

SMOKAでデータを公開しませんか？

要旨

データを公開・共有することで、研究結果を検証したりデータの再利用をしようという考えが科学研究コミュニティに広がっている。

OISTERに参加する望遠鏡のデータについても公開を検討する価値があるだろう。

ここでは、その手段としてSMOKAでの公開を提案する。

SMOKAについて簡単に紹介し、SMOKAでデータを公開するにはどうすればよいかを話す。

中島 康 (国立天文台・天文データセンター)

SMOKAチーム (樋口 あや、格和 純、市川 伸一)

SMOKAとは

Subaru-Mitaka-Okayama-Kiso Archive system

光赤外天文データアーカイブシステム (2001年～) @ 国立天文台

公開生データ (+ クイックルック画像、気象データ) を提供

すばる望遠鏡 : Suprime-Cam, FOCAS, HDS, OHS/CISCO, IRCS, CIAO, COMICS, CAC, MIRTOS, MOIRCS, Kyoto-3DII, HiCIAO, FMOS, HSC, CHARIS, IRD, SWIMS, MIMIZUKU

岡山観測所 188cm : ISLE, KOOLS, HIDES, OASIS, SNG, MuSCAT

木曾観測所 : 1kCCD, 2kCCD, KWFC, KISO Schmidt Plates digitized data

MITSuME (東工大) : MTA(明野), MTO(岡山)

東広島天文台 : HOWPoI, HONIR

西はりま天文台 : NIC

せいめい : KOOLS-IFU (共同利用枠のみ セメスタ20B~)

国立天文台は大学共同利用期間であるので、
国立天文台だけではなく大学の望遠鏡のデータも扱っている。

32装置 (2020年11月現在)

全天スカイモニタ画像の検索とダウンロード

岡山観測所 東広島天文台 MITSuME(明野) 木曾観測所

SMOKAのデータアーカイブシステム

観測所から**データ転送**

データを**ストレージ**で保管 サーバ室で管理された冗長化ストレージ

データベースでメタデータを管理

データベースと連動した**検索システム** ユーザはウェブブラウザで検索

ウェブブラウザおよびメールによる**データ請求システム**

ダウンロードシステム FTPサイトにデータ配置し、メールでユーザに通知

様々な条件を指定して絞り込み検索

Search Conditions

Object Name (for name resolve) **Resolver**
 SIMBAD NED
 Don't Resolve

Coordinate System **center (RA)** **Radius(arcmin)**
 Circular **center (DEC)**
Equinox

Field of View (arcmin) Rectangle
From (RA) **To (RA)**
From (DEC) **To (DEC)**

Observation Date **Exp Time (sec)** **Observer**

Frame ID **Exposure ID**

Output Format
 TABLE (max 5,000 rows, HTML) ASCII (max 20,000 rows, text)

Frame or Shot mode
 FRAME SHOT (SUP, HSC, and KWF ONLY)

