

# 宇宙科学データ解析研究のためのバーチャル・センターの構築

文部省宇宙科学研究所 長瀬文昭

Abstract: The development of the Data ARchives and Transfer System (DARTS) for the space science research was started in 1995 by the PLAIN center (Center for Planning and Information system) of ISAS (The Institute of Space and Astronautical Science). Since 1997, a part of ISAS data, i.e., the X-ray astronomy database obtained with "ASCA", solar physics database obtained with "Yohkoh", and geophysical plasma database obtained with "Geotail", have been archived for public release. Since 1998, we have organized a "virtual data center" (DARTS-VDC) which is an association of space science related universities and institutes. Members of DARTS-VDC work together for the development of archival databases and analysis tools, and for the public release of their products, under the support of ACT-JST. At present in 2000, we have developed significant new archive databases from data of ISAS satellites, e.g., high resolution magnetic field database from Geotail, infrared image and catalogue databases from IRTS, and near-Earth field and particle database from Akebono. These databases are released to the public this year. We have also been developing data analysis software tools for the aid of research that is performed using DARTS databases. Furthermore, we are now constructing a mirror site and gathering the major international space science data on the basis of mutual data exchange for the research convenience of Japanese scientists.

## 1. はじめに

宇宙科学研究所では1993年に宇宙科学企画情報解析センター（PLAINセンター）が設立された。そして1995年より宇宙研の衛星観測事業で取得されたデータを一般研究者の利用に供するためのデータアーカイブシステムDARTS (Data ARchives and Transfer System) 開発計画を立ち上げた。1998年秋よりはこのDARTSを拡充すべく、宇宙研の科学衛星観測事業に参加している全国の大学・研究機関の協力を得て、科学技術振興事業団、計算科学技術活用型特定研究開発推進事業のプロジェクト、「宇宙科学データ解析研究のためのバーチャル・センター構築」(DARTS-VDCと略す)が始められた。このプロジェクトの骨子は宇宙科学関連機関の協力を得て、(1)宇宙研の衛星観測データを一般研究者が容易に利用出来るデータに変換したアーカイブデータベースを作成しDARTS-VDC傘下で管理する、(2)DARTS-VDC傘下のアーカイブデータベースを利用する研究者に、その解析支援ソフトウェア群を整備、提供する、(3)国内研究者の総合的解析の便宜を図るため、海外の宇宙科学に関する主要なデータベースを集約したミラーサイトをDARTS-VDC傘下に設置する、ことである。本計画以前に公開していた「あすか」のX線天文データベース、「ようこう」の太陽物理データベース、Geotailの磁気圏プラズマデータベースは本計画期間に一層改訂・強化された。それに加え(1)Geotail衛星磁場データベース、(2)「あけぼの」衛星地球磁気圏データベース、(3)IRTS赤外線天体カタログデータベース等が新たに構築され、DARTS-VDCを経て公開される運びとなった。また当初計画した「多波長観測データ可視化ツール開発」等の解析支援ツールの開発も順調にすすんでいる。さらにミラーサイトには太陽地球系物理学データベース(CDAWeb)、X線天文データベース(ROSAT全天サーベイデータ)が組み込まれ既にデータ中継サービスを開始した。

## 2. データベース開発・構築

### (1) X線天文データベース

本データベースでは、従来より提供している「あすか」衛星のデータアーカイブを"revision 2"と呼ばれる最新のデータセットで置き換えた新バージョンを平成12年8月15日に公開した。これにより約3年にわたって更新が滞っていたデータを全面的に改訂すると同時に、これまで公開していなかった平成8年夏以降のデータも提供できるようになった。Revision 2 データアーカイブは日本側が各機関の協力を得て製作した新たなキャリブレーション

ョン情報を基にして、NASA/GSFC と共同して作成したものである。旧版に比べると情報が整理統合され、「ディレクトリ単位でのダウンロード」機能の追加、ファイル選択メニューを階層化するなど使いやすくなっている。また、新データ導入と並行して検索システムも調整し、ユーザーインターフェースを改善した。さらに、新着データの更新作業の大部分を自動化して現在は毎日データを更新している。本データベースの次期バージョンでは複数衛星のデータアーカイブを同時に扱うことを目標としており、手はじめに BeppoSAX 衛星のデータアーカイブの組み込み作業を行なっている。改修の本質的な部分は終了しているので、平成12年中にはユーザーインターフェースを整えて公開できる見通しである。当初「あすか」の次に本データベースに組み込む計画であった宇宙研の「ぎんが」衛星データに関しては、既存アーカイブの公開には問題が多いことが分かり、現在新版アーカイブデータを作成中で平成13年中には上記システムに組込まれる予定である。

