

天文学データセンターの運用について

中 嶋 浩 一

1 はじめに

筆者は現在、1993年4月～1994年3月の予定で、フランス、ストラスブールの「天文学データセンター」(Centre de Données astronomiques de Strasbourg, CDS)に滞在しており、このようなセンターについて、その意義、運用方法、その現状、将来の方向、等について調査中である。そこで、本報告において、これまでの調査の結果知られたこと、およびその運用に関する考察、などについて述べてみたい。なお、このようなデータセンターの、日本の天文学界における必要性・あり方などについても、最近、大型望遠鏡の利用方法等に関連して議論されており、これについてもあわせて検討する。

2 ストラスブール天文学データセンター (CDS) について

フランス、アルザス地方の主要都市ストラスブールは、現在ではECの会議場の所在地として知られているが、近代の歴史では、独仏国境の町としてたびたびその領有が変わったこと(19世紀後半から5度も変わった)、このため住民の使用言語や氏名までが度々強制的に変更させられたこと(ドーデー、『最後の授業』、など)、にもかかわらずこの地方には独自のアルザス文化がありアルザス語が使われていたこと(田中克彦、『国家と言語』、岩波新書、参照)、などでも知られているところである。またストラスブールは大学の町としても有名であり、欧州各地から多くの学生を集めている(ゲートもここで学んだ)。この旧ストラスブール大学は、現在、自然科学系、社会科学

系、法学系の3つの大学に分けられ、その自然科学系は、以前ここで細菌学のルイ・パスツールが教鞭を取っていたということから「ルイ・パスツール大学」と呼ばれている。ストラスブール天文台はこのルイ・パスツール大学所属の天文台、というわけである。現在は、約30人のスタッフと10人あまりの大学院相当の若手研究者の活躍の場となっている。

由緒ある天文台の例に漏れず、この天文台も都市近郊にあるために、望遠鏡等による天文観測は現在では行われていない。昔の天文観測は、星の見える方向を測定してその土地の経緯度や時刻を決定していたが、現在ではこれは原子時計や電波観測などによって行われるようになり、伝統ある天文台の業務は「その歴史的使命を終えた」と宣告された。このため、英国のグリニジ天文台のようにその活動を停止してしまったところもあるが、多くの天文台はそれぞれ個性ある天文学研究センターとして再出発した。そのなかでストラスブール天文台は、その地の利を生かして、天文学のコンピュータデータセンターとしてより一層の活躍を続けてきた、というわけである。ストラスブールの町の名前の由来は、ゲルマン語(ドイツ語とはあえて言わない)の“Strasse”の“Burg”, すなわちライン川による南北の往来とパリから東に伸びる東西の街道との十字路に発展した町、ということであるが、ストラスブール天文台はコンピュータのネットワーク、すなわち今様の「電子の街道」の要衝として発展している、ということか。

このストラスブール天文台の活動の中心としてのデータセンターは、“Centre Données astronomiques de Strasbourg”, 略称“CDS”と呼ばれる組織で、フランスの大学審議会のような INSU という機関とルイ・パスツール大学との合意に基づいて天文台の中に設置されたことになっている。最も、最初は“Centre Données Stellaires”という意味での CDS で、いわば「恒星データセンター」であった。設立は1972年である。その後、銀河系や電波天体などの、恒星以外の天体も含むようになって、現在の名称になった。

この組織の規模は、1991年の実績で、人員で20人弱、予算で5千万円程度、設備は比較的大型のワークステーションが実働で8台、ということで、

世界の水準からすれば決して大規模な組織ではない。このような機関が、設立以来 20 年余、世界の中心としての天文学データセンターとして活躍を続け、高い評価を得ているのはどのような理由によるのだろうか。

3 CDS のサービス活動の内容

現在の CDS の、データセンターとしてのサービス活動は、大きく分けて「SIMBAD」と「カタログアーカイブ」の 2 種類がある。

3-1 SIMBAD について

「SIMBAD」というのは、恒星や銀河系などのいろいろな天体に関する一種のデータベースである。しかしそれは単純なデータベースではなくて、必要なデータを高速で検索する特殊なユーザーインターフェイスのソフトウェアを装備しており、この点に大きな特徴がある。以下にこれを説明する。

