|------------------------|
| 175-1978 | 1 | 0 |
| 179-1982 | 0 | 0 |
| 183-1986 | 0 | 0 |
| 187-1990 | 0 | 0 |
| 191-1994 | 0 | 0 |
| 195-1998 | 0 | 0 |
| 199-2002 | 0 | 0 |
| 203-2006 | 0 | 0 |
| 207-2010 | 1 | 0 |
| 211-2014 | 0 | 0 |
| 215-2018 | 16 | 1 |
| 2019 | 0 | 0 |

arXiv: プレプリントサーバー

The screenshot shows the arXiv.org website in a browser. The browser's address bar displays "https://arxiv.org". The page header includes the Cornell University logo and the text "Cornell University". Below this, the arXiv.org logo is prominently displayed on a red background. To the right of the logo is a search bar with the placeholder text "Search...", a dropdown menu set to "All fields", and a "Search" button. Below the search bar are links for "Help" and "Advanced Search".

The main content area on the left contains a paragraph describing arXiv as a free distribution service and an open-access archive for 1,777,731 scholarly articles in various fields. Below this text is a "Subject search and browse:" section with a dropdown menu currently set to "Physics" and a "Search" button. There are also buttons for "Form Interface" and "Catchup".

Below the search section is a "News" section with the text "arXiv now processes new submissions and replacements with TeX".

On the right side of the page, there is a "COVID-19 Quick Links" section. It contains the text "See COVID-19 SARS-CoV-2 preprints from" followed by a bulleted list of links: "arXiv" and "medRxiv and bioRxiv". Below the list is an "Important:" notice stating that e-prints on arXiv are not peer-reviewed and should not be used to guide clinical practice or health-related behavior without consulting multiple experts.

arXiv is a free distribution service and an open-access archive for 1,777,731 scholarly articles in the fields of physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, electrical engineering and systems science, and economics. Materials on this site are not peer-reviewed by arXiv.

Subject search and browse:

Physics

Search

Form Interface

Catchup

News

arXiv now processes new submissions and replacements with TeX

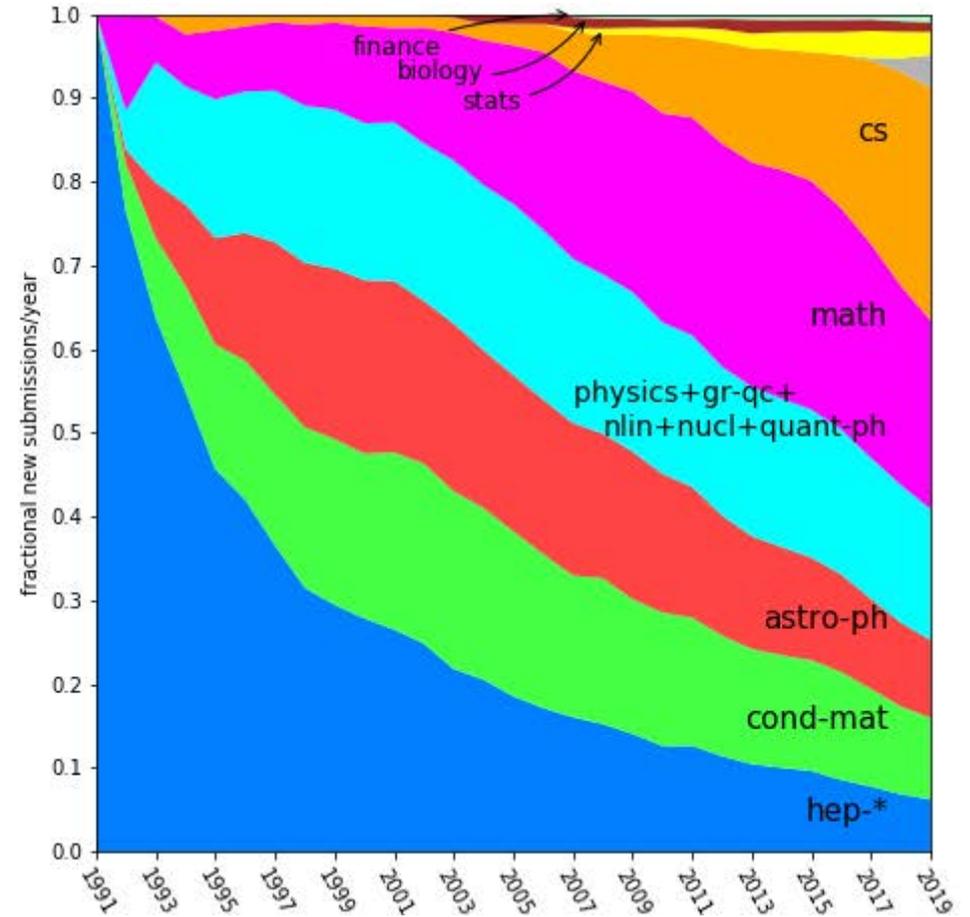
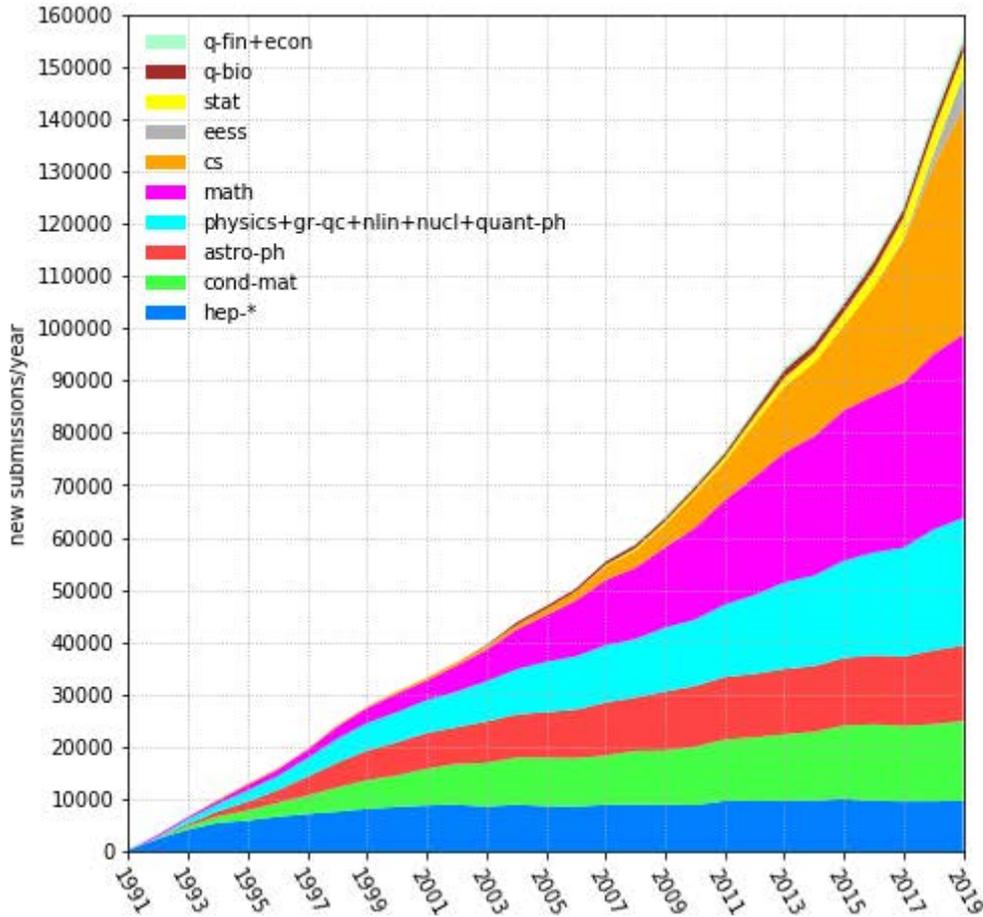
COVID-19 Quick Links

See COVID-19 SARS-CoV-2 preprints from

- [arXiv](#)
- [medRxiv and bioRxiv](#)

Important: e-prints posted on arXiv are not peer-reviewed by arXiv; they should not be relied upon without context to guide clinical practice or health-related behavior and should not be reported in news media as established information without consulting multiple experts in the field.