Instruments/Filters

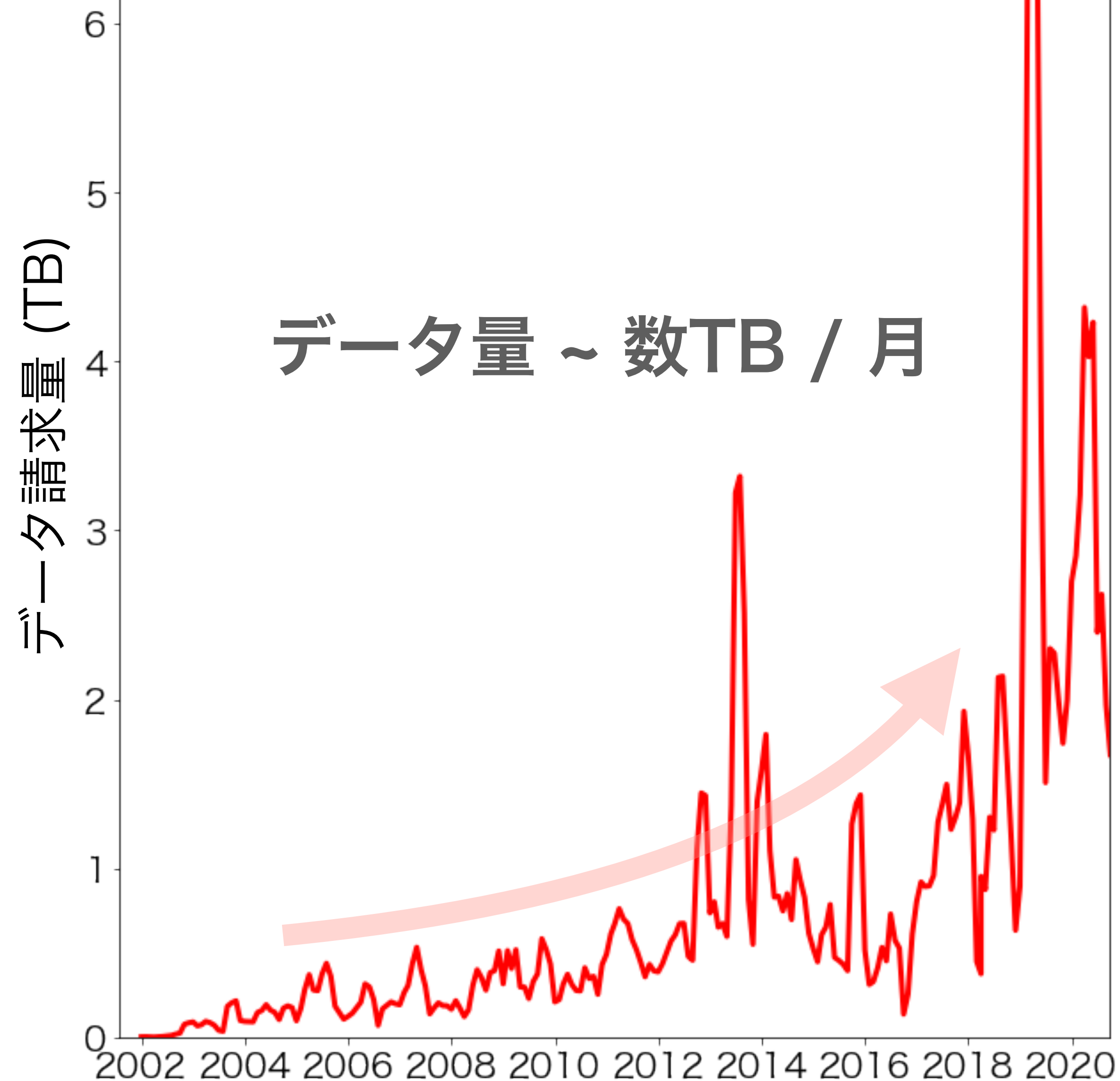
Instruments **Observation Mode** **Data Type** **Observation Category**