天体に関するデータとしては、たとえば恒星であればその名称(これも慣習により一つの星でもいろいろな呼び名がある。たとえば七夕星の「ベガ」は、 α Lyr など、計 53 もの名称が表示される！)、天球上の位置、等級、スペクトル型、などがあり、また銀河系などではその形状に関するパラメータなどが考えられる。しかし、これらの天体を研究する立場から言えば、このような定型的なデータよりもむしろ、この天体についてどのような観測データがこれまでに公表されているか、またこの天体についてこれまでにどのような研究論文が発表されているか、ということが最も重要である。そしてこの SIMBAD は、これらのデータを装備し、かつ高速で検索するところに大きな特徴があるのである。ちなみに、前述のベガの例で言えば 840 件の参考文献の所在が表示され、またアンドロメダ銀河を検索すると 1190 件の研究論文があることがわかる。SIMBAD の意味するところは“Set of Identifications, Measurements and Bibliography for Astronomical Data”であるが、これはその特徴を良く表わした名称となっている。

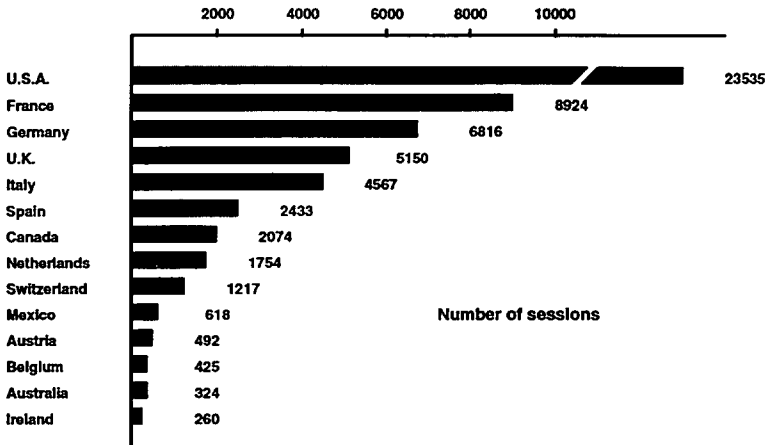
さらに重要なことは、天文学の研究論文がそれこそ時々刻々出版されるの

に合わせて、この最新の情報をデータベースに取り込んで行かねばならない、ということである。このためにはたくさんの研究論文にきちんと目を通すことが必要不可欠であり、ちょっと考えただけでも実に変な作業であることがわかる。CDSでは、これを、パリ天体物理学研究所、パリ天文台、ボルドー天文台などとの共同作業を組織して実行している。このほかにも、リヨン天文台、スイスのジュネーブ天文台、ローザンヌ天文研究所、ドイツのハイデルベルク天文研究所、米国のNASAなどが、観測データのアップデート作業に協力している。研究論文を調査することも大変ならば、共同作業を組織することもまた大変な努力を必要とする。実際、一部の機関で今後の共同作業の存続が危ぶまれているということであり、関係者が頭を痛めている。

SIMBADでもう一つの重要な問題は、膨大なデータベースをいかに効率良く高速で検索するか、ということである。ユーザーが調べたい天体は、文字どおり星の数ほどあるわけであり(現在75万程の天体が検索可能)、これに対してそれぞれの天体に対応する各種のデータがさまざまな形式のデータファイルとしてコンピュータに蓄えられている。このように、いくつかのデータグループがあって、それらを関連(relation)させつつ必要なデータを抽出するというデータベース管理のテクニックとして「リレーショナルデータベース」という技術が良く知られている。ところが、このSIMBADのように検索エントリの数が膨大になると、リレーショナル……の方法では検索に大変手間取ってしまい、時間がかかって使い物にならない。そこでCDSでは、独自のソフトを開発して、事実上は瞬時に応答が得られるようなシステムを完成した。これは、データベース管理ソフト技術としてはこれまでにないものであり、天文学者の片手間の仕事ではなく、専門のソフト技術者のスタッフがかなりのエネルギーをつぎ込んで開発したものである。

このSIMBADを利用するには、CDSとの間で利用契約を結び、何がしかの料金を払ってSIMBADシステムコンピュータのアカウントを取得する必要がある。しかし一度それを取得すれば、現在、世界中の研究者の間にはりめぐらされているコンピュータネットワーク(internet, DECNETなど)を

図1 SIMBAD への、国別アクセス数
(1990年7月～1993年4月).
出典：CDS, Lesteven 氏報告より.



