

JVO の研究開発 —プロトタイプの概要—

本田 敏志[†] 白崎 裕治[†] 田中 昌弘[†] 大石 雅寿[†] 水本 好彦[†]

安田 直樹^{††} 増永 良文^{†††}

[†] 国立天文台 〒181-8588 東京都三鷹市大沢 2-21-1

^{††} 東京大学 宇宙線研究所 〒277-8582 千葉県柏市柏の葉 5-1-5

^{†††} お茶の水女子大学 理学部 情報科学科 〒112-8610 東京都文京区大塚 2-1-1

E-mail: [†] {honda.satoshi,yuji.shirasaki,masahiro.tanaka,masatoshi.ohishi,mizumoto.y@nao.ac.jp,
^{††} yasuda@icrr.u-tokyo.ac.jp, ^{†††} masunaga@is.ocha.ac.jp

あらまし Virtual Observatory (仮想天文台) は様々な観測所において独立に管理されている観測データやカタログなど天文データベースへのアクセスを一元的に行い、連携した検索・解析を GRID 技術を用いて行うことを目標としたシステムである。我々は、これまでに JVO プロトタイプの構築を行い、天文データベースへのアクセスするための問い合わせ言語 JVOQL と、GRID 技術をつかって、実際に分散配置された天文データベースの検索ができることを確認した。本稿では JVO プロトタイプの概要とプロトタイプ第 2 版について述べる。

キーワード 天文学 DB, 分散 DB, DB 言語, 科学 DB

Construction of Japanese Virtual Observatory —Outline of JVO prototype—

Satoshi HONDA[†] Yuji SHIRASAKI[†] Masahiro TANAKA[†] Masatoshi OHISHI[†] Yoshihiko
MIZUMOTO[†] Naoki YASUDA^{††} and Yoshifumi MASUNAGA^{†††}

[†] National Astronomical Observatory of Japan 2-21-1 Osawa, Mitaka, Tokyo, 181-8588 Japan

^{††} Institute for Cosmic Ray Research, University of Tokyo, 5-1-5 Kashiwa-no-Ha, Kashiwa, City, Chiba,
277-8582, Japan

^{†††} Department of Information Science, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo, 112-8610, Japan

E-mail: [†] {honda.satoshi,yuji.shirasaki,masahiro.tanaka,masatoshi.ohishi,mizumoto.y@nao.ac.jp,
^{††} yasuda@icrr.u-tokyo.ac.jp, ^{†††} masunaga@is.ocha.ac.jp

Abstract The JVO aims to provide federated astronomical database and data analysis environment through the Grid technology. We have build the prototype of JVO, and we confirmed that it could search the distributed astronomical database using the GRID technology and query language for access astronomical database (JVOQL). This paper describes the outline of the prototype of JVO and the second version.

Keyword Astronomical DB, Distributed DB, DB language, Scientific DB

1. はじめに

天文学の研究は、古代より様々な天体を観測し、その観測データから天体の素性を推測してモデルを作るという手法によって進歩してきた。天文学では実験が出来ないため、その作られたモデルは、さらなる観測をおこなうことによって検証される。そのため、観測を継続するとともに、過去の観測データを収集し、比較することが最も重要である。特に、時間とともに変化する天体の観測データは、その瞬間の現象が再現さ

れるとは限らないので、二度と手に入らない極めて重要なデータとなる。その上、天文観測では観測する波長域が違えば同じ天体でもまったく違った情報を得ることができるため、複数の波長域で得られた観測データの比較や、過去に得られた観測データと比較することによって、天体の物理的性質を研究している。

このように、天文の観測データは様々な角度から利用される可能性があることから、写真技術の進歩やデジタル化の影響も手伝って、ほとんどの観測で、データを保存しておくことが習慣となっており、観測デー

タの客観性も高まっている。

2. 天文データベースの現状と VO の必要性

近年の天文観測は、望遠鏡や検出器の発展により、非常に精度の高いデータが大量に生産されている。たとえば、国立天文台野辺山観測所は～1 TB/年、ハワイのすばる望遠鏡は～20TB/年、現在チリで建設中の ALMA にいたっては～PB/年のデータが生産される。これらのデータはアーカイブデータとして公開され、その結果、実際の観測者によって得られた研究結果だけでなく、アーカイブデータを使うことによって得られた研究結果も数多く生み出されている。特に、HST（ハッブル宇宙望遠鏡）で得られるような地上の望遠鏡では得ることが不可能な貴重なデータは、観測者による一定の占有期間を経たのちにすべてアーカイブデータとして公開されることを前提としており、ネットワークを使って自由にアクセスすることが可能となっている。

