

# 天文学連携データベースシステム（ヴァーチャル天文台）の開発 — プロトタイプ第3版の開発 —

本田 敏志<sup>†</sup> 白崎 裕治<sup>†</sup> 田中 昌宏<sup>†</sup> 川野元 聡<sup>†</sup> 大石 雅寿<sup>†</sup> 水本 好彦<sup>†</sup> 大江 将史<sup>†</sup>  
安田 直樹<sup>††</sup> 増永 良文<sup>†††</sup> 石原 康秀<sup>††††</sup> 山崎 昭一<sup>††††</sup> 瓦井 健二<sup>††††</sup>  
中本 啓之<sup>†††††</sup> 小林 佑介<sup>†††††</sup> 吉田 徳夫<sup>†††††</sup>

<sup>†</sup>国立天文台 〒181-8588 東京都三鷹市大沢 2-21-1

<sup>††</sup>東京大学 宇宙線研究所 〒277-8582 千葉県柏市柏の葉 5-1-5

<sup>†††</sup>お茶の水大学 理学部 情報科学科 〒112-8610 東京都文京区大塚 2-1-1

<sup>††††</sup>富士通株式会社 〒261-8588 千葉市美浜区中瀬 1-9-3

<sup>†††††</sup>株式会社セック 〒150-0031 東京都渋谷区桜丘町 22-14

E-mail: <sup>†</sup>{honda.satoshi, yuji.shirasaki, masahiro.tanaka, satoshi.kawanomoto, masatoshi.ohishi, mizumoto.y, masafumi.oe}@nao.ac.jp, <sup>††</sup>yasuda@icrr.u-tokyo.ac.jp, <sup>†††</sup>masunaga@is.ocha.ac.jp, <sup>††††</sup>{ishi, yamazaki, kawai}@ssd.ssg.fujitsu.com, <sup>†††††</sup>{nakamoto, kobayashi, tko-yoshida}@sec.co.jp

**あらまし** 天文学では、γ線から電波まで様々な波長の観測データをデータベース化し、それらを利用することによって、観測対象の天体について研究することが重要である。そのため、世界中の天文台に分散配置された様々なデータベースを連携し、統合的に扱えるシステム（VO）の開発が世界中で進められている。我々はVOのプロトタイプを製作し、天文データベースを統合的に扱うための検索言語やXMLメタデータ検索機能などを設計、実装してきた。新たに開発中の実用化を目指したプロトタイプ3では、他国VOとの連携を目標として標準プロトコルの制定を進め、すでに一部のVOとは接続に成功している。本論文では新たに開発、実装した機能を紹介するとともに、会場から他国VOと接続し、天文研究への応用例を紹介する。

**キーワード** 天文学DB, メタデータ管理, グリッドコンピューティング

## Development of Astronomical Database System

### — Development of prototype 3 —

Satoshi HONDA<sup>†</sup> Yuji SHIRASAKI<sup>†</sup> Masahiro TANAKA<sup>†</sup> Satoshi KAWANOMOTO<sup>†</sup> Masatoshi OHISHI<sup>†</sup> Yoshihiko MIZUMOTO<sup>†</sup> Masafumi OE<sup>†</sup> Naoki YASUDA<sup>††</sup> Yoshifumi MASUNAGA<sup>†††</sup> Yasuhide ISHIHARA<sup>††††</sup> Shouichi YAMAZAKI<sup>††††</sup> Kenji KAWARAI<sup>††††</sup> Hiroyuki NAKAMOTO<sup>†††††</sup> Yusuke KOBAYASHI<sup>†††††</sup> Tokuo YOSHIDA<sup>†††††</sup>

<sup>†</sup>National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo, 181-8588 Japan

<sup>††</sup>Institute for Cosmic Ray Research, University of Tokyo, 5-1-5 Kashiwa-no-Ha, Kashiwa, City, Chiba, 277-8582, Japan

<sup>†††</sup>Department of Information Science, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo, 112-8610, Japan