増え続ける投稿量



https://arxiv.org/help/stats/2019_by_area/index

分野によるarXivの使われ方 (私見)

- 天文学: acceptedとなったものを投稿する
- 物理学の一部: submitted段階で投稿, reviseによりreplace
- 情報学: 国際会議用の論文(予稿)も結構多い

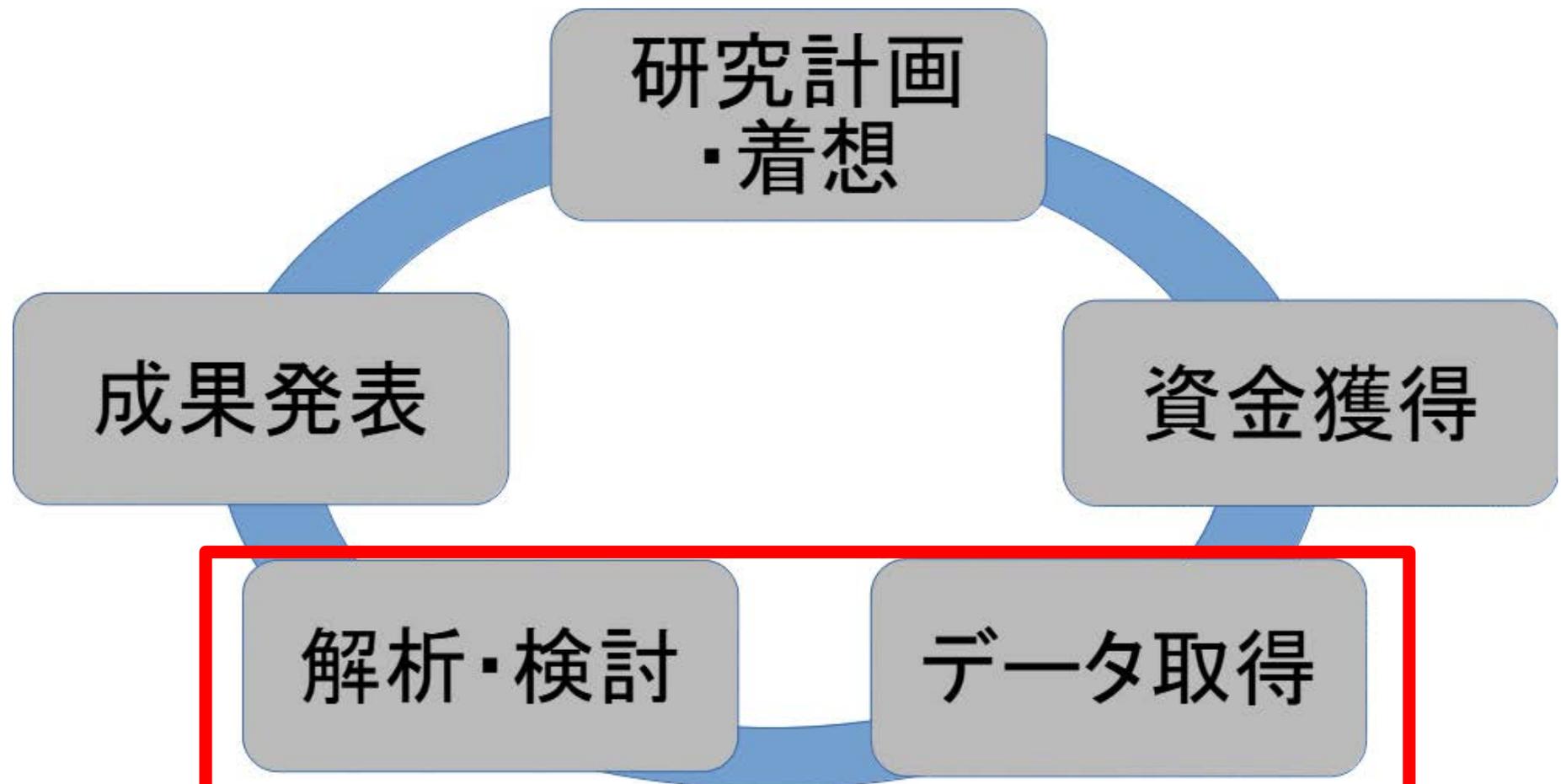
参考

林 和弘, 小柴 等 「arXiv に着目したプレプリントの分析」,
NISTEP DISCUSSION PAPER, No.187, 文部科学省科学技術・学術政策研究所.