介して、世界各地から24時間、リアルタイムで利用することができる。実際、現在最も頻繁に利用しているのは米国の天文学者である(図1参照)。

このように、SIMBADは大変優れたデータベースでありシステムであるわけであるが、ここで注意したいことは、これを利用してどの程度の天文学上の研究成果が上げられているか、ということである。これに関しては、本報告の主要検討課題でもあり、後に改めて検討する。

3-2 カタログアーカイブ

天文学で「カタログ」というのは、いろいろな天体についての各種データをまとめた一種の一覧表で、たとえば「天体位置カタログ」、「恒星スペクトルカタログ」などさまざまなものがある。これの「アーカイブ」はまさに、これらのカタログを記録保管することである。むかしのバビロニアでは粘土板で、またシルクロードの敦煌では紙の巻物で保管されていたが、現代ではコンピュータのメモリの中に保管されている。

このような、いわゆる「マシンリーダブル」なアーカイブも、そのかたちは短期間に次々と変わってきた。CDSが発足した20年前は、あの「パンチカード」でカタログデータを集めて保管し配布していたものが、まもなく「磁気テープ」に変わり、さらにデータ配布についても、以前は磁気テープを郵送していたものが前述のコンピュータネットワークを介した直接配布（いわゆるftpによる）になっている。このため、記録の保管という観点でいえば、現在はそれは大容量磁気ディスクの中に蓄えられ、またそのバックアップとして磁気テープ（8mmビデオテープが使われる！）に保管されるようになっている。さらに、とどまるところを知らないデータ膨張のために、「CD-ROM」による保管・配布や「ジュークボックス」式のテープ記録装置なども検討されている。

さて、このカタログアーカイブの意義とそのサービスの特徴は何だろうか。

現在でも、研究成果はすべて論文の形で印刷出版される。観測データについても、データ量が膨大なものを除いてすべて同様である。これを利用して何か研究を行おうとする場合、論文の著者に連絡して何らかのマシンリーダブルな形でデータを送ってもらうわけであるが、送る立場からするとこれは実にたいへんな労力を要する作業であり、しかも自分の研究に必ずしも役に立つわけではない。結果としてこのような取り引きはあまり成功しないことになる。さらにそのデータに関する説明やフォーマットなどの利用法に関する説明なども、なかなか手に入りにくい。このようなことから、カタログデータを一箇所に集めてしっかりした説明をつけ、スタンダードな方法で利用できるようにすれば、その利益はユーザ・供給者ともにはかりしれないものがある。

もちろん、多量のデータを集めて「情報検索」に資するというメリットもあるが、現在のCDSのカタログアーカイブはそのための機能は提供していない。情報検索に関しては、データファイルを自分のコンピュータ上に移した後、自身でプログラムを作成して検索しなければならない。「データベース」というのは情報検索の機能を備えているべきであるという考え方からす

れば、これはデータベースとは呼べない。「カタログアーカイブ」というゆえんである。

このカタログアーカイブのサービスのための、CDSにおける業務の内容を見てみよう。まず最初に、カタログが出版されるのを常にサーベイしていること、そして出版情報を発見したらその著者に連絡してカタログを取り寄せること、から始めなければならない。次に、このカタログにスタンダードな形のしっかりした説明文をつけて、コンピュータデータファイルに登録する。この際、できるだけデータ形式の整合性のチェックを行い、データオリティの向上に努める。さらに、著者によるカタログの改良などにも手違いなく対応し、いわゆるバージョンの管理を行う。このほか、これらのカタログの利用状況などに関しても記録を取って管理することが必要である。