<sup>††††</sup>Fujitsu Limited 1-9-3 Nakase, Mihama-ku, Chiba, 261-8588, Japan

<sup>†††††</sup>System Engineering Consultants Co., Limited sakuragaoka-cho, shibuya-ku, Tokyo, 22-14 150-0031, Japan

E-mail: <sup>†</sup>{honda.satoshi, yuji.shirasaki, masahiro.tanaka, satoshi.kawanomoto, masatoshi.ohishi, mizumoto.y, masafumi.oe}@nao.ac.jp, <sup>††</sup>yasuda@icrr.u-tokyo.ac.jp, <sup>†††</sup>masunaga@is.ocha.ac.jp, <sup>††††</sup>{ishi, yamazaki, kawai}@ssd.ssg.fujitsu.com, <sup>†††††</sup>{nakamoto, kobayashi, tko-yoshida}@sec.co.jp

**Abstract** In the field of the astronomical research, it is important to use the multi-wavelength database for investigations of astronomical object. Therefore the development of VO (Virtual Observatory) is constructing all over the world. VO is the federate system of distributed astronomical database. We have built the prototype of VO, and have development the query language and the XML metadata search function. In the prototype 3, we establish a standard protocol for the cooperation with other VOs, and we have already succeeded in connection with a part of other VOs. This paper describes the

prototype 3 of JVO.

**Keyword** Astronomical Database, Meta-Data Management, Grid Computing

## 1. はじめに

近年の天文観測は様々な観測装置の開発によって、可視光の観測だけでなく、ガンマ線から電波まで広い波長域で観測を行うことが可能になり、対象となる天体の性質にさまざまな角度から迫ることができるようになった。さらに国立天文台のすばる望遠鏡をはじめとする大型望遠鏡の建設によって、これまで観測できなかった暗い天体の観測が可能になり、一度に多数の天体を観測することが可能になったため、非常に効率よく観測データが得られるようになった。

このように様々な観測装置や望遠鏡を使って得られた観測データは、観測者などによって画像処理などされた後に、科学的議論に乗せることができるデータとなる。処理された画像データやスペクトルデータは主に天文観測で標準とされている FITS[1]とよばれるバイナリ形式で各観測所などに保管され、引き出された科学的データについてはカタログデータとして論文発表される。得られた観測データは、他の波長域のデータや過去のデータと比較したりできるように、ほとんどがデータベース化され、ネットワーク上で公開される。

しかしながら、観測装置の多様化、大型化によって日々生産される観測データは膨大なものとなり、従来の手法ではもはや処理できなくなっている。例えば、すばる望遠鏡で年に数十 TB の観測データが生産され、建設中の ALMA は年に数 PB にも及び、今後さらに加速して増加していくと考えられている。

このように、現在の天文学では世界中の観測所に分散配置されている膨大なデータベースから、いかにして必要な情報を迅速に検索し、取得するかということが重要となっている。バーチャル天文台 (Virtual Observatory : VO) は近年著しい発展を見せている情報学の研究によるネットワーク技術やデータベース技術を使って、分散データベースをあたかも 1 つの巨大データベースであるかのように見せるためのシステムであり、今後の天文学においては必要なものと考えられている。そのため VO の開発は世界中で進められているが、世界のデータベースを連携するには各国の協力が必須であり、IVOA (International Virtual Observatory Alliance) [2]が各国の VO 連携のために結成され、プロトコルの標準化などが議論されている。我々は 2002 年度より JVO (Japanese Virtual Observatory) の開発をおこなっており[3,4]、本稿ではこれまでの開発によって蓄積された技術を元に、世界の天文データベースの連携を可能にしたプロトタイプ 3 について報告する。