補足 arXivとcovid-19

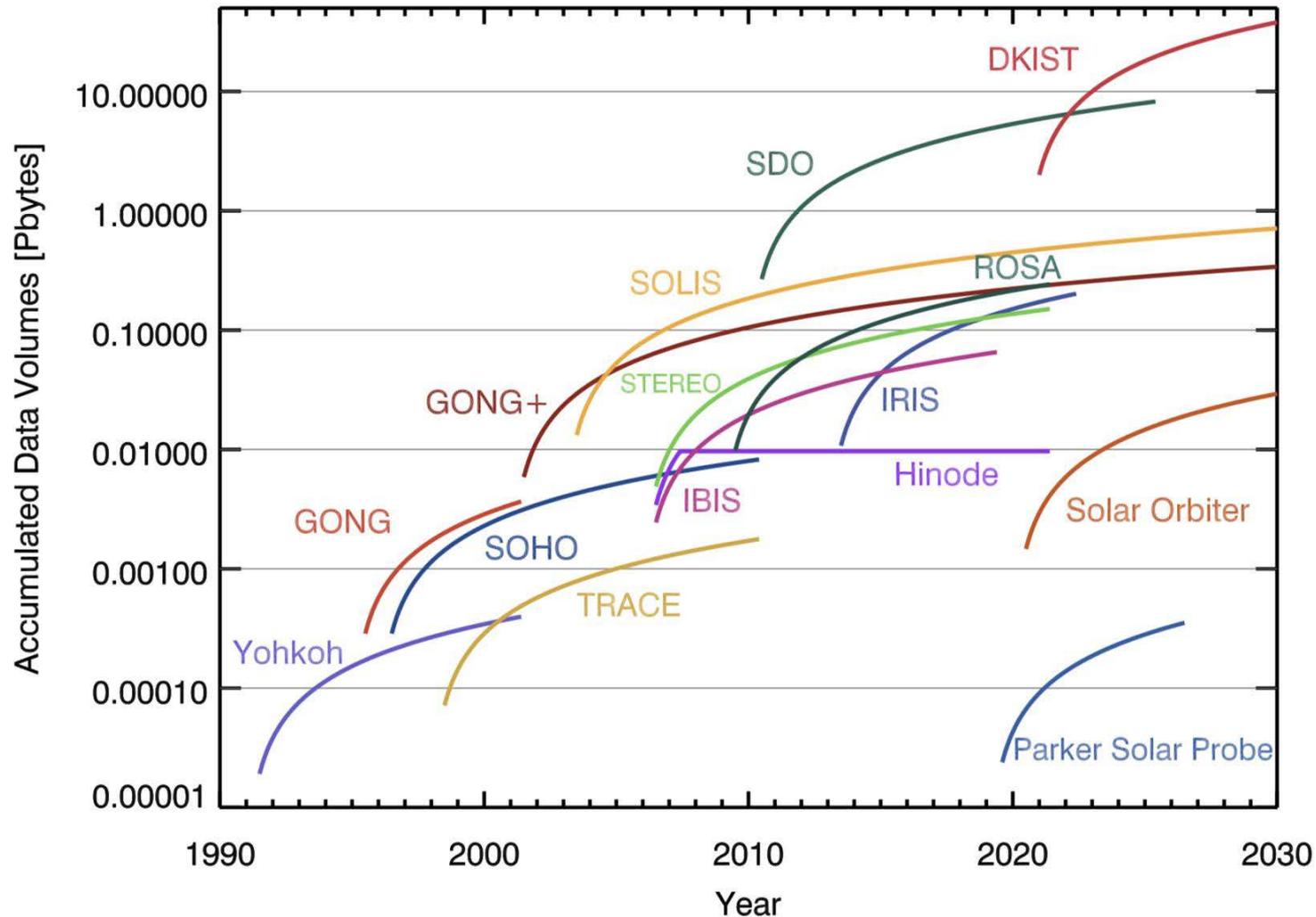
- 速報性が重要視されたのか, arXiv, bioarXivなどに関連preprintが投稿増
- arXivのトップページも現在はずぐに関連記事に跳べるように
- 一部タグのついた記事が急増
- physics.soc-phのタグがついたプレプリントは, 2019年1月より四半期ごとで141, 146, 131, 143, 148, 403, 370

