

すばる望遠鏡のデータベース

高 田 唯 史

〈国立天文台ハワイ観測所 650 North Aōhoku Place, Hilo, HI96720, U.S.A.〉

e-mail: takata@naoj.org

すばる望遠鏡の運用を支えるデータベースには様々なものがありますが、観測データを管理し、ユーザーに提供するためには様々な仕組みが用意されています。今回はこういった種類のデータをどうやって処理し、管理をしているかの現状を紹介するほか、ハワイと日本という離れた環境下でどうやってデータの一元的な管理を実現しているか、また、データベース天文学と呼ばれる大量データを用いた天文学の新しい取り組みに際して、すばるのデータをどのように生かして行くべきかなどについての我々の活動などを紹介したいと思います。

1. “すばる”を支える様々なデータベース

すばる望遠鏡は非常に大きくしかも複雑なシステムであり、出てくるデータも様々な場面、形態、利用方法がありますので、当然その運用を支えるために様々なデータベースが存在しています。観測データなどの管理についてはもちろんですが、その他にもオートガイド用の星を観測時に提供するため、USNO カタログ（約5億天体）の中から、ガイド用の星に適当であるかを示すパラメータを付加したカタログのデータベース¹⁾や、ファインディングチャートを提供するDSSの画像データベース、観測を実行するために必要な観測手順書を作成する場合に、それを補助するための標準星などのデータを集めたカタログデータベース、望遠鏡のポインティングアナリシスを行うときに必要な、非常に位置精度の高い星を集めたカタログデータベースなどのいわゆる天体カタログデータベース²⁾と呼ばれるものから、プロポーザルの管理や、観測者の計算機アカウントを管理しているデータベースのようなもの³⁾や、オペレーターのトラブルシューティング方法を集めて検索できるFATS (Fault Tracking System) と呼ばれるデータベース⁴⁾など、非常に様々なものがあります。

本原稿はデータベース天文学を目指す意味でのデータベースのあり方と現状の紹介についてが主目的ということなので、以後は観測データ及びそれに付随するデータを管理するためのデータベースに関連する事柄に絞って紹介することにします。

2. すばる望遠鏡のデータを管理するデータベース

すばる望遠鏡から生産される観測データ他を管理・ハンドルするためにSTARS (Subaru Telescope ARchive System)というシステムが国立天文台ハワイ観測所のヒロ山麓施設内で運用されています。また、これとほぼ同じ機能とデータ量を有するMASTARS (Mitaka Advanced STARS) が、国立天文台三鷹の解析研究棟内で運用されています。ここでは、これらのシステムで管理運用されているデータの種類やデータの管理方法、実際のデータの流れなどについて書いてみようと思います。その他に、公開用のデータについては国立天文台計算センターを中心とした有志の方々と共同で運用しているSMOKA (Subaru Mitaka Okayama Kiso Archive) システム⁵⁾が存在し、それぞれの役割を担っています。各システム間のつながりなどは図1に簡単に示してあります。

