

STARS/MASTARS

2025.02.06

国立天文台天文データセンター
高田唯史、八木雅文、北田千博

STARS/MASTARSとは何か？

すばる望遠鏡の観測者向けの
観測データアーカイブシステム。

STARS=SubaruTelescopeARchiveSystem

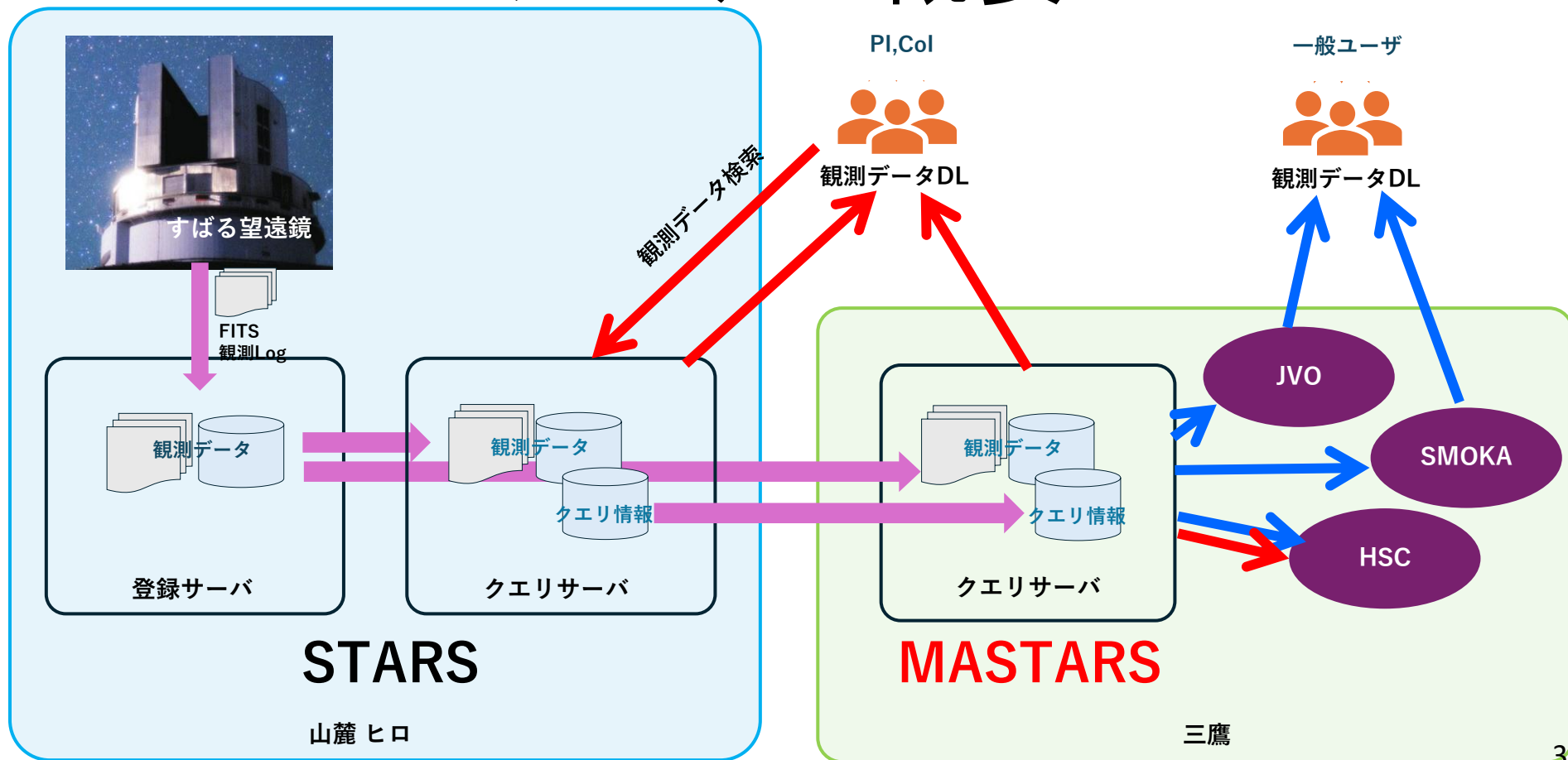