研究のサイクルから アーカイブの使われ方を見る



大きすぎるデータをどう探し, 集め, 扱うか

増えるデータ量 太陽物理学の場合



<https://twitter.com/NatSolarObs/status/113678572454252544>

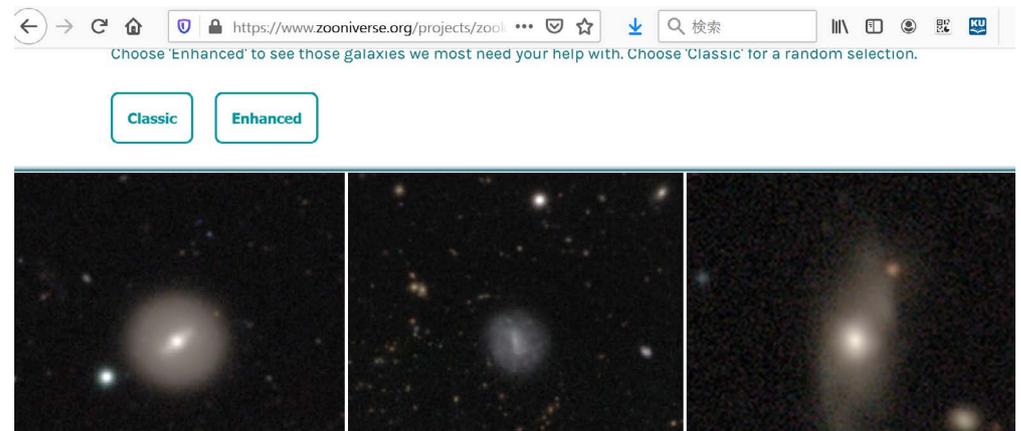
国立天文台 天文データセンター

- 多波長データ解析システム
- 天文データアーカイブセンター
- Hyper Suprime-Cam Subaru Strategic Program
- Japanese Virtual Observatory
- 天文データセンター講習会
- 国立天文台観測データ公開システム/解析システム
ム総合案内

Citizen Scienceのアプローチ

Galaxy Zoo

- 銀河の形態分類を一般市民に判断してもらう
- 研究者のみで行うより早く, さらに特徴的な発見もされるなど
- arXiv同様, オープンサイエンスの成功例として取り上げられる
- 現在はZooniverseの一部



日本のすばる望遠鏡データを使った Galaxy Cruise

Citizen Science Project
GALAXY CRUISE

日本語

ステージ 1 コース: 本番 お便り 操作ガイド ミッション クレジット
ケトウス・ケントルム街 ランキング ログアウト

渦巻銀河? 楕円銀河?