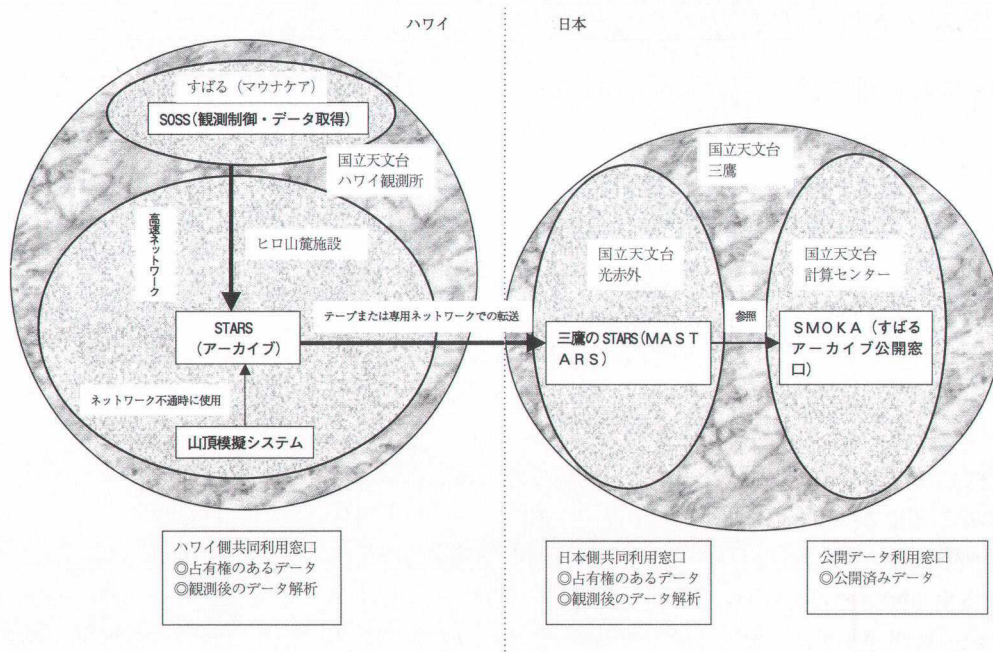


図1 すばる望遠鏡データに関連するアーカイブシステム構成図。基本的には、データはネットワークを介して流れる。マウナケア山頂と山麓施設間のネットワークが不通になったときはテープでデータをヒロまで運び、山頂模擬装置と呼ばれるシステムからデータを STARS に流し込む。

2.1 データの種類と現状のデータ容量

現在すばる望遠鏡において稼働している（または稼働した事のある）観測装置は10装置にのぼりますが、そのうちの9装置についてはデータはすばる望遠鏡のデータルールに従う形で FITS 化され、STARS に登録されて管理されています。それぞれの観測装置は表1に示すように様々な形式のデータを生産しますが、これらは全てすばる独自で定めたルールに従ったヘッダーキーワードを使っているため、その情報をうまく抜き出すことでデータベースで一元的に管理ができ、同じ枠組みで様々なデータが検索取得できるための基本となっています。

観測データから二次的に生産されるものとして、FITS のヘッダーの部分だけを切り取ったファイル (HDI: HeaDer Information file), より速くデータの性質などを知ることができるように、元データを20倍以上に圧縮して作られた QLI ファイル (Quick

Look Image file)⁶⁾なども同時に STARS 内で生産・保存され、管理されています。その他にも、主鏡鏡面の測定を行うためのシャックハルトマン (SH) 画像や、オフセットガイドに使われた星 (AG) の画像なども管理の対象となります。データセットと呼ばれる、将来的にデータ解析パイプラインに流し込むのに必要な情報を集めたファイルについても保存・管理も行っています⁷⁾。

このほかにも、地上望遠鏡では必要不可欠な環境情報の保存・管理が必要で、1日に数GBに及ぶ望遠鏡関連のステータスログファイル群や、スカイモニター (赤外線全天モニターカメラ) の画像なども保存管理されています。また、非常に頻繁に使われる環境データをログ群から切り出して日毎にまとめた簡易版ログも生産、管理を行っています。

このように管理運用するデータは多岐にわたり、データベースでの管理、検索用ユーザーインターフェースでの便利なデータ提供はすばる望遠鏡を運

用していく上で必要不可欠なものとなっています。

2.2 データの管理方法

観測データの管理は、必要な情報をデータ本体から抜き出すなどしてリレーショナルデータベースに登録しておき、データ本体については別途ハードディスク、もしくはそれと連動したテープライブラリに保存しておいて、データベースを検索することによってその保存場所がわかるようにしてあります。観測データについては、取得後1年半という占有期間内についてはそれぞれの観測を行ったグループメンバーのみがアクセスできるように、パスワードによる認証を通してデータにアクセスできるように設計されています。観測者は、STARS 検索用のユーザーインターフェース(図2)を介してデータを検索、取得できるため、自分の観測日時とその時に使われたプロポーザル ID と呼ばれる専用の番号さえ覚えていればデータは比較的簡単に手にすることができます。この UI は既に運用システムとして