先に述べたように、天文観測では複数の波長域での観測データは極めて重要である。しかしながら、一般に波長域の違うデータはそれぞれ異なった観測装置で観測され、異なった観測装置によって得られたデータを扱うには、それぞれの観測データがアーカイブされているデータベースにアクセスし、必要なデータを検索して取得した後に、観測装置固有の情報を用いた解析を行う必要がある。そのため、複数の波長域のデータを扱うには多大な労力が必要である。上述の HST では装置の特性パラメータを含む最適化された手法を解析パッケージとして公開していることもあり、アーカイブデータを使った研究が非常に多いが、一般にこのような作業を行うにはある程度の観測装置に関する知識と解析の経験が必要とする。

さらに、近年盛んに行われているサーベイデータを扱うにはデータの転送や解析に現状の計算機資源では不十分な場合が多い。そのため、貴重な情報が含まれている可能性があるにもかかわらず、ほとんど使われることの無いアーカイブ化されたデータも多数存在する。

そこで、複数の波長域のデータを使った統計的な研究を行うために、分散配置されたデータベースへ一元的にアクセスし、各データベース同士で検索や解析ができるシステムが必要である。このような研究を可能にするために、Web サービス、GRID 技術、データベース技術など様々な情報技術を用いた研究が進められており、特に、観測データを収め管理する計算機やデータベースシステムを「望遠鏡」、そのデータを解析するツールを「観測装置」として、データアーカイブに記録された宇宙を観測するという意味から VO(Virtual

Observatory)と呼ばれるシステムの開発が世界中で進められている[1],[2]。

VO の開発は主に以下のようなことを目標として進められている。

- ・ いつでも、どこでも観測ができる。
- ・ データ解析は VO 内部で行い、大量のデータ転送を行わない。
- ・ 観測装置によらない統一的な操作性により、多波長にわたる研究を容易にする。

これらを実現することにより、新しい発見や、大量データの統計的研究が進むことだけでなく、教育や啓蒙に役立つことが期待されている。

本稿では我々が開発を行っている VO のプロトタイプについて、概要と動作の流れ、そして実際に天文データを使ってサイエンスを行った結果について述べる。

3. JVO の開発と

我々は日本の VO(JVO)の開発を進めており[3]これまでにプロトタイプの構築を行った[4],[5]。

JVO の開発は、必要と考えられる最低限の機能のみを実装したプロトタイプを作成し、その評価を行った後に改良を加え、新たな機能を追加した次のプロトタイプを作成していくという方針で行っている[4]。

プロトタイプの第 1 版は、次の機能が実際に機能するのか評価することを目標として開発された。新たに問い合わせ言語として SQL を拡張した JVOQL を定義し、複数の分散データベースを統一的に扱う方法を開発した[6]。JVOQL は、不特定多数のデータベースを検索、天体の座標を指定して検索、と言った想定される検索方式を考慮している。観測データは世界中の観測所や研究所の計算機に置かれ、ネットワークで接続されているが、これらの計算機資源を透過的に使うために GRID 技術を採用した。これらの機能は十分動作することが実証された[4]。

プロトタイプ第 2 版では、第 1 版の評価を元にシステムの改良を行った。以下、それぞれプロトタイプ第 1 版、第 2 版について述べる。

3.1. JVO プロトタイプ第 1 版

プロトタイプ第 1 版は主に JVOQL を投入するメインコントロールウィンドウ(図 1)と、検索結果を表示するページ(図 2)からなる。

実際の使用は、メインウィンドウを起動し、JVOQL 入力フィールドに JVOQL を入力し実行することで、JVOQL が JVOQL パーサに送られる。パーサは JVOQL の定義に従って指定された検索要求を構文解析し、利用するサーバ、必要なカラムの情報などを抽出する。複数のサーバへの検索が記述された検索要求を個々の