## 2. これまでの JVO の開発

JVO の開発は、まず機能を限定したプロトタイプの開発を行い、用いた技術の有効性や問題点などを洗い出し、実用システムの基盤技術を蓄積していく手法で進めている。

### 2.1. プロトタイプ第 1、2 版

プロトタイプの第 1 版では天文分散データベースを連携検索するために、SQL を天文学用に拡張した JVOQL[5]を定義し、LAN 上に分散配置したデータベースに対し GRID 機能を用いて実際に検索実行することを目標として製作した。これらは十分動作することが実証されたが、GRID 機能を実現するために用いた Globus Toolkit 2[6]は長時間ジョブの発行といった利用を想定しているため、結果を問い合わせる間隔が長く、VO にとって実用的ではなかった[7]。

そこで、第 2 版では Globus Toolkit 3 の GridServices を活用することによって、ジョブ起動の待ち時間をほとんど無くすことに成功した。さらに、増加する天文データベースに対応するため、メタデータを XML 形式にし、レジストリとして XML データベース (カレアレア[8]) に登録する機能を持たせ、ユーザーインターフェースを改良し、SQL の知識が無くても検索が行える JVOQL エディタを実装した。実際にすばるの観測データを使って、天文学研究の検索を行ったところ、非常に有効であり実用的なことが証明された[9]。

## 3. プロトタイプ第 3 版の開発

これまでの 2 つのプロトタイプの評価などを踏まえて、プロトタイプの第 3 版では運用モデルとして、海外 VO との連携を主な目標とした開発を行っている。海外 VO との接続が可能となれば、VO の目指す、「いつでもどこからでもあらゆるデータを得ることができる」システムに大きく近づくことになる。

### 3.1. システムの概略

JVO は主に以下の 4 つから構成される。

- ユーザからの検索や解析の指令を受け、結果を表示するユーザーインターフェース部
- 検索、解析指令を受け付け、メタデータを管理するレジストリサーバやデータサーバの処理を呼び出し、結果を返すポータルサーバ部、
- ポータルからの検索サービスや解析サービス要求に対して処理を実行し結果を返すデータサーバ部、
- メタデータの管理検索を行うレジストリサーバ

部

システムの開発はマシン環境に依存しない Java を用いている。

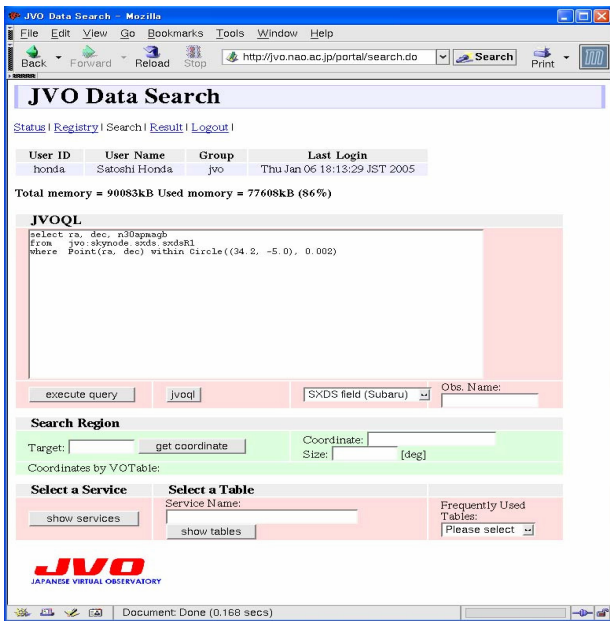


図 1 JVO データ検索画面。  
Fig.1 The screen of data search.

JVO の利用者はポータルサーバにアクセスし認証を行った後に、JVOQL を記述するか JVOQL エディタから検索条件を入力し、検索指令を送る (図 1)。ユーザーインターフェース部は Web ベースで作成しており、JVOQL エディタの他に、検索結果の表示機能、特定のデータ解析用画面、ステータス表示機能などがある。

入力された JVOQL はポータルサーバ内のパーサによって NativeSQL に分解され、これを元に一連の処理を記した Workflow を生成する。天文データの検索や解析作業は、1 度の検索で終了することではなく、様々な検索や解析の組み合わせが必要となるため、実行手順はダイナミックなものでなければならない。クロスマッチ検索を行うときには検索条件に合う天体の個数が少ない方を先に検索することが求められ、1 つのサーバに対する検索が失敗した場合には別の似たサーバへ検索の再実行といったことに対応するためである。