#### (2) 「ようこう」太陽物理データベース - 新インターフェース「Solar Theater」の開発

我々の星・太陽は活力に満ちた星である。中でも、ようこう衛星が観測するX線領域では、種々の時間スケール、空間スケールの変貌をまのあたりできる。この営みを解き明かすためには、太陽が放つ多次元の情報といかに向かい合うかが鍵となる。これまで、ようこうでは衛星のテレメトリデータ及び、解析ツール [商用ソフトIDLの上で使用] を配布という形でデータの公開を行ってきた。この方式では、一枚の絵をみることにさえ、幾つものステップを要していた。そこで新たに、Webを通じ、画像をムービーなど種々の対話型形式でブラウズするためのシステムを構築している。ユーザーインターフェース (UI) は JAVA の applet として作成し、標準のインターフェース (JDBC) を通じデータベースにアクセスする [構築済み]。コンテンツは FITS 形式などの画像ファイルと、観測条件を検索するための SQL Server として用意する [準備中]。長期的には JAVA のオブジェクト指向性、JDBC が有する分散性を活かし、各種観測データを融合する分散型データベースへの統合 UI へ育てていく。

#### (3) GEOTAIL 磁気圏プラズマデータベース

GEOTAIL 衛星は磁気圏の電磁気・プラズマ環境を計測する衛星であり、プラズマ粒子の観測データは磁場データと並び中核的な要素を占める。このデータの一般公開用データベースの構築は、プラズマモーメント (即ち、密度、速度、温度) の12秒値データについては低時間分解能 (3秒値) の磁場データとともに1998年秋にすでに完成し公開が開始されている。本プロジェクトにおいては、観測から得られたプラズマ3次元速度分布関数データの一般公開用データベースの作成を目指している。これは世界でも初めての試みであり、宇宙プラズマ科学の発展に大きく寄与できるものと考えている。具体的な3次元分布関数の可視化技術、プログラム開発を東京大学で担当し、宇宙研においてはデータの配信システムを構築する。現在、初期的なデータ配信システムに対応するデータベース構築がほぼ終了している。プロジェクト後半期間では東京大学で開発された可視化プログラムとのインターフェース部分の開発を行うことになる。これまでに、JAVA等のネットワークベースのプログラミング言語を用いた実験を行っているが、分布関数のデータサイズは大きくネットワーク経由で対話型の可視化プログラムをユーザに実用的に提供できるかどうかなお検討の余地がある。しかし、この問題は将来のネットワーク環境の向上によって自然に解消されることでもあるので、当初の目的通りプロジェクト期間中には (少なくとも) 最低限の情報を持った分布関数データベースの公開をスタートさせたい。

#### (4) Geotail 高時間分解能磁場データベース

Geotail に搭載されている磁力計 (MGF) は、1992年の打ち上げ以来、すでに8年にわたってほぼ全期間について観測を継続している。この磁力計は、磁場3成分を1秒間に16サンプル (16 Hz) という計測をおこない、地球の磁気圏内および太陽風での磁場データを提供してきた。衛星からのデータは、各種の補正をおこなったのち、研究にそのまま使えるように座標変換をほどこし、全期間にわたって一般に公開している。この物理量変換過程について見直しを行い、すべての処理が順調に進むように解析システムを構築した。磁場のデータは、3秒平均値と16 Hzの値の両方を公開できる体制とした。特に、3秒値の磁場データについては、web上でのオープンな

公開により、国内の研究者はもとより、外国の研究者にも簡単に利用できるようにしている。本格的な研究のみならず、この3秒値の磁場データは、地球近傍での太陽風のモニターとして国内外の研究者に広く利用されている。一方、16 Hzの磁場データについては、このデータを使った研究を進めながら、データの品質そのものの詳しい解析をおこなった。そして、高時間分解能データを用いて初めて解析の対象となる、地球定在衝撃波面や地球磁気圏尾部でのリコネクション領域の磁場構造とそれに伴うプラズマの振舞いの研究に利用されるようになってきた。