サーバが処理可能な検索要求に分解する。

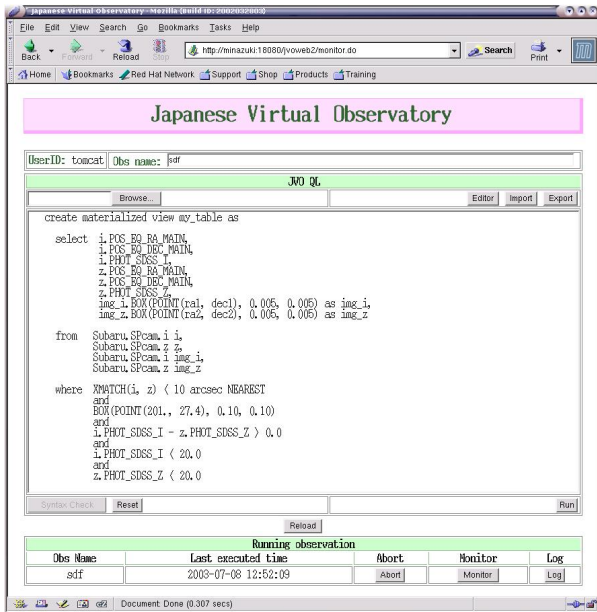


図 1 メインコントロールウィンドウ
Fig.1 Main Control Window

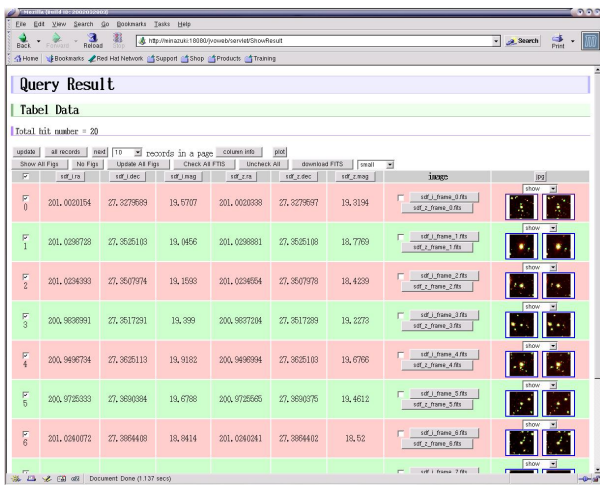


図 2 検索結果の表示
Fig.2 Result page

サーバごとに分解された検索要求は順番に並べられ、順に個々のサーバに検索要求が送られる。各サーバに実行要求が送られ、実行されていく様子はモニターすることができる(図 3)。

JVOQL の入力には QL を直接入力することもできるが、利用者が指定する検索条件を JVOQL に変換するユーザーインターフェースを開発した(図 4)。検索条件として、対象となるカタログ・画像データベース、天球座標、観測時刻、波長、画像取得、出力形式、クロ

スマッチ条件が指定できる。これらの機能により利用者は SQL の知識を必要とすることなく、必要な検索条件を指定できる。カタログを指定する部分では、データベースにアクセスし利用可能なカタログ・画像データベースの検索が行える。また各カタログのカラム情報の表示も可能である。検索・表示は汎用性を Java servlet、Java Beans、JSP から構成される Web サービスにより実現している。

得られる結果は、天文学データを XML 形式で記述した VOTable[7]と呼ばれる VO では国際的に標準化された形式で得られ、HTML に変換して表示される(図 2)。データの転送に XML 形式である VOTable を採用することによって、将来的に VO の互換性が見込まれる上、得られた結果をデータベースに登録して、さらに検索を行うことも考えられる。

実際に、国立天文台すばる望遠鏡[8]の観測データをつかってカタログデータベース、画像データベースを異なるサーバに配置し、データベース検索、クロスマッチ検索、画像の切り出しなどを行い、性能評価を行った。その結果、解決すべき課題も明らかになった。例えば、GRID 機能を実現するために利用した Globus Toolkit 2[9]は認証の部分でオーバーヘッドがあるために遅い。これはすぐに終わる処理を考慮して作られておらず、結果を問い合わせる間隔が長いのである[2]。利用者の検索結果などのデータ管理や稼働状況の把握も必要である。複数の利用者を想定した場合に、一部の利用者によって得られたデータが、ディスク容量を圧迫するといった、計算機資源の利用において互いに干渉する可能性が考えられる。