前項の内、「カタログの出版のサーベイ」については、CDSでは論文雑誌の発行者との協議を行って、ある程度ルーチ的にデータが入手できるようにする、などの努力を行っている。なおこの点に関連して興味深い提案がなされている。すなわち、論文雑誌の記事そのものをマシンリーダブル化して（「電子出版」と呼ぶ）カタログ同様にアーカイブしよう、という試みである。実際、いくつかの雑誌記事について、このカタログデータファイルに試験的な登録がなされている。これについては、後で再度検討する。

「データ形式の整合性のチェック」というのは、数字の入るべき位置に英字や記号などが入っていないか、カラムのずれはないか、昇順・降順などが守られているか、というような、すべて形式上の矛盾のチェックである。CDSではこのためのいろいろな検査プログラムを開発しており、興味深いテーマを提供している。

このほか、「スタンダードな説明文の付加」など手間のかかる作業がいろいろあるわけであるが、現在、これらはすべてCDSのスタッフだけで行っており、ほかの機関との共同作業は行われていない。これを何とかして幅広い共同作業に組織することが、当面の課題となっている。

このサービスの利用法に関しては、SIMBADと同様アカウントを取得す

る必要があるが、アカウントを持てば、前述の ftp によって容易にカタログを取得することができるだけでなく、ファイルシステムが UNIX 形式であるために、telnet によって login してデータの構成・所在などを見ることもできる。

4 CDS の近未来プロジェクト

上述のようなサービスを絶えず維持してゆくだけでも大変な作業であるが、CDS ではさらにこれを改良・拡大する作業を、すでに開始している。それは、全天の景色 (!) をすべてディジタル化し、これをデータベースに加えて、SIMBAD などで検索した天体が天球上でどのように見えどのような環境の中にあるか、ディスプレイ上で一目で見られるようにしようという壮大なプロジェクトである。コードネームは「ALADIN」である。

天文学で使われる特殊な望遠鏡に「シュミット望遠鏡」というのがある。これは、通常の望遠鏡が、文字どおり「ヨシのズイから天井をのぞく」ようにごく狭い範囲しか見ることができないのに対し、特殊な光学系を用いて広い範囲を大判の写真に一度に撮影する、という望遠鏡である。空の範囲としては 6° 四方を、また写真の大きさとしては 36 cm 四方というのが通常のシュミット望遠鏡である。米国パロマ山のものが有名であるが、日本にも、長野県の本曾山中に優れた性能のシュミット望遠鏡がある。この様な望遠鏡を用いて全天をくまなく撮影した写真集が、パロマ山や、南半球の欧州南天天文台などから出版されている。 6° 四方というのは、満月が 150 個近く入る大きさであるが、それでもこれで全天をカバーするのに 1000 枚以上必要である。これをすべてディジタル化してコンピュータで扱えるようにしようというのだから、ALADIN は壮大な計画であるというわけである。

とはいっても、どのように壮大なのかは、少し説明を要するかもしれない。この計画では、角度の $1''$ のメッシュで全天をディジタル化することになっているが、この大きさで全天をカバーすると、 5×10^{11} ドットになり、1 ドットを 1 バイトのデータとしても 500 GB のメモリが必要となる。これはフロ

ッピーディスクで約 40 万枚、光磁気ディスクでも 1000 枚という膨大なデータ量である。フィルターを変えて全天を観測することを考えると、最低でも 1000 GB（これを 1 テラバイト、1 TB と呼ぶ）が必要になる。このためには、ジュークボックス式のテープ装置など、新たな設備が必要になることは、前述したとおりである。

もっとも、全天の写真をディジタイズするプロジェクトは、世界の天文学者が協力して実行することになっており、昨年あたりから少しずつ開始されている。この計画では、最終的には 1 テラバイトの 1000 倍の“1 ペタバイト (1 PB)”を考えており、新たな意味での天文学的数字が出現した。