生成された Workflow は VOSpace と呼ばれるユーザごとの個人用データ保管領域で管理され、再検索を行う場合や、解析処理の追加などに利用することができる。将来的に VOSpace には各ユーザ用のデータベースを備え、他の VO と共有することも可能にすることを目標としている。また、利用者管理や認証、登録情報については LDAP サーバ[10]を用いて管理される。

Workflow に記された様々な検索要求はそれぞれに対応するデータベースを検索用レジストリサーバへ問

い合わせ、その結果から、各データサービスに対応する形式に変換され、検索が実行される。データサーバは世界中に存在している天文データベースや解析サービスであり、通常これらはリレーショナルデータベースであるが、異機種のハードウェアであったり、問い合わせの形式は個々に違っていたりすることが多い。この問題に対し、後で述べるようにインターフェースの標準化が進められている。

検索結果は VO におけるデータ交換の標準フォーマットとされている XML 形式の VOTable[11]として変換される。結果を表示するユーザーインターフェースには VOTable をユーザの見やすい表示に変えることのできる機能も実装した。

### 3.2. レジストリ

VO は世界の分散データベースを扱うことから、アクセスするサーバは様々で、ダイナミックに変化する。そのため、VO はそれぞれのデータベースについて URL やテーブル情報などのメタデータをレジストリとして登録し公開する必要がある。レジストリはデータサーバがメタデータを登録する Publishing Registry と、Publishing Registry からメタデータを取得し、登録されているメタデータを検索することができる Searchable Registry の 2 種類が存在する。

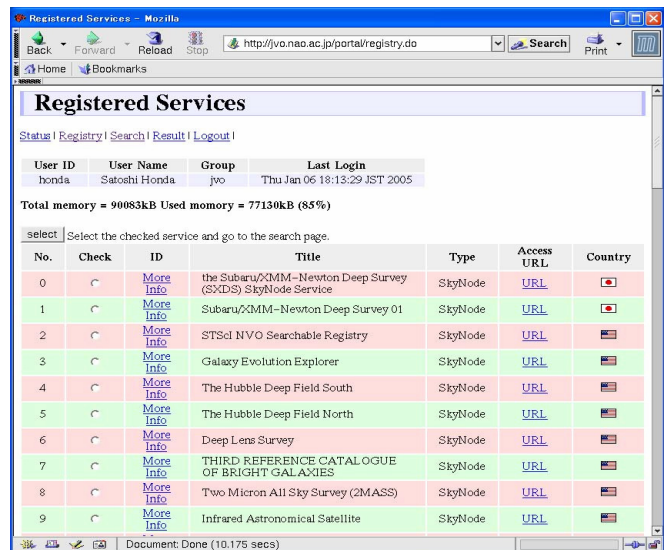


図 2 現在登録されているデータサービスの一覧。  
Fig.2 The list of registered data services at present.

天文データベースでは、メタデータも変更される可能性があるため、常にメタデータの交換を行わなければならない。そのため IVOA の取り決めにより、OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) [12]プロトコルを利用してメタデータの交換を行うこととし、登録される項目も決められ、データ構造を定義するスキーマも用意されている。

レジストリサービスは、Publishing Registry に対して、起動時、及び定期的にサービス情報の更新を問い合わせ、差分情報を取得し、自身の持つサービス情報レジストリに登録して管理する(図2)。差分情報の取得には、問い合わせ時にタイムスタンプを指定した選択的ハーベスティングにより実現する。

レジストリに登録されたサービスやメタ情報は、ポータルサーバからのカタログ検索要求やデータサービスの呼び出しに利用される。レジストリサービスは Grid Service のインターフェースを有するが、外部呼出し用に Proxy を用意することで、外部からの呼び出しには Grid Service を意識する必要がなくなる。