Observation Band **Filter lists / Wavelength**

ある日付のフレーム一覧から選択

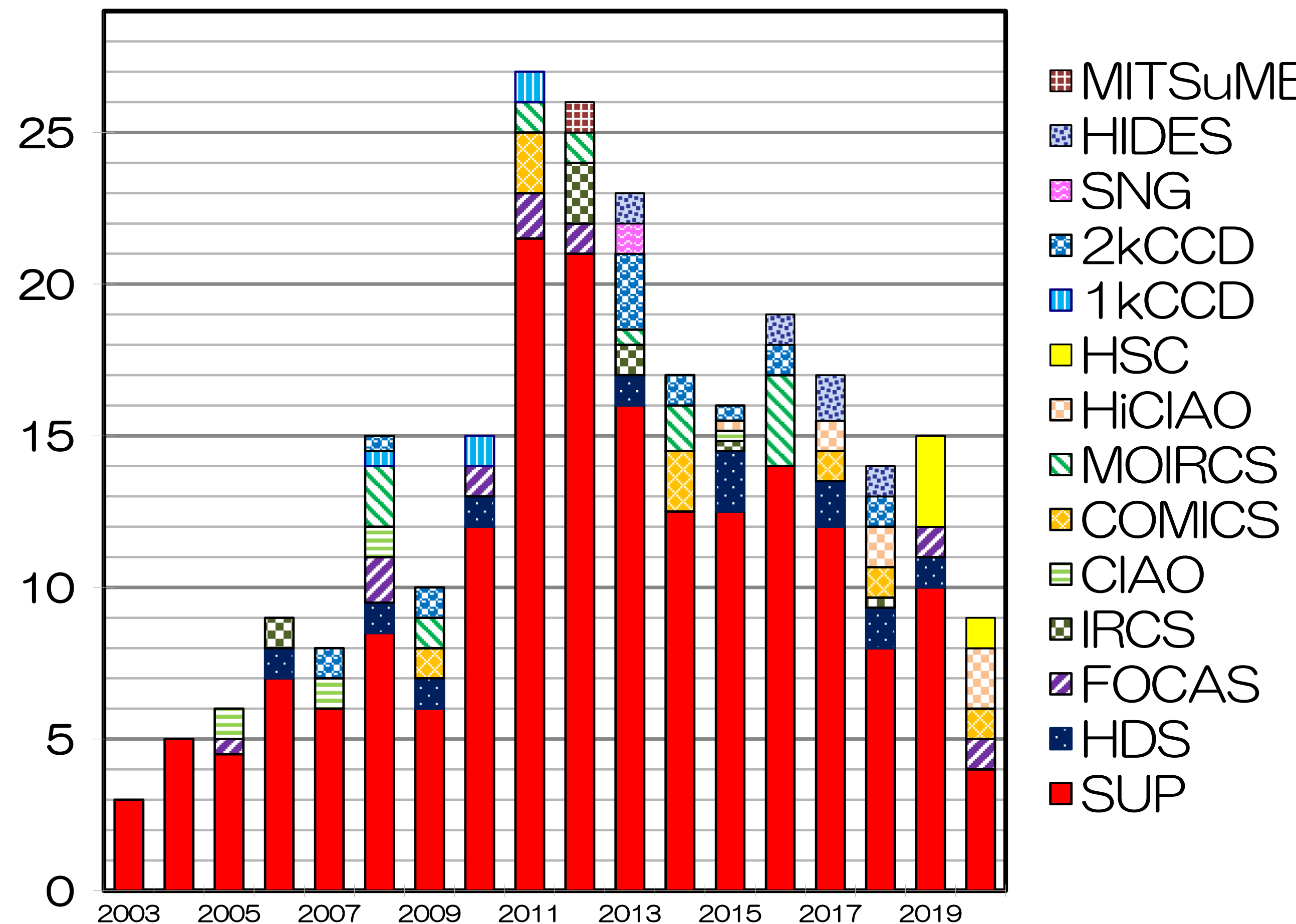
No.	Doubt	RawData	FRAMEID	DATE OBS	FITS SIZE	OBS
1		<input type="checkbox"/>	HSCA17172000	2019-04-01 weather	17	IMAG
2		<input type="checkbox"/>	HSCA17172001	2019-04-01 weather	17	IMAG
3		<input type="checkbox"/>	HSCA17172002	2019-04-01 weather	17	IMAG
4		<input type="checkbox"/>	HSCA17172003	2019-04-01 weather	17	IMAG
5		<input type="checkbox"/>	HSCA17172004	2019-04-01 weather	17	IMAG
6		<input type="checkbox"/>	HSCA17172005	2019-04-01 weather	17	IMAG
7		<input type="checkbox"/>	HSCA17172006	2019-04-01 weather	17	IMAG
8		<input type="checkbox"/>	HSCA17172007	2019-04-01 weather	17	IMAG
9		<input type="checkbox"/>	HSCA17172008	2019-04-01 weather	17	IMAG
10		<input type="checkbox"/>	HSCA17172009	2019-04-01 weather	17	IMAG
11		<input type="checkbox"/>	HSCA17172010	2019-04-01 weather	17	IMAG
12		<input type="checkbox"/>	HSCA17172011	2019-04-01 weather	17	IMAG
13		<input type="checkbox"/>	HSCA17172012	2019-04-	17	IMAG