MASTARS=Mitaka Advanced STARS

 主に観測から通常 1 年半の占有期間の間に
観測者(PI,CoI)が生データを検索・ダウンロード
するためのシステム

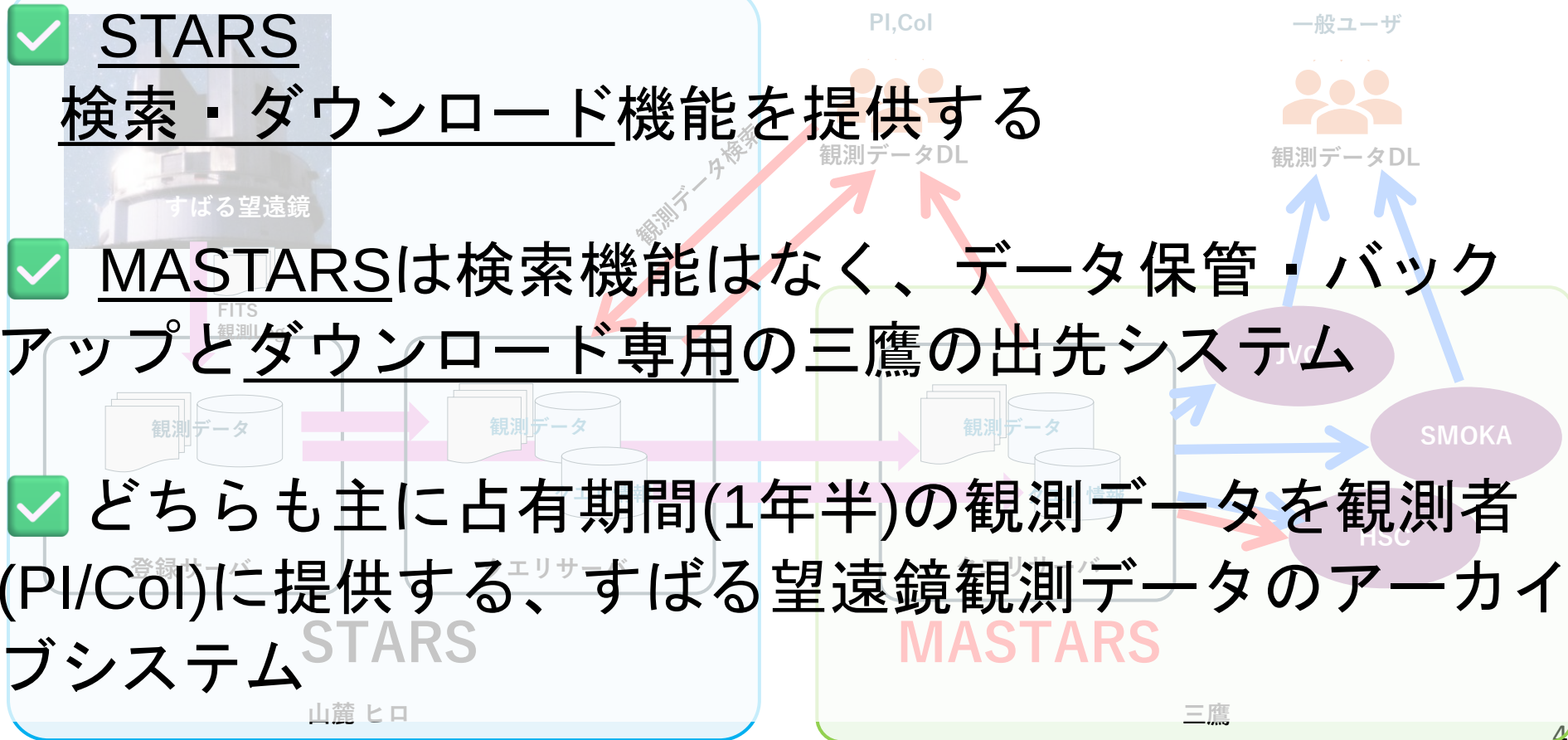
 すばるの生データを未来まで保管し続けることも重
要な役目

 望遠鏡や環境のログも運用開始時から
25年分すべて保管

システム概要



つまり...



