

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

文部科学省宇宙科学研究所 ○長瀬 文昭

Development of a Virtual Data Center for the Research of Space Science

Fumiaki Nagase (The institute of Space and Astronautical Science)

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"ASCA", solar physics database obtained with "Yohkoh", and geophysical plasma database obtained with "Geotail", has been archived for public release. Since 1998, we have organized "virtual data center" (DARTS-VDC) which is an association of space science related universities and institutes. Members of DARTS-VDC worked 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 September 2001, we have developed significant new archive databases from data of ISAS satellites, e.g., high time-resolution magnetic field database from "Geotail", infrared image and catalogue databases from "IRTS", and near-Earth field and particle database from "Akebono". These databases have been released to the public use. We have also developed data analysis software tools for the aid of research that are performed using DARTS databases. Furthermore, we have now constructed a mirror site and are now 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を構築し、そこには現在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/BeppoSAX）等が組み込まれ、既に宇宙科学データベース中継サービスを開始している。

2. データベース開発・構築

宇宙研の衛星で取得され、各機関で公開データベースとして開発されたアーカイブデータは宇宙研のDARTSデータアーカイブシステムで一元管理をされる。ここにある宇宙科学データベースを利用しようとする研究者は

インターネット経由で URL:<http://www.darts.isas.ac.jp> にアクセスする事により、その傘下にある全てのアーカイブデータと解析支援ツールを無料で自由に利用する事ができる。現在宇宙研が所有する過去の衛星観測データ資源は約3.5Tbであり、このうち約2.5Tbが科学データベースとして編集公開されている。最近1年間のDARTS全体での総アクセス件数は1日平均1100件である。科学研究目的でのアクセスはこの一部であろうが、科学研究用データベースとしては良く利用されていると言えよう。以下に各データベース開発状況を簡単に述べる。

(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。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、更新を行ってきた。この更新作業は、2001年10月5日に最後のデータが公開された時点で終了し、あすかで観測されたデータ全てが、本データベースへ組み込まれた。また、本データベースには、BeppoSAX 衛星のデータアーカイブも組み込まれ、2001年1月から公開を行っている。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、プロジェクトチームにより、新版アーカイブデータの作成が行われた。データ部分については、処理が終了したことから、本データベース内にコピーを置いた。しかし、現在解析ソフトがプロジェクトチームにより改定中であることから、これの完成を待ってデータを公開する。

(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目の当りできる。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ザするためのシステムを構築した。このシステムではユーザーインターフェース (UI) として JAVA の applet を用い、データベースのアクセスには標準のプロトコル (JDBC) を用いる。これを用いると、世界各地に分散する太陽観測データを統合的に活用するシステムの構築が、容易に可能である。

(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中時間分解能 (3秒平均値) 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している。現在、最低限の基礎情報を持った分布関数データベースの公開を始めている。

(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

2000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した。現在公開しているデータは、1) 衛星の基礎情報：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、2) 衛星軌道情報：各観測装置のON/OFF情報（各種条件設定によるデータ有無の検索が可能）、3) 観測データ：全機器の統一データ、である。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）」「熱的電子計測装置（TED）」など5機器の公開を実施、他の4機器も公開準備が出来たものから漸次追加予定である。本衛星は1989年の打ち上げ以来、1太陽活動周期（約11年）の長期に渡り、同一観測装置で取得した貴重なデータを提供している。

(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・整備を行った。中心となるデータベースは赤外線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図である。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。このデータは世界的にユニークで価値の高いものであり、今後国内外の天文研究者に広く利用されると期待される。

(7) 「はるか」電波天文学データベース

「はるか」は、1997年に宇宙研で打ち上げた電波天文衛星で、世界中の電波望遠鏡と協力してこれまでに600近くの天体が観測された。このデータを、プロジェクトチームの許可を得て、本データベースを利用して公開することになった。この「はるか」のデータを組み込んだ検索システムの構築と、データリストの更新プログラムの作成は終了した。今後プロジェクトチームが、データを指定場所に置けば、自動的に検索システムに組み込まれるようになっておりデータ量がある程度そろった時点で、データ公開を開始する。

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・カタログサーバに接続し、それぞれから画像データやカタログデータを取得して、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や画像とカタログ位置との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「MAISON」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、の重ね合わせ表示機能を提供した。引き続き開発さ

れた第二版「jMAISON」では、Java言語のアプリレット機能を活用した開発により、眼視比較を行う上での機能強化やカタログデータプロットなどの新機能追加を行なったサービスが提供されている (<http://maison.isas.ac.jp>)。

(2) Geotail衛星・プラズマ速度データ可視化システム

Geotail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する。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。

(3) Geotail衛星・プラズマ波動可視化ソフトの開発

Geotail衛星の波動観測器には、SFAと呼ばれるスペクトル受信機があり、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、ホームページから会話型で日時などを指定して表示できるシステムを、データ公開に向けて開発した。観測開始、終了時刻、表示するバンド、出力形式などを指定すると、衛星の位置とともに表示できる。さらにWFCと呼ばれる波形捕捉受信機がある。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。データを見るのが容易ではなかったが、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った。

(4) γ 線バースト地上速報システムの開発

理化学研究所 宇宙放射線研究室では、2000年10月に米仏の研究機関と協力して打ち上げた γ 線バースト探査衛星HETE2で、 γ 線バースト発生直後のわずかな残光を観測し、その発生位置(対応天体)を知るための地上速報システムを開発した。我々は、NASA/GSFCからインターネットのソケット通信を用いて発信される、HETE2により得られた γ 線バーストの発生位置等の情報を元に、 γ 線バースト発生後数分以内に、美星天文台や宮崎大学等の国内の天文台で可視光追観測を自動で行なう地上システムを構築した。そのシステムは、現在稼働中であるがHETE2のデータを数十分ごとの地上交信でダウンリンクされるたびに自動的にMITのミッション運用センターから理研に転送し、自動処理を行なって、光度曲線やスペクトル、およびX線位置分布のプロットを製作して蓄積し、WWWインターフェースによって研究者が閲覧できる。また、ガンマ線バーストがあった場合には、地上詳細解析も自動的に行なわれ、その結果も自動的にデータベースに蓄積されるシステムも合わせて構築した。

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

ROSAT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。2000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、2000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。また他のミラーデータとしては、前に述べたように、X線天文衛星BeppoSAXのデータが組み込まれ、2001年1月から公開されている。

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を目標とした。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多くなり、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、一般ユーザー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（海外機関との伝送を含む）が重要な位置を占め、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た。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極めて有効に稼動している。

6. まとめ

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は真に日本の宇宙科学のデータセンターと言えるレベルに達した。さらに（1）当初のプロジェクト計画では検討事項であった「ぎんが」、「はるか」など宇宙研衛星で取得した他の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の構築を始めることができた。（2）DARTSアーカイブデータを利用するための当システム独自の各種解析支援ツールの開発を行う事ができた。（3）利用価値の高い外国の宇宙科学データベースの中継公開（ミラーリング）を開始する事ができた。宇宙研PLAINセンターでは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針である。

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た。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・運用班には校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た。さらに海外からは、米国NASA /GSFC、ドイツのマックス・プランク研究所、イタリアのBeppoSAX Science Data Center の協力を得た。科学技術振興事業団で雇用され各機関に派遣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た。