ALADIN プロジェクトのもう一つの特徴は、最近の高性能ワークステーションのグラフィック機能をフルに生かした「グラフィカルユーザーインターフェイス」のソフトウェアを開発することである。これを大画面の X 端末装置に搭載すれば、空の任意の部分の景色を背景に、目的の天体がどのような位置、どのような状況にあるのかを目で確認しながら、じっくり研究できるわけである。もちろんこれは SIMBAD などとも連携して、画面上の天体がそれぞれどんな天体なのか、識別表示することもできる。角度の 1" のメッシュは地上での眼視望遠鏡観測の精度に匹敵するものであり、これまでの天体観測と同様のことが、コンピュータの画面上で可能になってしまうということである。

データストレージの問題、ほかのサイトからの利用の問題等、まだまだ解決すべきことは多いが、来年あたりには試験的に利用できるようになる予定であり、その成果が期待される。

5 データセンター運営に関する諸問題

さて、現在ではこのように大活躍中の CDS であるが、将来のことを考えた場合、いろいろな問題が予想される。以下、この点について筆者のまったく私的な検討を行ってみたい。

5-1 支持基盤の確保の問題

CDSのような事業は、事業といってもやはり天文学の研究活動の一環であり、当然のことながら日本の国会図書館のような公的な事業ではない。したがって、少なくとも世界の天文学者の間にその支持基盤を確保しておかなければ、これを長期間存続させることはできない。また、天文学者の広汎な支持があっても、それをバックにして公的機関に積極的に働きかけ、予算や人員を確保しなければ活動できない。このため、このような事業はそれを推進しようとする人の「外交努力」ともいうべき活動が、本質的に重要になる。そしてこのような活動は、「天文学者になりたい」と考えるような人にとっては、通常、あまりやりたくない「雑用」と思われるわけである。

たしかに、一つの事業を行うにあたっては、多かれ少なかれ、外交努力や地道な雑用は必要不可欠であるが、同じ天文学の事業であっても、ロケットを打ち上げたり、新しい望遠鏡を建設したりする事業に比べると、CDSの事業はいろいろな面で異なる要素を持っている。すなわち、新しい望遠鏡の建設などは、数年間という比較的短い限られた期間だけ地道な努力をすれば、その後は宇宙的な天文学研究に没頭することもできるわけであるが（またすぐ次の事業を始める人もいるが）、CDSのようなデータベース管理やカタログアーカイブの仕事は、長期間地道に続けることが本質的であり、なかなか雑用から離れられない。また、新しい装置の建設は、完成したあかつきにはこれまで知られなかった新しい大発見も大いに期待できるわけであるが、アーカイブの場合には、はっきりいってこのような予想は困難である。

これに加えて、近年のように、研究のために限られたリソースを取り合うという傾向が顕著になってくると、（物理学の世界などではかなり以前からそうであったということであるが）研究業績の評価の問題が大変シビアになってくる。そして（アーカイブのように）研究者の研究環境を整備するという面でたとえ大きな貢献があったとしても、それを支持してくれる基盤を培っておかなければ、はなばなしい新発見に評価を奪われてしまうということになる。それやこれやで、結局、雑用ばかりということになりかねない。

アーカイブデータを利用して次々とはなばなしい研究成果をあげられるのであればあまり考慮する必要はないかもしれないが、必ずしもそうでないとするならば、データセンターの支持基盤を確保するためにその指導者の外交能力や人脈などの個人的・個性的な要素が重要になる、ということが大きな問題である。このような人材がいなければ、この事業は存続できないのではないか。

5-2 共同作業の組織作りの問題

前述のように、SIMBAD は多くの他の天文台・研究機関の協力を得て維持されているわけであるが、その文献調査のことを考えると、共同作業機関なくしては現状の SIMBAD の維持は不可能であると思われる。ところが、前項で検討したように、他の天文台にしてもこのような作業は個人的な能力と努力に依存する要素が大きいわけで、このことは、ある特定の個人がリタイアしてしまうと共同作業が続行不可能となってしまう、ということの意味する。実はこれはまさに現実に行っていることであって、前述のように関係者にとって頭の痛い問題となっている。やや乱暴な言い方をすれば、かつての同志が一人一人消えて行く中で、どのようにして残った構成員を引き締め、あるいは仲間を増やしていった運動を続けるべきか、という状況にもたとえられよう。データセンターの指導者には、このような意味での指導者の資質が求められるのではないか。