なぜADC-UMで話すのか？

MASTARS は2022 年にハワイから
データセンターに移管されたためです。
今回(2024/07-)の新機器への移行も
データセンターが主体で行われました。

将来的にはSTARSの大部分の機能も含め
データセンターが開発する次期光赤外アーカイブに
統合・集約されていくことが想定されています。

MASTARS 新機器への移行

- 2024/02-06 にかけて前システム(2018-2024)のレンタル終了と RHEL7の終息(2024/06)のため、新レンタル機器へ移行。
- 前システムでテープバックアップしたデータは、新システムでは読むための有償ソフトウェア用の予算削減で古いテープは全部捨て、オンラインで新システムのディスクへコピー。
- 念のため一時的にアマゾンのクラウドストレージにバックアップをとっていた。とても大変だった。
- 現在新機器のテープと無償ソフトウェアで新規にバックアップを作成中。

MASTARSが停止する(した)のは

- ・ 年2回のメンテナンス (5月と11月)
- ・ 移行時の停電対応試験(2024/05/27)
- ・ 電源系トラブルに伴う停電(2024/07/30)
- ・ 次はUPSのメンテナンス (2025/2~3 のどこか1日)
- ・ キャンパスの法令電気点検(原則毎年11月の第二土曜)
金曜停止・翌月曜起動