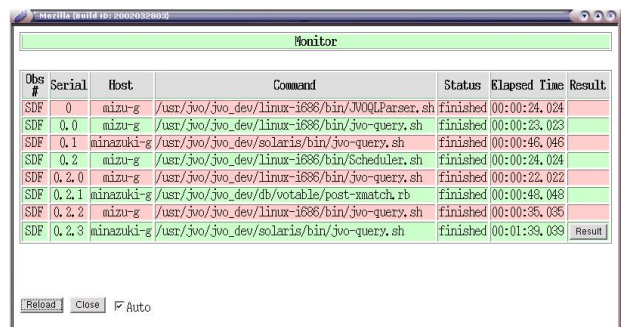


図 3 モニター画面
Fig.3 Monitor window

3.2. JVO プロトタイプ第 2 版

観測データは常に追加されていくため、ユーザが利用可能なデータベースのリストは登録や更新が容易に行える必要がある。第 2 版では、観測データが保管さ

れているデータベースサーバの情報や観測データなどのメタデータを XML 形式にし、レジストリとして XML データベースに登録する機能を持たせた。これによって、現在利用可能なアーカイブデータとそれらからどのようなデータを取得できるか等、利用可能なサービスの登録情報を JVOQL で検索できるようになり、アクセスできるデータベースのリストなども動的に生成することができるうえに、大量のデータベースを検索する場合にも有効である。

このメタデータは IVOA[1]の案に従って設計している。これは近い将来、世界の VO と連携することを考えたものである。

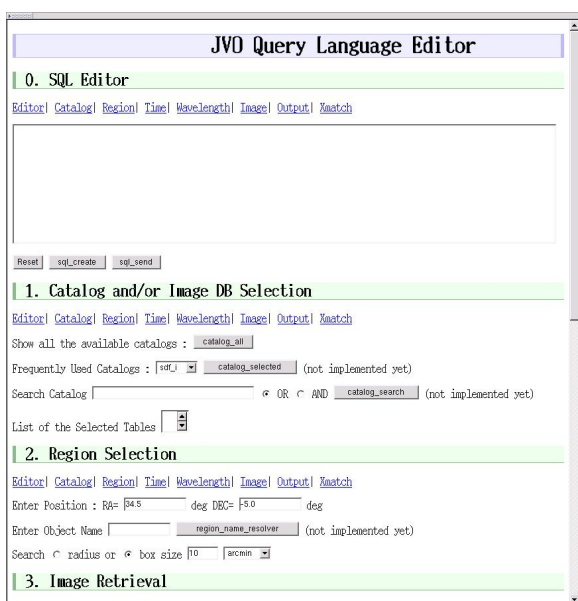


図 4 JVOQL エディタ
Fig.4 JVOQL Editor

第 2 版は、構成される基本画面は Top page、DB Navigator、Search、Analysis、Data Display の 5 つで、各基本画面にはサブ画面が存在する。Top page はユーザ認証を行うページであり、進行状況も確認できる。利用者認証機能を持たせ、利用者を管理することによって、利用者ごとのジョブ管理、データ管理を行うことを可能にし、検索の中断機能を持たせることも可能となった。検索を行う場合は Search ページで JVOQL エディタから検索条件を指定し、検索を実行する。検索は、入力された検索条件をもとに XPath を生成し、レジストリへ渡される。検索結果は XML 形式で出力され、パーサおよびスケジューラがそれを元に実行手順を作成し、エグゼキュータによって各サーバへ問い合わせる。GRID 機能の実現には Globus Toolkit の最新版である Ver3(OGSA)を採用し、そこで新たに導入され

た「グリッドサービス」を遠隔実行に用いた。これにより、第 1 版で課題となった認証でのオーバーヘッドが短くなることが期待される。さらに、グリッドサービスに合わせて、各サービスも再構成を行った。各サーバとの間のデータ転送には認証付きのファイルシステム SFS を新たに採用し、各データベースサーバ間の転送には RFT を採用している。各サービスのインターフェースは XML 形式に統一され、今後の汎用性、拡張性もあがった。

検索結果は Data Display ページで表示されるが、得られた結果を各ユーザに用意される個人データベースに保管することができ、それを使って新たに解析を行ったり、あらかじめ登録しておいた解析エンジンを呼び出すことも可能である。Analysis ページには特定のサイエンス専用の検索、解析機能を組み込んだページを追加することも容易な設計にした。