渦巻銀河 楕円銀河 どちらでもない

1 2 3

$\alpha=02:20:13.1413$ $\delta=-04:38:46.5622$ $L=3.4377'$

研究のサイクルから アーカイブの使われ方を見る

アピーリングにするには
他の分野は使うか

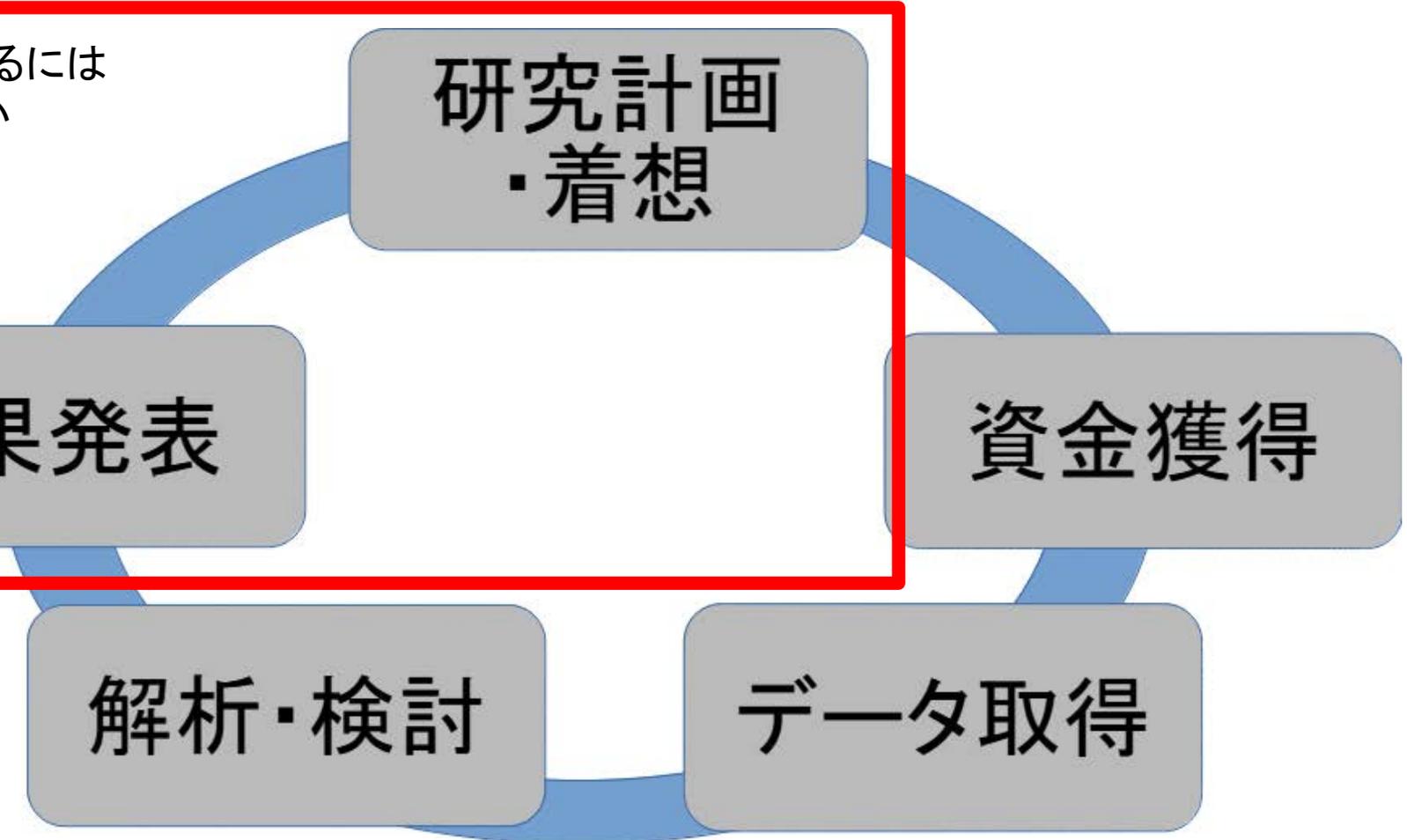
研究計画
・着想

成果発表

資金獲得

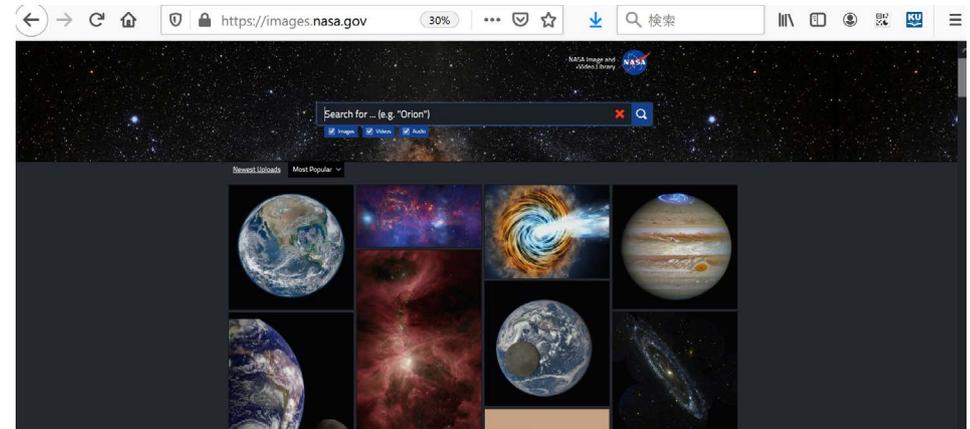
解析・検討

データ取得



NASA image and video library

- NASA content - images, audio, video, and computer files used in the rendition of 3-dimensional models, such as texture maps and polygon data in any format - generally are not subject to copyright in the United States.



<https://www.nasa.gov/multimedia/guidelines/index.html>

「自然史、理工系デジタルアーカイブの現状は人文系に比較して認知度は低い」か

- (そもそもアーカイブという概念の認知度はどうか)
- arXivもADS(の一部)もアーカイブだが使っている人たちはあまり意識していない
- arXivに関してはcovid-19関連でやや認知度が向上したかも
- NASA imageについてはあまりにフリーなので使われているが意識されていない
- 他分野の人が天文観測データを使って何かやる, というのはどうい
う場合か. 逆は?