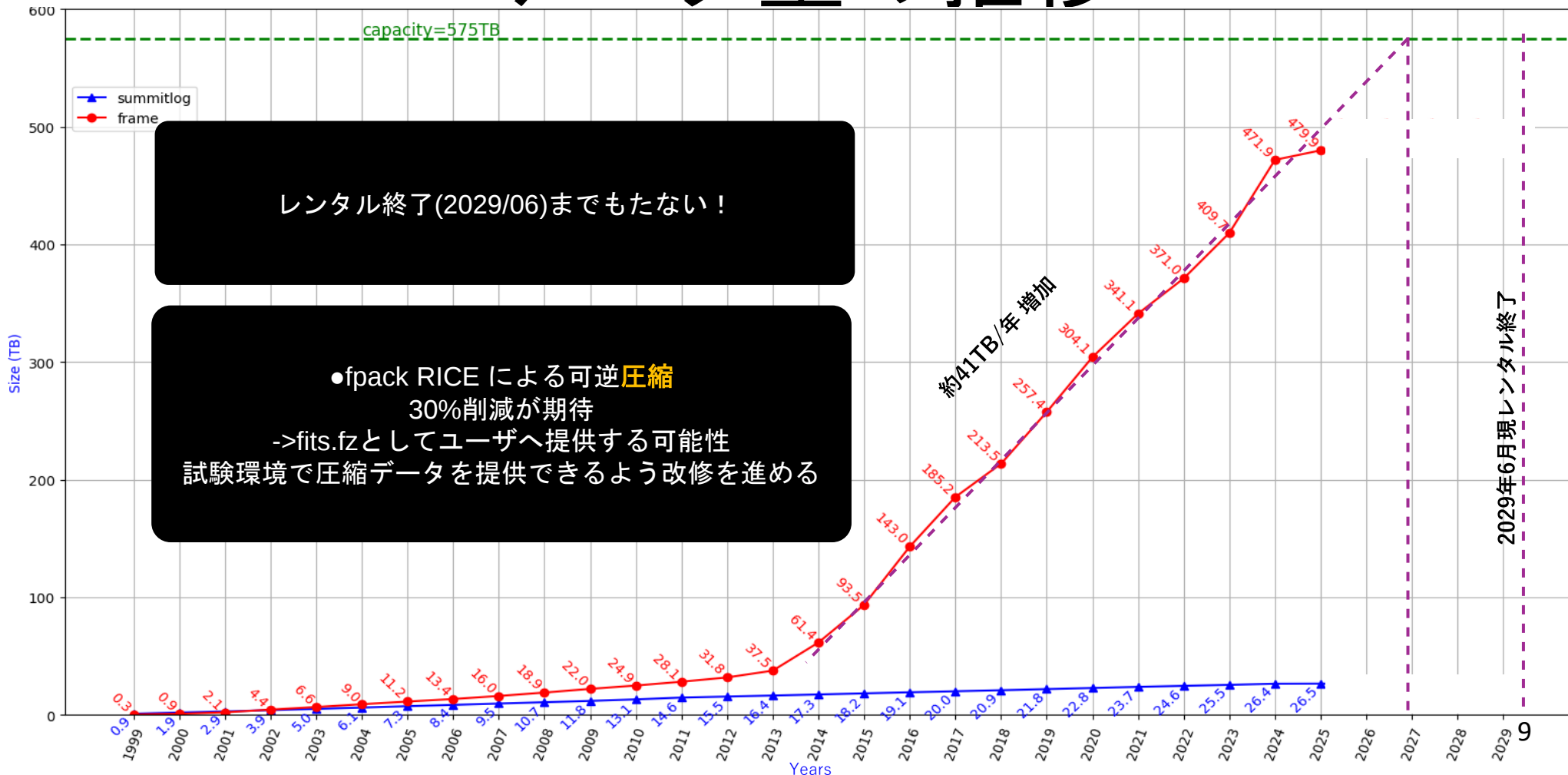
事故以外はSTARSのトップページで広報されるはず。

- ・ その他、天文台内・天文台上流のネットワークの不具合でアクセスできなくなることもある。

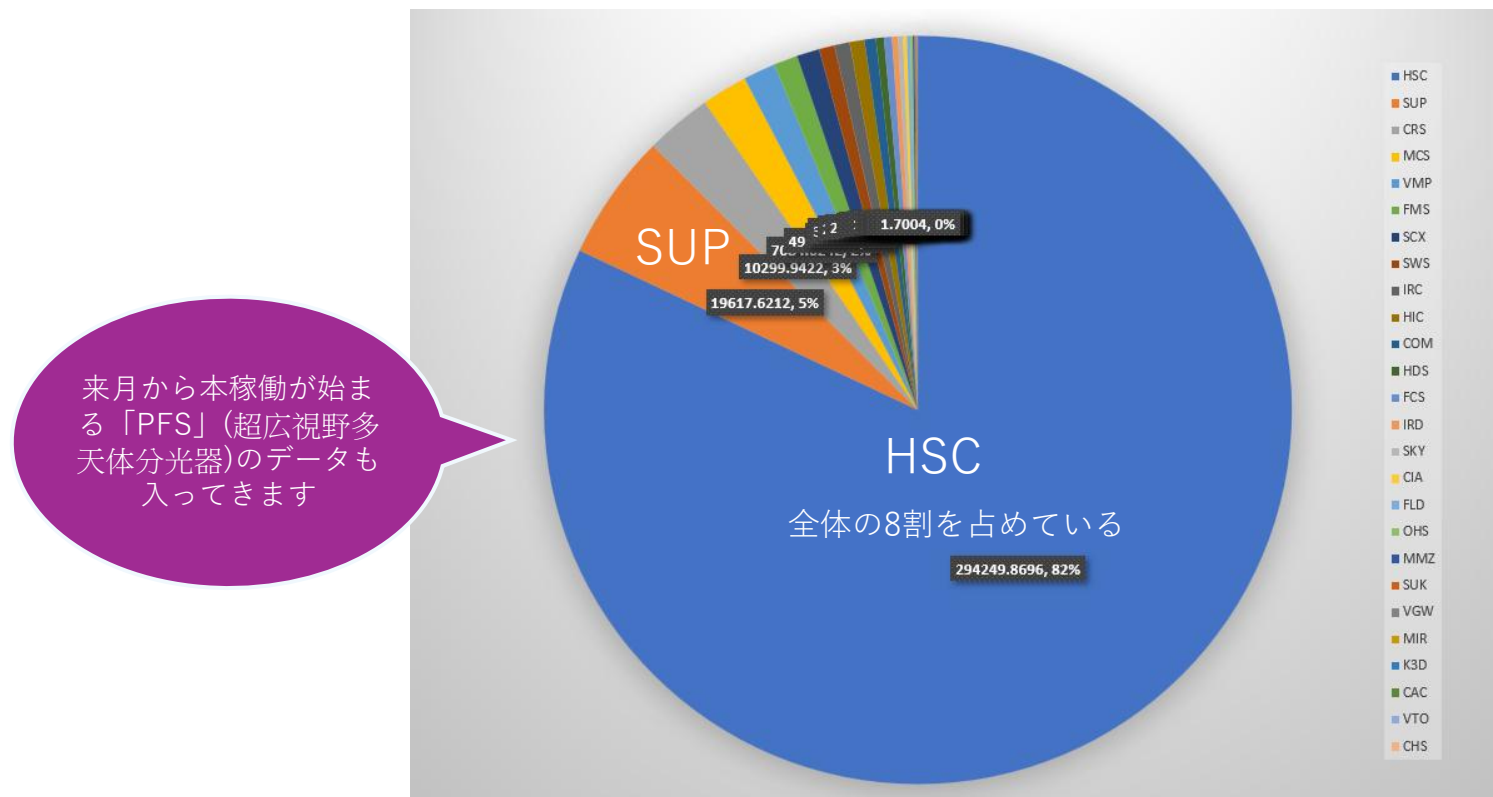
今保存しているデータと置き場

- ほぼすべての観測生データ、望遠鏡ログ(~500TB)
付随情報(野帳やマスクパターン、観測手順書など)
は保存していない。
- ハワイのオンライン、三鷹のオンライン
ハワイのテープバックアップは最新状態。
- 三鷹ではLTO-8 の LTFS にバックアップ中
- 移行時にアマゾンのクラウドに
バックアップしていたが、毎日課金されるので
テープにバックアップが完了したものは削除中。

データ量の推移



観測装置別FITSファイルサイズ占有率



利用実績と内訳

2023年(2023/01-12)での情報を調べると、ダウンロードされたファイル数約80万、データ量約15TB

アカウント数は検索の際に用いられたアカウントがダウンロードでも使われると仮定して 61人

うち、ダウンロード量の多かったトップ5アカウントで 8.5TB と全ダウンロード量の約半分

利用実績と内訳

アーカイブシステム別ダウンロード数

(2018/03 - 2024/06)

アーカイブシステム	ダウンロード数
STARS	7,345,672
MASTARS	14,792,179

