

次期光赤外(生データ)アーカイブ構想

Concept/Plan of The Next Generation Data Archive for
Optical/Infrared (raw) data of Japanese ground-based telescopes

高田唯史(国立天文台天文データセンター)

Tadafumi Takata(ADC/NAOJ)

構想に至った背景(Background)

- **予算の逼迫に伴うシステム効率化への対応(the need for system efficiency due to budget constraints)**
 - ✓ 各観測所の運用費縮小や人員体制の変化（特にデータ運用）
 - ✓ ストレージ（特にオンライン）容量をなるべく安く得る
 - ✓ 電気利用効率を高める（ストレージの共有、データ圧縮、省エネストレージデバイスの活用（例えばテープ）、データ多重化の見直し）
- **取り扱うデータの巨大化や複雑化に伴うデータ利用手法の変化(Changes in data usage methods due to the increasing size and complexity of the data handled)**
 - ✓ データの移動(コピー)が大きなコストとなる(時間、お金)
 - ✓ データが巨大すぎて自前の計算機にダウンロードできない
 - ✓ 大プロジェクトのついでに取り扱える規模ではないデータの増加
 - ✓ 収集した情報から新たな“気づき”を得るための手法の進化（AIに代表されるデータ科学的な取り組みへのより親和性の高い対応）
- **管理情報の統一を図る必要性(The need for uniform management information)**
 - ✓ 分散アーカイブ独自の情報管理の問題点
 - データ修正情報の矛盾（データアクセス権やデータ解析に直結する情報の場合には致命的になり得る）
 - ✓ エビデンス・サプライヤーの増大する重要性
- **単純に古くなってきたというのも理由**

目標 (The Goals)

- **次期(次世代)光赤外データ用アーカイブシステムの開発を行う。**
(Develop a next-generation optical and infrared data archive system.)
- **現在のすばる望遠鏡観測者用アーカイブ (STARS/MASTARS) の次世代版を開発し、SMOKA (もしくはその後継システム) との総合的な運用に向けた調整を行う。特にデータ運用モデルの設計に重点を置きながら進める。**
(We will develop the next generation of the current Subaru Telescope archives (STARS/MASTARS) and make adjustments for comprehensive operation with SMOKA (or its successor system), with a particular focus on designing a data operation model.)
- **最終的には統一的な光赤外データアーカイブが構築できれば良い**(It would be great to build a unified optical and infrared data archive.)

2024年度活動(Activities in FY2024)

- **オープンアクセス推進化予算による新アーカイブ試験開発用サーバ計算機システムの予算獲得（3月中旬までに順次導入）** (Obtaining budget for computer system aiming test development of new archives through OA promotion budget (to be gradually introduced by mid-March))
- **Proxyサーバーのテスト機も別途購入** (Proxy server test machine also purchased separately)
- **新アーカイブ試験開発用サーバ計算機（テープドライブ込み）とネット、ワークスイッチ、および無停電電源装置等を本年度中にセットアップし、来年度からの本格的な開発試験開始に向け準備を整える。** (The new archive test development server computer (including tape drive), network switch, and uninterruptible power supply will be set up within this fiscal year, and preparations will be made for the start of full-scale development testing from next fiscal year.)
- **CDR資料に向けた情報（今までの調査結果も含めた）の整理（まとめ）やアップデート（STARS/MASTARSに限定）** (Organizing (summarizing) and updating information (including past investigation results) for CDR documents (STARS/MASTARS only))

2025年度以降の活動計画

(Tentative activity plans for FY2025 and beyond)

- **STARS/MASTARS**については**現行のアーカイブシステムの機能を分析**(Regarding STARS/MASTARS, we analyzed the functions of the current archive system.)
- **将来の他アーカイブ（具体的にはSMOKA）との統合に向け、異なるDBシステム間の移行に向けて調査検討試験を行う。統合のレベルについても技術的・時間的な要素も考慮した計画を模索する**(In preparation for future integration with other archives (specifically SMOKA), we will conduct research and testing to transition between different DB systems. We will also explore plans for the level of integration, taking into account technical and time factors.)
- **ユーザ管理・アクセス制限の他プロジェクト（具体的にはまずはハワイ観測所）との責任分界点の確認と明確化を進める**(Confirm and clarify the division of responsibility with other projects (specifically, Subaru Telescope) regarding user management and access restrictions.)
- **アーキテクチャ（基本となるソフトウェア（開発言語やDBMS等）候補）の調査と選定**(Research and selection of architecture (candidates for the underlying software (development language, DBMS, etc.)))
- **開発 + Deployスケジュールの策定を行い、具体的な見通しをチーム内で共有する**(Formulate a development + deployment schedule and share specific forecasts within the team)

今後の予定（あくまで目論見）



- 2024/12/23 ADC内関係者によるキックオフ会議開催（Kick-Off Mtg on Dec 23 2024）
 - ✓ 今後の打合せの頻度(Frequency of future meetings)
 - ✓ 情報共有方法の決定(Deciding how to share information)
 - ✓ 次回は2025年3月を予定(Next Mtg in March 2025)
- **2025/1～ 概念設計開始（次期STARSに限定）** (Conceptual Designing limited to STARS/MASTARS)
- **FY2025にCDR(Conceptual Design Review)を開催予定**(Hold CDR in FY2025)
- **FY2026 プロトタイプ作成** (Proto-typing in FY2026)
- **FY2026EndまでにFDR(Final Design Review)を開催予定**(Hold FDR in FY2026)
- **FY2027本番ソフト作成、後半で試験運用開始**(Development and test operation in FY2027)
- **次回の計算機リプレイス時には新アーカイブを立ち上げて運用しておきたい**(We hope to set up and start operation of new archive sufficiently prior to next computer replacement(FY2029))
- **並行して次世代SMOKAとの併存もしくは統合についての本格的な検討や試験的な実装を行いたい**(In parallel, we would like to seriously consider coexistence or integration with the next generation SMOKA and conduct a trial implementation.)
- **どんな人員体制で臨む（臨める？）かは今後の検討課題**(The personnel structure that will be (or can be) put in place is an issue to be considered in the future.)