

5. 3 素粒子物理学分野における情報の流通

高エネルギー加速器研究機構計算科学センター長
渡 瀬 芳 行

素粒子物理という研究の現場では、情報とは具体的にどのようなものがあり、どう流れているのかを、研究分野の1例としてお話します。そして、大学とは異なる研究所での図書館（部門）に対する期待を述べたいと思います。

1. はじめに一素粒子物理の研究

従来の図書館では、主たるメディアである紙による印刷物としての情報を整理し、提供していく活動が中心であった。しかし、ここ数年の間に、電子的メディアの台頭が著しく、図書館に求められる機能も大きく変わってきている。研究の現場での活動では、それより更に早くから電子メディアによる情報の交換が行われ、いわゆるマルチメディアによる情報の交換になっている。

研究情報の流れを述べるには、研究分野の研究課題は何であって、具体的な研究活動はどんなことを行っているかについて、少し触れなければならない。素粒子物理学は、物質の最も原初的な構成要素は何か、また、それらがどのような法則で、相互に作用しあい、この物質宇宙を作っているのかを明らかにする学問分野である。それは、宇宙の始まりから物質がどのように生成されてきたかを研究することにもなり、宇宙物理学とも密接に関連している。研究的手法には、実験と理論、および近年では、計算機によるシミュレーション研究の3つの様態がある。研究者は、いずれかの専門家

になって研究するのが通例である。実験研究では、大きな加速器を作り、電子や陽子を高エネルギーにまで加速して相互に衝突させ、宇宙初期の非常に高温な状態を再現して、極微の世界での物質の相互作用について研究している。より高いエネルギー状態を得るには、より大きな加速器や実験施設が必要で、その建設には大きな資金をを用意しなければならない。このため、20世紀の半ば頃から、加速器施設の共同建設や運営、また実験装置の建設から運転までを国際的に共同して行うことが慣例となっている。現在、もっとも物理的に大きな加速器は、スイスのヨーロッパ連合核物理研究所 CERN にある、周囲 27 km の円形加速器である。このような国際的な共同研究グループには、世界中から多くの研究者が参加している。最近では、1 研究グループが数百人にもなる。そのため国際的なコミュニケーション環境や情報の交換には、専用のネットワークを構築したり、多くの人材を情報分野に当てるなど、大変重要視している。一方、理論の研究は、基本的には個人ベースであるが、実験と同じ環境で研究しているため、お互いの情報交換は非常に早い。それに基づいて、理論研究のスピードも早くなっている。これが良い結果を生むかどうか別として、研究者は情報の流れには非常に関心をもっている。これは、世界の研究者の総数は、数万人くらいであるが、研究対象がかなり、はっきりしていることにも起因していると思われる。

2. 情報の種類

情報の流通は、研究の対象によって、また研究スタイルによって、種類や量など大きく異なる。素粒子物理であっても、大型の加速器を使って行う実験による研究と純粹に理論的考察にもとづく素粒子論の研究とでも異なっている。研究情報として、どのようなものがあるかを簡単に種類分けしたのが表1である。この中で、計算物理研究とは、大型のスーパーコンピュータをつかって行うシミュレーションにもとづく理論研究である。この場合は、実験データの代わりに、シミュレーション計算によって

得られるデータが、実験データに対応している。研究現場で、研究の進行とともに変動し生きている情報は、3つの研究スタイルによって、異なっている。しかし、一旦、論文となって公開されたもの以降は、ほぼ同じである。どのように研究が行われたとしても、論文の形になったものは、固定した情報といえ、アーカイブされるものである。ここまでは、研究者にとって必要な情報であるが、さらに研究成果を一般社会に還元するとか教育に役立てるとなると、また別に、研究の成果を加工して専門以外の人にもわかる情報にする必要がある。このような一般への広報活動も、基礎科学研究の使命の一つである。

表 1 研究情報の種類

	実験研究	理論研究	計算物理研究
変化し、 生きている情報	日常会話	日常会話	日常会話
	実験情報	国際会議などの情報	シミュレーション手法などの情報
	実験生データ		計算結果のデータ
	実験解析用データ		解析用データ
	実験結果のデータ	実験結果の情報	解析結果のデータ
		他の研究者の投稿論文	他の研究者の投稿論文
固定した情報	電子投稿論文		
	ジャーナル論文		
	書誌データベース		
	研究成果の情報データベース		
一般への広報	WWWによる発信情報 広報誌 教育用教材 一般メディア、新聞情報		