カタログアーカイブについても、共同作業機関の獲得は急務である。現状ではこの作業は、CDS の中の特定の個人の超人的な努力に大きく依存しているように見え、健全な状態ではないように筆者には思われる。また、貴重なデータが一箇所のセンターに集中したのでは、データ保全の観点からいっても大きな問題であり、世界各地に分散保管する必要がある、という指摘もある。その意味でもアーカイブの新たな共同作業機関の獲得が重要である。「貴重なデータ」という点に一つの哲学があるわけであるが、これについては次に検討する。またこれに関して「日本の役割」も問われるわけであるが、

後で検討する。)

5-3 アーカイブの哲学の問題

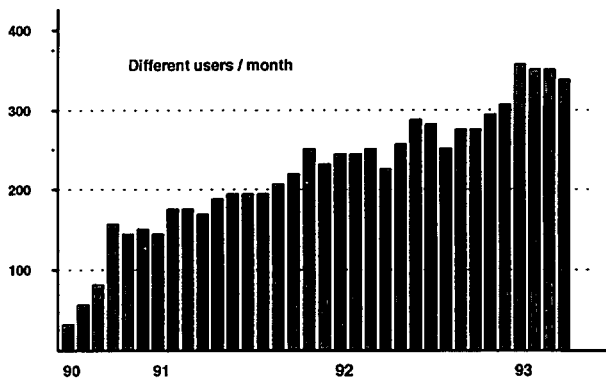
バビロニアの粘土板や敦煌の巻物などは、人間の歴史の研究にとって計り知れない大きな意味を持つものであり、この意味でアーカイブの重要性は疑いの余地はない。またお役所仕事などもすべて前例に基づいて執行されるわけであり、過去の記録の保存は大変重要な作業である。しかし、一般的に保存するとはいっても、どのデータが将来具体的にどのような意味を持ってくるのかは、現時点ではまったく判断できないわけであり、どれが重要なのかについては何ともいえないからすべて細大もらさず保存しよう、ということになる。ここに一種の哲学があると、筆者は思うのである。

この哲学の一つの特徴として、「完璧性」の思想があると思われる。これは前述の「細大もらさず」ということに関連する。現実には何でもすべて残すということとはできないから、ある一定の明確な基準を設けて、それにあてはまる記録はすべて保存する、ということになる。しかし一旦処分してしまった記録は「とりかえしがつかない」ことでもあり、また場合によっては「記録がない」ということも一つの積極的な情報になりうるわけで、この意味で「できるだけ多くまた完璧に」という考え方が成り立つ。また「完璧でなければ意味がない」という考え方にもつながる。このように考える人々からの積極的な支援によってアーカイブ活動は成立している、という面もあると思われる。

しかしひとたび「限られたリソース」ということを考えると、このような考え方はいずれ破綻をきたすであろうことは明らかである。そして最近特に、リソースの問題がシビアになってきていることも事実である。

日本の国会図書館などは、国家的な事業として巨費を投じて地下8階という巨大なスペースを確保したが、これも比較的近い未来にはふさがってしまうだろうということで、すでに関西方面に新たな施設を検討しているということである。本学の図書館はいったいどうなるのだろうか。

図2 SIMBAD への、月毎アクセスユーザ数。
出典：CDS, Lesteven 氏報告より。



リソースは、スペースだけでなく人員の問題もあり、こちらのほうが深刻である。スペースの問題は、「仕方がない」ということで比較的アーカイブ支持者の納得が得られ易いが、人員の問題は「なせばなる」というような考え方に傾きやすい。そして「なしてもなかなかならない」という場合には、積極的な支持者がそっぽをむいてしまうという最悪のケースになるわけである。