#### (5) 「あけぼの」地球磁気圏データベース

今年9月よりオーロラ・地球磁気圏観測衛星「あけぼの」のデータ公開を開始致した。現在公開しているデータは以下の通りである。1) 衛星の基礎情報：これは関係者連絡先・論文リスト・衛星情報など、あけぼの衛星データを利用する上で必要な一般情報である。2) 衛星軌道情報：これは各観測装置のON/OFF情報を含んでおり、条件設定によるデータ検索が可能である。3) 観測データ：地球・惑星系の衛星は、一般に国内外の多数のグループが複数の観測装置を独立に搭載する形となっており、従来データ公開ポリシーが観測装置によって異なっていた。DARTSでは、公開の条件を「公表時の事前連絡」・「適切な謝辞」・「別刷送付」の3点に留め、全機器のデータを一般の研究者に自由に使用可能とする。これまでに「低エネルギー粒子計測器(LEP)」、「オーロラ撮像装置(ATV)」、「放射線モニター(RDM)」の公開を実施、他の6機器も公開準備が出来たものから漸次追加予定である。1989年の打ち上げ以来、1太陽活動周期(約11年)を生き抜いてきたあけぼの衛星のデータが、更に広く使われることを期待している。

#### (6) IRTS 赤外線天文データベース

IRTSは、1995年3月に打上げられた日本初のスペース赤外線望遠鏡である。焦点面には赤外線の全域をカバーする4つの観測機器が搭載されており、全天の7%の天域の赤外線サーベイ観測を行った。本プロジェクトでは、IRTSで得られた観測データのデータベース化とDARTS上で公開するため独自の検索システムの開発・整備を行っている。中心となるデータベースは全波長域の画像イメージと近中間赤外点源カタログである。画像データは掃天領域を88の区画(1区画あたり約13度角)に分割した空間強度分布である。ブラウザ上では、任意の領域において、任意の観測機器・チャンネルの強度分布が表示され、異なる波長との比較を容易に行うことができる。近中間赤外点源カタログは、1.4-11.7ミクロンの波長帯における、56チャンネル分光スペクトルデータで、ブラウザ上では、天体リスト、天体固有名、座標のいずれからでも検索が可能となっている。データは世界的にユニークで価値の高いものであり、アーカイブの結果、国内外の天文研究者に広く利用されると期待される。

### 3. プログラム開発・製作

#### (1) 多波長天文画像データ検索・閲覧サービス「MAISON」の開発

宇宙科学研究所/宇宙科学企画情報解析センターと国立天文台/天文学データ解析センターでは、WWW上で稼動する「多波長天文画像データ検索・閲覧サービス」の共同開発を行っている。このサービスは、ユーザの指定に基づいて、Web上で稼動する独立した画像サーバに接続し、それぞれから画像データを取得して、オンラインでそれらの重ね合わせ表示を行なう、というサービスである。これにより、ユーザはWeb上で簡単に複数の画像の眼視比較を行うことが可能となる。類似のシステムはアメリカ・NASAのSkyViewシステムやフランス・CDSのALADINシステムがあるが、MAISONは、サーベイ観測画像データのみではなくポインティング観測画像データにも対応する、という独自の特長をもつ。このことは、すばる望遠鏡やASTRO-E2、ASTRO-Fなどの、日本の来るべき高品質ポインティング観測データを最大限に活用するために極めて有用である。開発第一版は既に公開されており(<http://maison.isas.ac.jp>)、この版ではX線天文衛星ASCAのポインティング観測画像、DSS(可視光)、IRAS(遠赤外)、FIRST及びGreenbank(電波)のサーベイ観測画像、が利用できる。また、眼視比較を行う上での機能を更に強化し、更にカタログデータのプロットなどの新機能を追加した新しい「MAISON」の開発を平成12年度から開始している。

### (2) ジオテイル衛星・プラズマ速度データ可視化システム

ジオテイル衛星・プラズマ速度分布関数の可視化システムの作成について、東京大学理学部と宇宙研が共同した種々の試行の結果、全国のユーザとの共用には分散型システムが最適と結論した。それによれば、宇宙研に中央システムを置き、ネットワーク経由でのユーザからのデータリクエストに応じてデータを管理・配信する。そして、実際の解析作業を支援するためユーザの手元にデータ解析処理ソフトを配布して東京大学では配布ソフトの雛型作成を担当し、Kylixシステムへの移植作業を行い、その大半を終えている。作成する可視化ソフトは特定の衛星観測に依存せず汎用であるべきだが、粒子データはその特性上、観測装置の個性を強く反映するため、データ入力部分の完全汎用化は未だ現実的ではない。しかし新たに開発したソフトウェア・パッケージは環境に合わせた書き直しが容易であり、多様な観測装置からのデータ形式にすみやかに適応させることができる。適応能力の試験のため、最近、急遽再解析が必要となった15年前の「すいせい」ハレー探査機からのイオン粒子データの解析ソフトの再開発(かつてのソフトはそれを搭載したハードウェア自体が消滅しており再利用不可能)を試みた。その結果、ソフトを2週間の間に完成することができ、移植性の高さを実証することができた。