3. 変化し、生きている情報

研究の現場での情報の流れには、固定したものはなく、あらゆる情報交換システムやメディアが駆使される。情報というものは、供給側と受け取る側がないと意味がないが、研究現場で、研究の初期の段階では、供給側と受け取り側が閉じた一研究グループ内である。しかし、ここは、もっとも量が多く、また多様な質の情報がある。実験研究の初期の段階では、ごく限られた需給範囲での情報の流れになっているが、この辺の最近の動向について、少し詳しく述べる。

3. 1. 共同研究環境と WWW

素粒子物理の実験的な研究を高エネルギー物理とも言うように、素粒子の極微の世界を研究するため、電子や陽子を光の速度近くまで加速し、高エネルギー状態をつくる必要がある。そのための大型の高エネルギー加速器実験施設を、国際的に共同して建設し、実験して、さらに実験のデータ解析まで行っている。このような共同研究には、国内外とのコミュニケーションが容易にできる環境が必須である。1980年代から、国際的なネットワークによって、世界中の主要な研究機関および大学を結ぶ専用のネットワークの整備と維持を協力して行っている。初期には電子メールのみであったが、1990年ころより、テレビ会議システムが普及し、国際間での頻繁な打ち合わせをこれで行って来ている。大きな加速器や実験装置を共同で建設するには、日常的に、会議を持ち、物事を決定して先に進むことが必要である。その度に海外に出張することは不可能で、このテレビ会議システムの果たす役割は大変大きい。

同じ、1990年頃、World Wide Web (WEB) が CERN [1] で、発明されて急速に普及していったのは、このような研究現場のニーズに、触発されたものである。これは、情報の共有手段として大変有効で、WEB で図面

とか文書をみながら、研究の打ち合わせをテレビ会議で行うのが、現在では普通のことになっている。ここで言う 1992 年頃のごく初期の WEB には、たった 20-30 の極く限られた高エネルギー関連機関のサーバしかなかった [2]。しかし、1990 年代半ばからは、急速に普及し、インターネットの代名詞として WEB が利用されている。すでに、一つのコミュニケーションのメディアとして、定着した。最初の頃のことを考えると、想像を絶するような、発展のスピードと社会的な衝撃の大きさである。WEB によって、静的な情報の共有は、国際的な共同研究の多くの問題を解決してくれた。情報としては、図面、プレゼンテーション、ドキュメントなど、それまでは、郵便やファックスで交換していたものが、世界中に散らばった共同研究者によって、即座に共有できることは、情報の交換にとって計り知れない恩恵をもたらしている。

しかし、WEB が一つのメディアとして、確立されるには、後日参照が可能な状態に維持していくことや [3]、Indexing とか検索機構などが担保される必要がある。この点は今後の課題である。

3. 2. 実験データの交換と GRID

情報としての実験の生データは、供給と需要が極く限られた共同研究グループの世界でしか有効でない。これは、他の研究分野でも同じであろうが、実験データのようなものは、複雑な実験条件を加味した解析をしない限り、意味を持たないためである。しかし、同じ実験グループと言っても、高エネルギーの場合は、グループメンバーが数百人から、多いグループは千人にもなるため、実験データの交換は、非常に重要なことである。と言うよりも、膨大な量の実験データの解析には、それだけの人員と同時に、世界の各地の研究機関が持っている計算機資源を活用して解析を分担しなければならない。従来は、実験データを磁気テープなどにコピーして、各地の共同研究グ

ループに配布し、分担して解析してきた。しかし、ここ1-2年の間に、国内のみならず、国際的にもネットワークが高速化し、それを十二分に使って計算機資源を接続することで、データを広域に分散して処理することができる環境になって来た。このようなネットワーク上の資源を共有していくためのシステム技術を、電力網になぞらえて、グリッド (GRID) [4] という。このグリッド網にアクセス権があれば、計算サービス、特殊なソフトウェアの利用、データアクセス、データ検索、またそれらを組み合わせたサービスが受けられるようになる。多数の研究機関の共同実験グループの持つ計算資源を協調して運用し、大規模なデータの解析を行えるシステムの構築が可能となる。この技術は、単にデータ解析に留まらない多くの応用が目されている。この技術は、ネットワーク上のハードウェア、ソフトウェアを共有するメカニズムを提供する技術で、装置の遠隔操作による実験装置や加速器の共同運転から、実験データの収集、その解析まで統合したヴァーチャルラ

ボラトリーを構成することも視野にある。

これは、WEB のような静的な情報の共有でなく、クライアントからの要求によって、一次情報を加工して発信する、新たな情報の流れが生まれることになる。従って、実験データとか、シミュレーションデータなどの一次データも加工処理込みで、発信することが可能である。この方向での動きは、未だ、明確になっていないが、近い将来、大きな変化をもたらす可能性がある。現状では、実験の一次データは、極く限られたグループの中に、その利用は限定されている。