6 天文学データセンターの将来について

CDS の、特に SIMBAD の利用については、現状では図2のように増大しつつあり、CDS の基盤はここ当分の間は安泰であると思われる。しかし上述のようにデータセンターの運営はいろいろな問題ををはらんでいるものであるから、現状でも、その将来像について明確なプランを設計し、またそれを内外に示してゆくことが必用不可欠である。以下、筆者なりに検討したことを述べてみたい。

6-1 共同作業組織の拡大

将来プランとして第一に必要なことは、やはり世界的な共同作業組織の拡大であろう。そのためには、データセンターのあり方、データアーカイブの

新しい基準などに関し、明確なイメージを作って各地の天文台等に示してゆくことが必要である。

確かに、このようなデータアーカイブの作業は、中央集中的に行わないと完璧性が損なわれる面もあり、また新しいソフトの開発なども集中的に行うことが必要である。しかし筆者の見るところでは、現状のCDSはどうも一箇所で働き過ぎていて、いずれ息切れをしてしまうのではないと思われる。また中央集中は、支持基盤の拡大の面からも好ましくない。そこで、ある程度アーカイブの基準を柔軟にしてほかの機関の参加を可能とし、またソフト開発の分担作業の方式なども検討してゆくべきであると思うのである。

この点に関し、特に現在コンピュータ施設を拡充しようとしている機関や天文学の発展途上の地域などが、共同作業の可能性が大きいのではないか。データセンターとしての業務は、比較的容易に取り組むことができても対外的な評価を得られるという側面があり、コンピュータ導入の当初に取り組むテーマとして扱いやすいと考えられるからである。ただしこのようなケースでは、パーマナントなデータセンターとして業務を確立できない可能性がある。いずれにしても、絶えず共同研究機関を開拓する努力が必要である。

さらに、internetなどの国際的なネットワークのトラフィックのことを考えても、CDSと同様なノードの機能を持つローカルなデータセンターが今後は複数箇所必要である。これはある程度コンピュータ技術の確立した地域に設置したいが、前項とは逆になかなか実現しないことが予想される。たとえば図書館や博物館などの公的な機関に働きかけるなど、発想の転換が必要かもしれない。

6-2 天文学研究成果の公表

将来プランとしてさらに必要なことは、アーカイブカタログやSIMBAD, ALADIN システムを利用した具体的な天文学研究成果を世に問うことである。これは前述の業績評価の問題に関連することであるが、近年の傾向とし

で、各人の研究が（宇宙の神秘の解明というような）究極の科学的目標の達成にどのように結び付くのかを、具体的かつシビアに問うようになってきている。業務の支持基盤の拡大のために、また研究者自身のためにも、これは大変重要である。

残念ながら筆者のこれまでの研究テーマは、SIMBADのようなCDSの中心的データ（伝統的に恒星に関するデータが中心）との関連が薄く、研究テーマの具体例を直ちに示すことはできないが、たとえば最近欧州の天文学者が集め取的に取り組んでいる「ヒッパルコス」や「ROSAT」などの人工衛星観測研究の関連で大きな仕事はできないだろうか。

この点に関してCDSの現状を言えば、今のところ研究者スタッフも皆、アーカイブの仕事やソフトの開発に取り組んでしまっていて、宇宙的な天文学研究が手薄になっているように見える。ただこれについては皆も十分認識しているようで、いろいろ議論はしているが現状はやるべき仕事が多すぎて……ということであるらしい。外部から客員研究員を呼んできて成果を発表してもらうことも可能であるが、内部の人達の研究テーマの整備も早急に考えねばならない。

6-3 電子出版の取り組み

現在最もルーチン的な作業となっているSIMBADの文献サーチは、これからますます作業量が増大するであろうが、共同作業については逆にますます困難になることが予想される。したがって、これを新しい観点から再検討し、何らかの将来像を明確にしておかなければならない。ここで注目されるのが、前にも少し述べた「研究論文の電子出版」である。

「電子出版」というのは、研究論文の書き方をコンピュータのデータベース中に登録できるような形で標準化し（米国数学会などではこれをかなり以前から実行している）、それをデータセンターに集中蓄積することによって、学会員誰もがオンラインで直ちに論文が入手できるようにすることである。これを印刷出版に比べると、発行までの必要時間の短縮、入手スピードの向上、