活躍していた MOKA (Mitaka Okayama Kiso Archive)⁸⁾を原点としているので、比較的簡単に操作ができるような工夫も色々な場面で施されてい

図2 STARS のユーザーインターフェース画面の一部。検索条件を指定することで、山麓施設コンピュータシステム内の一時領域にデータを転送してことができ、その後テープに収めたり、ftp によってデータを持ち帰ることが可能となる。

ます。もちろん、そのユーザーインターフェースは、どのアドレスに WEB ブラウザを介してアクセスする必要があるかを知っておく必要はあります。例え

表1 2002年2月5日現在の、各観測装置のデータフォーマットとデータ生産量一覧。

装置名	装置の特徴	データフォーマット	ファイル数	ファイル総容量(GB)	1晩での最大生産量(GB)
OHS/CISCO	近赤外撮像分光	2次元	81,742	164.9	3.6
SuprimeCam	可視撮像	2次元	82,273	1284.6	28.2
COMICS	中間赤外撮像分光	4次元	41,784	127.8	1.7
IRCS	近赤外撮像分光	2次元	59,516	238.9	4.9
CIAO	近赤外撮像分光	2次元	13,420	49.1	4.3
FOCAS	可視撮像分光	2次元+ATE	7,530	67.2	2.1
HDS	可視分光	2次元+2つのATE	6,358	91.9	3.4
MIRTOS	近中間赤外撮像	3次元+ATE	11,342	48.8	5.3
CAC	可視光撮像	2次元	5,916	9.7	1.6
総量			309,881	2082.9	

※データ量は1998年末のファーストライト時からの積算で、テスト用の画像などは計算には含んでいない。

ば、ハワイ観測所内部でしたらば、現状では STARS ホームページ (<http://s03.naoj.org>) にアクセスする必要があります。管理の詳細などについては以下の文献などが参考になるでしょう^{2), 9)}。

2.3 データハンドリング (データの流れ)

データは山頂で取得された後、ヒロの山麓施設内のコンピュータシステムに高速ネットワークを通じて即座に転送され、データは瞬時に STARS のデータベースに登録され、検索取得が可能になります。さらに、1日1回の頻度でヒロ三鷹間の専用高速ネットワーク回線を介してデータは三鷹に転送され、約1日をかけて MASTARS に登録されます。これによって、同じデータを、ハワイと三鷹で保持でき、検索、取得も可能になります。実際、日本国内からアクセスする場合にはネットワークの速度の関係からも、MASTARS にアクセスする機会が多くなってきているようです。取得後1年半が過ぎたデータについては、基本的にアーカイブデータとして公開が可能になるので、それらについては SMOKA システムに情報が転送され、SMOKA より、観測者以外の人々にも検索・取得が可能になります。このように太平洋を挟んで2つのデータベースを分散化させているため、アクセスの分担によるロードの軽減や、2重化によるデータの安全

性の更なる確保などを行っているわけですが、反面、ヒロから三鷹またはその逆というように、管理情報の交換には手間がかかり、管理者にはそれなりの負担もかかります。自然に放っておけば良いというわけにはなかなか行かないのが問題といえは問題です。が、逆に言えばこれは機械や植木などと同じで、毎日丹念に様子を見てあげれば安定して動くようにもなるものです。結局は、こういった地道なあまり目立たない作業が明日のデータベース天文学を支えていく基盤となるのです。