### (3) ジオテイル衛星・プラズマ波動可視化ソフトの開発

ジオテイル衛星の波動観測器には、SFAと呼ばれるスペクトル受信機によるダイナミックスペクトル表示で、先ず観測された波動の概要を知ることが行われる。全体で5つチャンネルから成り立つが、その高周波部の3つのバンドのダイナミックスペクトルを衛星の位置とともに表示できるソフトウェアをIDLで開発した。衛星のデータをあけぼの衛星の公開データベースのPWS、VLF波動観測装置と同様の形式のファイルを作成した上で表示するようになっており、共通部分を多くしている。もう一つは、WFCと呼ばれる波形捕捉受信機である。これは電磁波の5成分を受信し、スペクトルのみならず、偏波や波形に関する情報が得られるが、データを見るのが容易ではなかった。これらを可視化するソフトウェアが作られ、源波形を見た上で、観測時間を設定し、波形、スペクトル、ダイナミックスペクトル、偏波などを会話型で表示でき、柔軟にデータ解析が行えるようになり、飛躍的に能率が向上した。

### (4) 線バースト速報・光同定システムの開発

理化学研究所では、今年度10月に米仏の研究機関と協力して線バースト探査衛星HETE2を打ち上げた。現在、HETE2により得られた線バーストの発生位置等の情報をいち早く全世界に通達し、また、美星天文台等の国内の天文台で即座に可視光追観測が自動で行なえるよう地上システムを開発中である。また1年以内に、宇宙研PLAINセンター等の協力を得て、線バーストが発生した位置のDDSの光学画像と、先述の国内の天文台で線バースト発生直後に得られた光学画像をインターネットで即座に公開するシステムを開発する。

## 4. ミラーサイトの構築

### (1) CDAWeb太陽地球系物理学データベース

CDAWebとは"Coordinated Data Analysis (Workshop) Web"の略称でNASAのデータセンターで運営されている太陽地球系物理学データベースである。これはNASA(米国)、ESA(欧州)、IKI(ロシア)、ISAS(日本)の4宇宙機関による国際的なプロジェクト(ISTP等)をサポートすることを目的に始められたデータベースで、このデータベースから世界各国の太陽地球系物理学に関連する人工衛星や地上観測データ等のデータセットが公開されている。データベースはすべての公開データ(key parameter)を含んでおり、例えばWIND衛星やPOLAR衛星などのISTP衛星群のデータが簡単に閲覧できる。データセットは共通データフォーマット(CDF)で作成されており、これらのデータセットから複数の観測項目に渡って対話的にプロットやダウンロードができるような機能がある。本プロジェクトでは太陽地球系物理学関連のデータとしてGEOTAIL衛星の高時間分解データを提供しているが、CDAWebのミラーサイト運営との連携によって国内外の研究者に大きな利便性をもたらすことが期待できる。本ミラーサイトは2000年6月から実運用が開始され、国内の研究者からのアクセスを中心に利用されている。

## (2) ROSAT/BeppoSAX - X線天文データベース

ローサットは、1990年から、1999年まで活躍した、X線天文衛星である。このうち、1990年7月30日から6ヶ月の間、全天走査観測が行われた。今年3月に、このデータがマックスプランク研究所より一般公開された。X線波長領域で、約1分角というこれまでにない、細かい画像分解能で観測されたデータは、非常に利用価値が高く、X線天文の分野で、これまで多くの成果を挙げてきた日本の研究者に有用であることから、ミラーリングの許可をとり、以下のアドレスにミラーサイトを構築した (<http://www.darts.isas.ac.jp/indexj.html>)。プログラムを日本語環境で動かすための変更や、当サイトのセキュリティポリシーに合うように環境を設定した上で、今年9月よりミラーサイトの公開を始めた。特に日本人の研究者が利用しやすいように、トップページとヘルプページは日本語ページも作成した。また他のミラーサイトとしては、前に述べたように、X線天文衛星 BeppoSAX のデータについて、現在ミラーサイトを構築中である。