経費の削減, 論文検索の効率化等々, どれをとってもよいことづくめといつてよい。しかしまだまだ, どのような標準化を採用するか, セキュリティの問題をどうするかなど, 実現までに解決すべき問題が山積しており, 天文学の世界では実現に今少し時間がかかりそうである。

この電子出版が実現すれば, これによる論文についてはコンピュータによる情報検索が可能であり, SIMBAD の文献サーベイ作業を自動化できると考えられる。これは前記の問題の解決になると同時に, さらにこの論文データをアーカイブに加えれば, このデータベースの需要は飛躍的に高まることが期待される。これはデータセンターの支持基盤の拡大にもなり, またこれに関するソフト開発などの新しい研究テーマが生じて, センターの技術者の活性化にもつながると思われる。

このテーマに関しては, 前述のようにすでに試験的に登録を開始していると同時に, CDS と論文出版社, その他の関係者との間で, 活発な検討が行われている。

以上, CDS の将来について, 筆者のいくつかの考えを述べてきたが, もとより CDS の各スタッフはそれぞれこの点について早くから真剣に考えてきているわけであり, とりたてて第3者がいろいろ心配することもないと思われる。ただ筆者が考えるのは, 日本においても同様な事例がいろいろ有り得るのではないかということである。このことを念願に置ながら CDS のことを考えてみたわけである。

7 日本の天文学とデータセンター

さて, 以上のように, CDS の仕事, および世界の天文学におけるデータセンターの役割等を検討したところで, 日本の天文学界では何ができるか, 何をすべきかについても考えてみたい。ただしこのことは, 今回の筆者のストラスブール滞在のメインテーマであり, 来年3月までじっくり時間をかけて検討すべき課題であるので, ここではまだ結論めいたことを書くことはでき

ない。以下、2つ3つ、現時点で気のついたことを書いてみることにする。

日本の天文学者は、地理的条件および言語の問題があって、国際舞台で外交活動することが苦手な人が多い。したがって、日本に、CDSに代わるような天文学データセンターを設立・維持することは、まず不可能であろう。(外交に関しては、フランス人には天性の能力があるように思われる。)しかし日本は、地理的な条件もあり、ファシリティにも比較的恵まれているので、CDSを中心とする世界の天文学データセンターの事業の一ローカル協力機関として大いに働くべき立場にある、と筆者は考える。

日本でそれが実現するための条件としては、この活動を熱心に推進しようとする才能ある研究者が存在すること、およびファシリティを有する研究機関がこの事業に理解を示してくれること、であろう。最近では、日本でもリソースの分配の問題がシビアになってきており、実現は必ずしも容易ではないと思われるが、さいわい、日本にはこのような事業を容認する風土(!)がまだ残っているように見受けられる。

特に日本の天文学界では、ハワイの大型望遠鏡の建設を軸に、機関の組織・研究グループ・ファシリティなどが大変革の時期の真っただ中にある。この中で、コンピュータ利用およびデータ処理の立場としては、望遠鏡操作のための膨大なソフトウェアの開発作業と同時に、完成後の望遠鏡から出てくる未曾有の膨大な観測データのことも今から考えておかなければならない。ここに現在、アーカイブとデータベース技術の「特需」があるわけであり、この観点で、いまこそまさに世界の天文学データセンターの事業との緊密な関係をうちたてるべき時であると筆者は考える。

このほか、電子出版の問題などは日本は特に恩恵を受けるわけであり、積極的に検討に参加しなければならない、等々、多くの検討事項が残っているが、以下の検討については別の機会にゆずる。

8 おわりに

本報告の骨子であるストラスブール天文台滞在研究は、本学後援会の提供する「中堅教官海外学術研究派遣基金」によって行われているものである。ここに改めて謝意を表する。また筆者の不在中の職務代行をお引き受けいただいた理科エリア各教官、および受入先のストラスブール天文台長 M. クレゼ博士に、記してお礼を申し上げる。

(1993年6月10日、ストラスブール天文台にて)

(一橋大学教授)