4. JVO 使ったサイエンスユースケース

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う[10]。重力レンズ天体とは巨大質量天体の強大な重力によって空間がひずみ、天体の向こうにある別の天体の像がその巨大天体の周囲に複数見える現象である。以下の作業を行うこととなる。

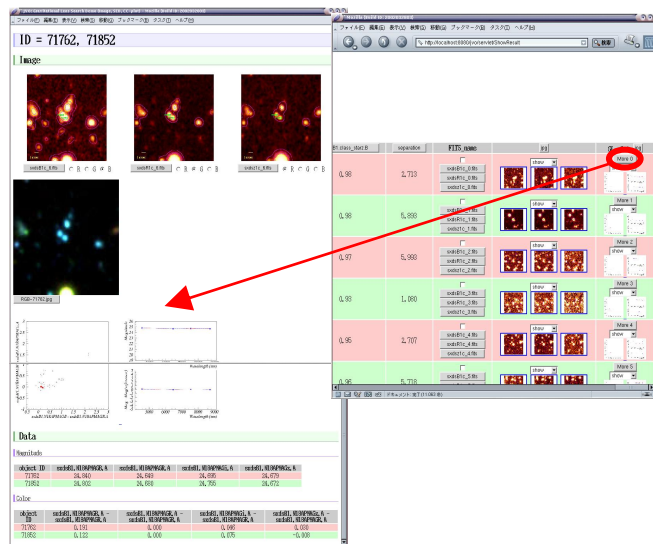


図 5 結果と解析のページ
Fig.5 Result and Analysis page

1. すばるのカタログから指定した領域のデータを取ってくる。
2. とってきたデータをもとに明るさを計算する。
3. クエーサーを選ぶ条件を決める。
- 4.

取ってきたデータからクエーサー候補のペアリストを作成する。5. ペア天体の周辺の画像データを検索する。6. 画像データを解析し、さらに候補を絞り込む。

これらの作業を JVO を使わないで行うと、何時間もかかる作業であるが、実際 JVO を使うと 5 分程度で候補天体を見つけて表示することが出来た(図 5)。さらに検索条件を変えて再検索も簡単に行うことができ、きわめて有効であることが実証された。

5. まとめ

本稿では JVO プロトタイプについて、第 1 版と第 2 版それぞれのシステム構成と動作の一連の流れについて述べた。第 1 版では複数のデータベースを連携検索する問い合わせ言語として JVOQL を定義し、実際に分散配置されたデータベースの検索を GRID 機能を用いて実行することができた。第 2 版では Globus Toolkit 3 を導入するなど最新の技術を導入し、それに合わせて各サービスも再構築した。さらに、メタデータをレジストリに登録する機能を持たせたことで、必要な検索サービスや解析サービスが存在するサーバの情報を容易に得ることが出来るようになり、拡張性や互換性を持たせた。サイエンスに対する有効性も実証された。しかしながら、個人データベースのデータ領域の扱いや、ユーザ登録、複雑な検索・解析を一括処理する方法など課題も残されており、今後プロトタイプ 2 の評価を行った後、実用システムの開発を始める予定である。

文 献

- [1] International Virtual Observatory Alliance
<http://www.ivoa.net/>
- [2] 洞口俊博, “世界の「仮想天文台」計画” 天文月報, Vol.95, No.7, pp.318-324, 2002
- [3] 大石雅寿, “Japanese Virtual Observatory の構築” 天文月報, Vol.95, No.12, pp.566-575, 2002
- [4] 田中昌宏他, “JVO プロトタイプシステムの実装” DEWS2004
- [5] Ohishi, M, et al. “A Prototype toward Japanese Virtual Observatory (JVO)” SAINT2004, 2004
- [6] 白崎裕治他, “天文学 DB 用検索言語の開発” DEWS2004
- [7] VOTable Documentation
<http://vizier.u-strasbg.fr/doc/VOTable/>
- [8] Kaifu, Norio, “Subaru Telescope” Proceedings of SPIE, Vol. 3352, pp.14-22, 1998
- [9] <http://www.globus.org/>
- [10] Shirasaki, Y, et al. “Searching for a cosmic string through the gravitational lens effect: Japanese Virtual Observatory science use case” ADASS XIII, 2004