### 3.3. サーバ連携

既に述べたように、世界各地の天文データベースは独立に開発されているため、そのインターフェースは多種多様なものとなっている。そのため、データベースの連携には、インターフェースの国際的な標準化が必要である。IVOA では VO の標準プロトコルの制定を行っており、Simple Image Access Protocol (SIAP) と Astronomical Data Query Language (ADQL) といったものがこれまでに策定され、VO ではこれらのプロトコルに対応することが求められる。

SIAP は天体撮像データの検索を目的としたもので、天球上の領域等を指定する複数の検索パラメータを HTTP の GET メソッド等で送信することにより、取得可能な画像データへのアクセス方法やそのメタデータを VOTable として取得することができる。

ADQL は SQL 言語を拡張したもので、天文データ検索が容易に記述できるよう定義されている。ADQL 自体は XML 文書として表現され、Web サービスを使って渡される。現状の ADQL の仕様では画像検索を行う機能はもっていないが、先に述べた JVOQL では SIAP と ADQL を統合した仕様策定も進めており、これによる検索機能も有している。

既存のデータベースを大幅に変更することなく、これらの機能を追加させることが望ましいので、VO に対応したデータサービスとして公開するためのツールキットの開発を行っている。

### 4. サイエンスへの応用

JVO を実際に使ったサイエンスユースケースとして、すばる望遠鏡で得られたデータから重力レンズ天体を探す作業を行い、JVO の実用性を検証した。これはカタログ検索、データの解析、再検索、画像検索、画像の切り出しと表示等の作業が必要で、従来の手法では数時間かかる作業であった。しかし、JVO を使うと、数分ですべての処理が終了した。このことから、JVO

が天文研究にきわめて有効であることが実証された。プロトタイプ 3 では海外のデータベースも利用して、QSO 周辺のイメージ検索による母銀河の検索や、コスミックストリングの探査[13]と言った、天文学的にインパクトの大きなテーマに取り組む準備をしている。

### 5. さいごに

プロトタイプ 3 の開発によって海外の VO と相互接続が可能になったことで、JVO の基盤技術は整ったと考えられる。今後は実際にユーザが使用して、あらゆる要求に対応できるか調査し、必要な機能を追加しながら運用システムを構築することとなり、データベース天文学は大いに進歩すると予想される。さらに JVO によるデータベースへの容易なアクセスは、これまでの限られた天文研究者だけでなく、多くのユーザが最新データを使うことが可能になり、教育の場での活用、天文学の底上げにもつながると期待される。

### 謝辞

これまでの研究開発を支えていただいた国立天文台スタッフの方々に深く感謝致します。本研究は、文部科学省科学研究費補助金特定領域研究「情報学」(15017289,16016292)及び独立行政法人日本学術振興会先端研究拠点事業による支援を得た。

### 文 献

- [1] FITS Support Office <http://fits.gsfc.nasa.gov/>
- [2] IVOA <http://www.ivoa.net>
- [3] 白崎裕治, 世界中の天文データベース連携を実現するバーチャル天文台 “情報処理, Vol.45, No.12, 1219, 2004
- [4] 田中昌宏, 他 "JVO プロトタイプシステムの開発", DEWS2004, 2004年3月.
- [5] 白崎裕治, 他, "天文学 DB 用検索言語仕様の考察", DEWS2004, 2004年3月.
- [6] <http://www.globus.org/>
- [7] 田中昌宏, 他, "バーチャル天文台 JVO プロトタイプシステムの開発" DBSJ Letters, Vol.3, No1, 81-84, 2004
- [8] <http://www.sec.co.jp/products/karearea/>
- [9] 本田敏志, 他, "JVO の研究開発 -プロトタイプの概要", DEWS2004, 2004年3月.
- [10] <http://www.openldap.org/>
- [11] VOTable <http://vizier.u-strasbg.fr/doc/VOTable/>
- [12] <http://www.openarchives.org/>
- [13] Shirasaki et al. "Searching for a cosmic string through the gravitational lens effect: Japanese Virtual Observatory science use case", ADASSXIII, Vol.13, 46, 2004