3. データベース天文学へ向けての取り組み

データベース天文学を推進する上で、システムとして備えている必要があるのは以下のようなものだと思います。あくまでも理想的な事を書いていますが、使いやすいデータ、システムがこの分野の発展を大きく左右するのは事実です。この数年間で、世界的な規模でバーチャル観測所 (Virtual Observatory) 構想が進められ始めているのも、やっと“使える”データがたまりだし、これから巨大データをうまく使って、今までにない天文学のパラダイムを作り出そうとする動きの現れです。これらの動きにも目を向けながら、一体今後どうやって我々のシステムを発展させて行くかの早急な方向付け

が現在大きな課題となっています。

●提供されるデータが比較的誰にでも使いやすいものであること

データが処理済みで、測光情報及び位置、時刻といった、天文学において基本的と思われる情報を一目で見ることができる、もしくは簡単な手続きで知ることができるような有用なソフトウェアと一緒に提供されること。これには、データそのものに付随する形の情報以外に、例えば、地上観測の望遠鏡であるすばる望遠鏡などの場合には、データの取得中に周りの環境はどういった変化を見せたか(空の透明度は?シーイングは?)などの付加情報も必要になることでしょう。

●データの検索がユーザーの欲しいデータの選び方を提供できるようになっていること

難しいデータベース用の言語などは知らなくても、ある程度の天文学的な知識さえあれば、ユーザーが自分の欲しいデータを自分で定義できるよう、いわゆる独自の「データ検索モデル」を構築できることが重要で、これを使って、ユーザーが自分独自の観点でデータベースの世界から取得(観測)したサンプルを選んでこれる事が理想的な形といえるでしょう。現在日本でも SMOKA グループがこういった取り組みを行い始めています。将来的にはすばるに限らず様々な種類のデータがこのようなユーザーインターフェースで検索取得できる日も来るのではないかと期待できると思います。

●他のデータとの比較が簡単に行えるようになっていること

多波長のデータを単純にオーバープロットして使える機能から、カタログ同士の cross-correlation をとってみたり、やりたいサイエンスによって様々なデータの加工、表示方法が必要となります。全てをシステムで網羅するのは無理としても、上記のような基本的な機能を備えたシステムの整備は必要不可欠です。海外では NASA の Skyview¹⁰⁾ や CDS の Aladin¹¹⁾ などが有名ですが、日本でも jMAISON¹²⁾ がそのような機能を備えているシステムです。このよ

うな国際的な流通に乗せられるようなデータを生産できるかが、データベース天文学を展開できるかの鍵を握っています。

観測所のデータベースとしてまず第一に確立されねばいけないものが先の1番目にあげた項目です。観測所は自分の観測所の生産するデータについて責任を持ってその性質や有効性を証明し、その情報を公開して行かねばなりません。これらの情報は観測を立案する上でも不可欠な情報ですので、アーカイブデータを使う以前に既に必要となってくるものも多いのです。また同時に、安定して使えるデータを提供できるよう努力する責任もあります。このためには、常に装置や望遠鏡の状態を把握し、出力されるデータを日常的にある決まったルーティンを使って処理しながら、経年変化的な情報を見て行くなどの日々の努力と、それらを整理して蓄えていくいわゆる2次的な情報用のデータベースシステムの整備が必要です。ただためるだけでなく、うまく情報を使いこなせるような様々なソフトウェアとの融合の上にそういった機能は成り立つのです。すばるでも少しずつではありますがそういった試みが行われようとしています。まだまだハッブル宇宙望遠鏡や ESO の VLT のような本格運用システムにまでは発展できていません。我々に課せられた今後の課題の大きなものの1つです。そして、これらの責任を観測所として果たしていけるということが、利用できるアーカイブデータを生産していけることに直接つながります。これらの土台を積み重ねることで、観測装置や望遠鏡の癖などが除去できたいわゆるキャリブレーション済みデータが提供できるようになります。これらのキャリブレーションが、ある程度自動化され、人的な手をあまり介さなくても行えるレベルに達しないとなかなか「使いやすいアーカイブデータ」にはなっていきません。現在、我々はそういった土台作りの真っ最中といえるでしょう。そのための土台として、共通化された FITS フォーマットや、データセットの概念などを含めたデータのくくり出しの機能が実装さ