MASTARSはSTARSの約2倍の利用数

ドメイン別ユーザ分布

日本、オーストラリア、カナダ、オランダ、チェコ、アメリカ、
スペイン、インド、台湾など

ハワイSTARSは2023年の一時期にシステム更新中だったので、
その期間に三鷹にアクセスが集中していた可能性はある

ユーザがダウンロードに使うツール

httpd のログに書かれている USER-AGENT は
2023年での情報を調べると以下の通り。

(※ただしUSER-AGENTには偽情報を書くこともできる)

- Wget 80.2%
- curl 8.2%
- lftp 6.1%
- Python-urllib 3.3%
- ブラウザ 2.1%

殆どのユーザがSTARSから自動配信されるダウンロードスクリプトを利用していると推定
(s2query.tar：ファイル一覧＋スクリプト)

STARSの検索結果から直接リンクをクリックしてブラウザでデータアクセスする人は少ない。1ユーザがMacのFirefoxで17000ファイルをダウンロードしていたが、その他は全部足して全アクセスの0.02%ほど

将来

今後数年、いくつかのタイミングで

「STARS3(次世代アーカイブ)について意見をください」
というお願いをしたいと思います。

小さなことでも構いませんので、声をお寄せください。
技術的あるいはポリシー的に難しいことでなければ
できるだけ実現するよう努力します。

将来

皆様の声を使い勝手のよい
次世代アーカイブシステムを作ります！