利用状況



登録ユーザ数 : **206** (2020年11月現在)
(毎年更新)

総論文数 **254** (近年 ~15/yr)



SMOKAの目指すもの

データ公開することで

(1) 研究成果の検証を可能に

研究成果が画期的であればあるほど
第三者による検証は重要。

(2) 観測データの当初の提案とは異なった発想・目的や較正・
解析方法で新たな研究成果を創出

再利用・発展

近年、この考えが世界の科学研究コミュニティの間で広がっている

データ公開の潮流

(1) Data Availability Statement

論文中で報告する知見の解釈、再現と発展に必要な「最低限のデータセット」を第三者が利用可能であるか、可能であればどのように入手できるかの記述。

Nature、MNRAS、ApJ、AJ

A&A はデータ公開も要求

研究助成機関でも

(2) オープンサイエンス

政策レベルで研究データのオープン化が重視されている。

(2013年G8や2016年G7など)

「オープンサイエンスの深化と推進に向けて」 2020年5月28日 日本学術会議による提言

公開してみる？

自前でデータアーカイブシステムを構築・運用

個々の観測所で行うのはハードルが高い → ノウハウと人的リソース

SMOKAで公開

- SMOKAのインフラを利用してしまおう
- システム構築時にデータのチェックも行われる
- SMOKAは世界的に名が知られている。

公開にあたって必要な品質についてのノウハウがあります

公開してみる？

自前でデータアーカイブシステムを構築・運用

個々の観測所で行うのはハードルが高い → ノウハウと人的リソース

SMOKAで公開

- SMOKAのインフラを利用してしまおう
- システム構築時にデータのチェックも行われる
- SMOKAは世界的に名が知られている。

公開にあたって必要な品質についてのノウハウがあります

SMOKAで公開

- まずはSMOKAチームに連絡・相談 [smoka \[at\] smoka.nao.ac.jp](mailto:smoka@smoka.nao.ac.jp)
- データ転送方法、データ公開ポリシー、生データファイルの内容（特に FITSヘッダなどのメタデータ）などの確認の打ち合わせを重ねる
- 観測所と連携しながらSMOKAチームがシステム構築 -> 公開へ

SMOKA公開時に観測所に求められるもの

(1) 観測生データとともに、マシンリーダブルな形で、較正と解析に必要なメタデータ（時刻、フィルター、積分時間、赤経赤緯、観測モードなど、）の正しい情報を提供してください。

- 観測所が情報が**正しい**ことを確認



- 最も理想的なのはFITSヘッダに必要なメタデータが全て正しく入っていること

- そうでない場合も（手書きログではなく）**マシンリーダブルな形でリスト提供**

-> **SMOKA側でFITSヘッダに埋め込み**

- FITS規約違反のFITSファイル、極端な場合FITSファイルでなくてもOK

-> **SMOKA側で正しいFITSファイルに変換**

(2) どのデータをいつ公開するかを観測所側で管理

- 「どのデータをいつ公開するか」は観測所が決定

SMOKAはそれに従って公開

- 全てのデータを公開する必要なし
- 公開しない方が良いデータは非公開

試験観測データや不適切なデータ（読み取りエラーなど）など

- 公開フレームの公開開始日を設定し、そのリストを観測所側で作成・管理
- SMOKAがそのリストにアクセスできるように

（松：データベースにSMOKAからネットワークアクセス可能、

梅：メールでcsvを定期的に送付）

必須ではないが望ましいこと

(3) データ転送サーバを用意

- 観測データをネットワーク経由でSMOKAのサーバに転送
- 定期的にHDDなどのメディアを郵送もOK

(4) 観測環境データの提供

- 気象データや全天スカイモニタ画像など

まずはご相談ください

smoka [at] smoka.nao.ac.jp

SMOKAの枠を外れた相談も、、

- 自前で公開
- OISTER内部で共有