れてきているわけです。また、装置や望遠鏡関係者のたゆまぬ努力で、少しずつデータが整理され始めているわけです。先は長いかもしれませんが、すばる望遠鏡のスタッフはそういったたゆまぬ努力で日進月歩望遠鏡を進化させていっているわけです。2番目、3番目に書かれていることは、国立天文台の計算センターなどの色々なデータを集めているいわゆるデータセンターなどとうまくタイアップしながら作り上げて行くべきものです。HSTのアーカイブデータのように非常にうまくキャリブレーションされて、すぐに使える形でデータが提供されると、こういったツールの効果がありありと見えるようになります。観測所としてはその枠組みに乗せられるような付加情報を観測データと共に提供できるようにする努力が必要です。

2002年6月からはすばる望遠鏡の共同利用観測で取得されたデータも公開され始めます。8メートルクラス望遠鏡としては最大の視野を誇る主焦点カメラのデータなどは、様々な場面で利用される事になるでしょう。また、2002年度からは、観測所プロジェクトとして、1平方度に達するような非常に広く、しかもハッブルディープフィールドに匹敵するような非常に深い撮像データを取得するプロジェクトが走ることが決定し、これらのデータが迅速に公開がされる予定になってきています。こういったデータが観測者以外の人々にも色々なところで有効であって欲しいというのが、アーカイブシステムを作る側の希望ですし、今後もそういったデータ作りのための努力を怠らないようにしていかなければならないと思っています。もちろん、望遠鏡を使った観測によって世界に誇る成果を生み出すことが観測所の第一目標ですが、それを追求していく上で、自然と使えるデータを作り出せる仕組みを少しずつでも組み上げていくことができれば、非常に良い事だと思っています。

参考文献

- 1) 安田直樹 他, 2000, 「すばる望遠鏡ガイド星カタログの作成」, 国立天文台報 第4巻, p.195
- 2) Takata., et al., 1998, "Data archive and database system of the SUBARU Telescope", Proc. SPIE 3349, 247
- 3) Ogasawara R., et al., 2000, "Proposal management system of the Subaru Telescope", Proc. SPIE 4010, 1590
- 4) Winegar T., Noumaru J., 2000, "Subaru FATS (fault tracking system)", Proc. SPIE 4010, 136
- 5) 市川伸一, 2002, 天文月報 第95巻, p.266
- 6) Hamabe., et al., 2000, "New Image Quick-Look System for Subaru Telescope Data Archive", Proc. ADASS 9, 482
- 7) Kosugi., et al., 2000, "Data quality control for Subaru Telescope", Proc. SPIE 4010, 174
- 8) Horaguchi., et al., 1999, "An Astronomical Data Archive System with a Java-Based User Interface", PASJ 51, 693
- 9) Takata., et al., 2000, "STARS (Subaru Telescope archive system) for the effective return from Subaru Telescope", Proc. SPIE 4010, 181
- 10) <http://skyview.gsfc.nasa.gov/>
- 11) <http://aladin.u-strasbg.fr/>
- 12) Watanabe M., et al., 2001, "MAISON: A Web Service of Creating Composite Images On-the-fly for Pointing and Survey Observational Images", Proc ADASS 10, 70

Databases in Subaru Telescope

Tadafumi TAKATA

Subaru Telescope, National Astronomical Observatory of Japan, 650 North Aohoku Place, Hilo, Hawaii 96720, USA

Abstract: There are various types of database which are helping the operation of Subaru telescope, especially some useful functions are prepared for providing users the observational data. In this article, I will introduce the current status of data archive system of Subaru telescope, and how we are managing and handling the enormous amount of data between Hawaii and Japan. Our works and future plans for using Subaru data in coming "database astronomy era" will also be presented briefly.