一例：芸術分野での利用

- 衛星データを使った作品制作，展示を企画（衛星写真含む）
- 専門家とのコミュニケーションを通してお互い何がしたいか理解



図3 (左) Bamy [天空]



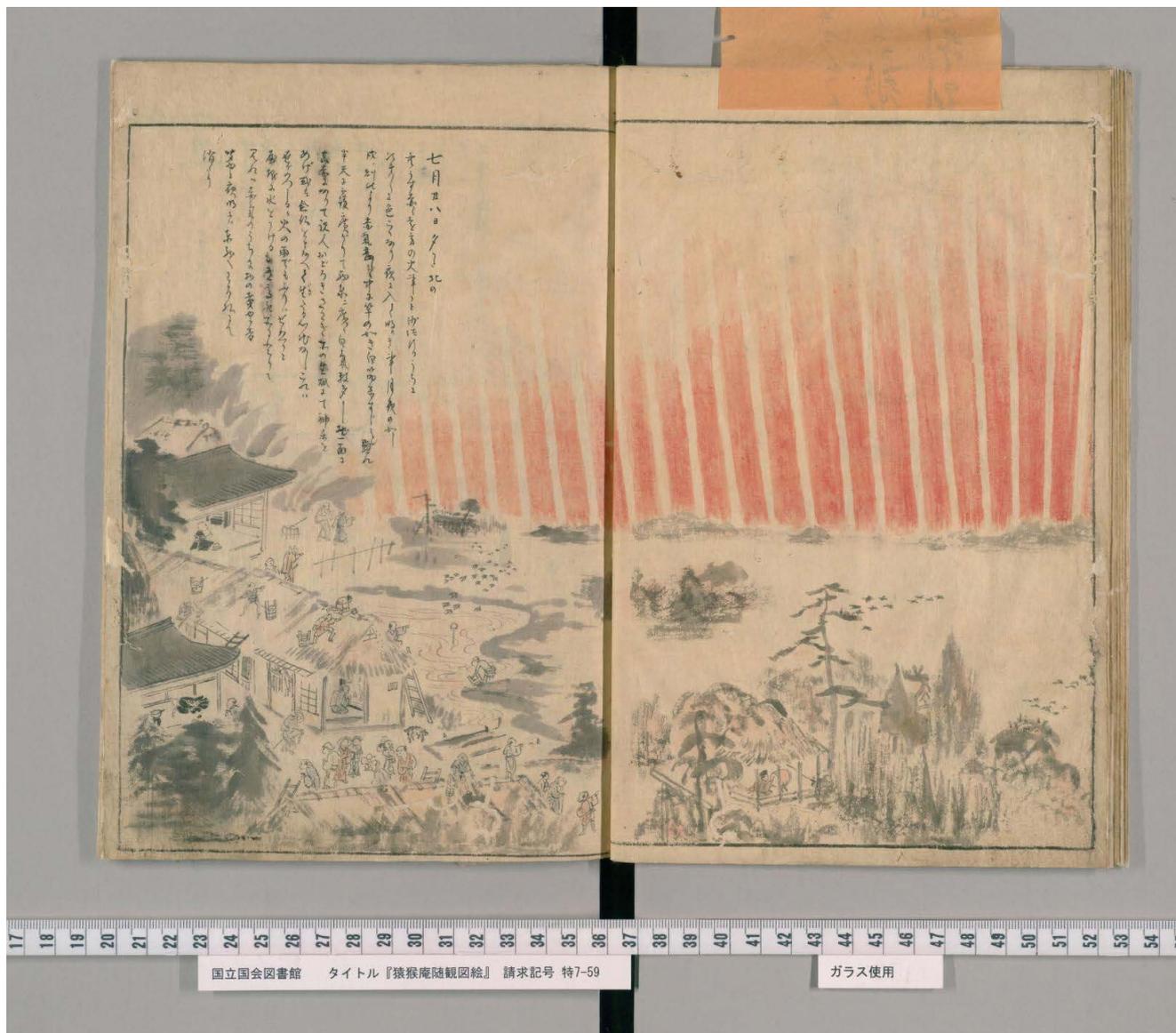
図4 (右上) 森田存 [宇宙(そら)の数だけ] (一部)



図5 (右下) 高原秀平 [跡の庭]

歴史文献の記述から過去の太陽活動を調べる

- 2014年より着手
- 一番最初は
- 「新漢籍全文」によるサーベイ



まとめと捕捉

- 公開性が高い:「データベース天文学」が可能
- データが大きい:解析環境の提供, ツールのコミュニティ内での開発 (pythonでの書き換え進む)
- 検索が独特 (google scholar と同様にADS活用)
- 新しいものの「アーカイブ」:各地で行われていた小規模のモノ, やや古いものは後手になりがち (地方の昔の写真を集めるとかは大変)