## 5. ネットワークの活用について

本プロジェクトではネットワークを介して国内(国外)各機関を結び、全体として1つの大きなバーチャルデータセンターを構築することが目標である。宇宙研の科学衛星プロジェクトは国内外の多くの大学・研究機関が共同してその製作・試験・運用・観測を担当する。それぞれの担当した機器の観測データの編集・校正や観測データを総合した科学的解析等のために、大量の衛星観測データセットがネットワークを通じて関連機関と宇宙研の間を往復伝送される。一方、近年衛星観測機器の高性能・高機能化にともない観測データの量は非常に大きくなり、また、学問の深化とともに複数の関連データを総合的に解析したいという研究者の要望はますます強くなり、データベースの集中管理が必要となってきた。したがって、本プロジェクトにおけるネットワークの利用形態としては、一般ユーザへのデータ配信サービスとともに、衛星プロジェクトチーム内での物理データ作成作業に利用する内部的なデータ交換作業が重要な位置を占めることになる。多くのプロジェクトにおいてはこの作業は海外機関との伝送を含んでおり、ネットワークの利用は作業の効率化の点で大きなメリットが生じる。幸い中心機関である宇宙研が数年前に SINET のノード校化され、国内の機関大学との通信回線は高速化されている。また、SINET の日米回線が増強されたことにより宇宙研 - NASA 間のデータ伝送も実用的な伝送速度でデータのやり取りができるようになっている。本プロジェクトで既に公開を開始している衛星データのデータベースは、現在ではほとんど総てのデータがデータベースの構築段階から一般研究者によるアーカイブデータの利用までネットワーク経由で配送されており、ネットワークを活用したバーチャルデータセンターとしての機能はすでに有効に稼働していると思われる。

## 6. まとめ及び今後の方向性

本プロジェクトは順調に当初計画を遂行してきている。宇宙研の DARTS アーカイブデータベースシステムも今回のバーチャル・データセンター・プロジェクトによる開発研究により大幅に増強され、DARTS-VDC は日本の宇宙科学のデータセンターと称するに恥じないレベルに達したと考えられる。今後当プロジェクトの残余期間中に、(1)「ぎんが」、「はるか」など宇宙研衛星で取得したデータでアーカイブ化の進んでいないものにつて順次科学データベースを構築し、公開していく、(2)現在開発中の磁気圏プラズマデータ可視化ツールの公開を始め、磁気圏波動データの可視化ツールの開発、X線全天モニターデータ処理システムの開発等 DARTS アーカイブデータを利用するための DARTS システム独自解析支援ツールの開発を行う、(3)利用価値の高い外国の宇宙科学データベース (Einstein, BeppoSAX X線天文データベース、Trace 太陽紫外線データベースなど) のミラーリングをさらに推進する、ことを目標とする。さらに、本プロジェクト終了後もこのバーチャル・センターの理念・構想を継承し DARTS-VDC システムを継続的に運用していく方策をたてる。また、Astro-F 衛星以降数 10 Tb に達すると予想される宇宙研の大容量衛星観測資源を効率的に校正・編集し迅速に公開サービスを可能とするための準備を進める。これには、高速入出力の大容量データ蓄積装置の導入、所内 LAN、国内回線、海外回線を包括する高速通信網の end-to-end の整備が今後の課題となろう。

## 7. 研究実施体制

本研究は宇宙科学研究所宇宙科学企画情報解析センター（代表：長瀬）及び、東大理学部（共同責任：寺沢）、東工大理学部（同：長井）、名大理学部（同：芝井）、京大宙空電波（同：橋本）、理化学研究所（同：河合）、国立天文台（同：近田）が中心となって研究開発に当たっている。その他に宇宙研及び国内外の宇宙科学関連大学・研究機関で構成される各衛星プロジェクトの実験・運用班には較正された観測データや解析ソフトの提供を始め多大な協力を得ている。さらに海外からは、米国NASA /Goddard Space Flight Center, ドイツのマックス・プランク研究所、イタリアのBeppoSAX Science Data Center の協力を得ている。科学技術振興事業団で雇用され各機関に依託された若手研究員は、本研究開発プロジェクトの推進に当たって中心的な役割を果たし、非常に大きな貢献をしている。