3.3. メタデータとしての実験値データ

実験データから求めた特定の物理量について、いろいろな条件パラメータとともにデータベースにした情報がある。素粒子反応の断面積とか、粒子の崩壊率とかの数値データを編集したものである。

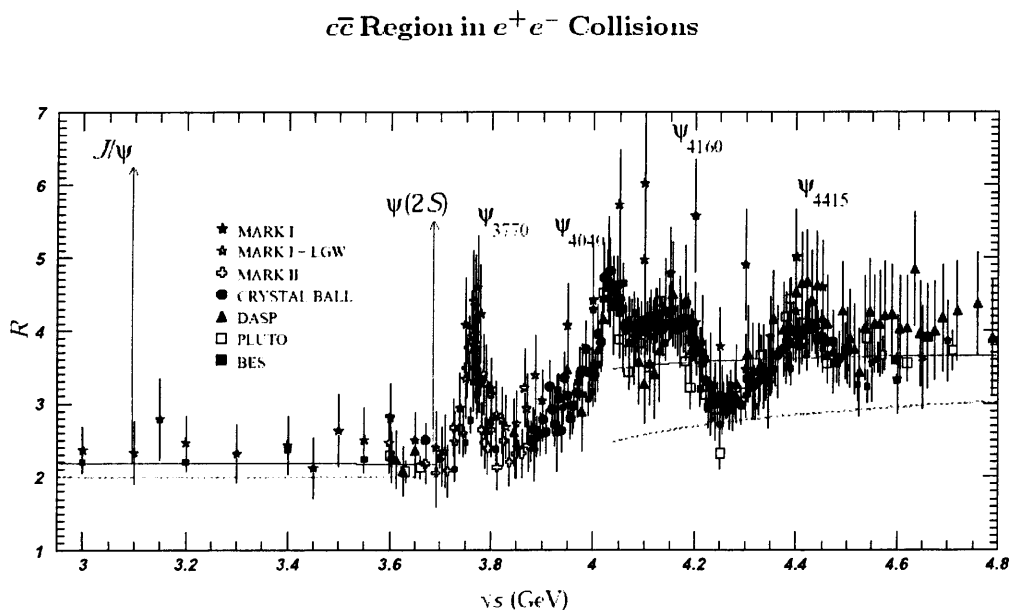


図1 実験結果の収集とその整理の例

電子と陽電子との正面衝突の確率を示す実験データを多くのグループのものを集めて整理して、現在までに蓄積されたデータの最新を示すものである。(Physical Review D66, 2002)

個々の研究者が自分に必要になる物理量だけを集めてデータベースにしていることもあるが、素粒子物理の分野では、世界の主要な研究機関が協力して、人手および資金を提供してデータベースを作成し、維持している活動がある。Particle Data Group (PDG) [5] という団体が、米国ローレンスバークレイ国立研究所 LBNL を拠点に、世界中の研究機関から研究者が交代で、1-2 年間、多数の論文をよみ、そのデータを数値にして、データベースに入れる作業を担当している。日本からも数名参加している。図 1 に、一つの例を示した。実験の生データを解析して、電子と陽電子との正面衝突の確率を示す相互作用断面積の数値を、重心系エネルギーを横軸にして示したものである。異なる実験グループが、異なる時期に、異なる加速器、測定装置で行ったものを収集している。測定誤差を含んで、実験値の平均などを求めることができる。

4. 論文の情報

4. 1 プレプリント情報

実験のデータ解析は、その実験グループの研究成果として、通常は多人数の著者名をつけた論文として発表する。理論の研究の場合は、一人か、数人の著者であるのに対して、実験の場合は、数十ないし数百人にもなる。論文は、ジャーナルに投稿され、審査委員の査読をとおして、印刷されたジャーナルの論文として発表になる。このプロセスは、従来、かなりの時間を要するプロセスである。半年-1 年にもなる。過去には、それでも価値はあったが、関連した国際会議やワークショップなどが毎年、かなりの頻度で開催されるようになると、そのような場で、すでに情報が交換されており、研究のスピードが速い。したがって、

ジャーナルの論文の価値が限定されてきていた。このため、素粒子、高エネルギー分野では、ジャーナルに投稿し、審査が通れば、ジャーナルの発行まえに、自分達で、印刷して関係研究者や主要な研究機関の図書室に、送付することが慣例であった。これをプレプリント論文として、各図書室では、重要な研究資料として扱っていた。受け入れしものの保存、データベース作成、そのコピーサービスなど。図 2 に、高エネルギー加速器研究機構 (KEK) が扱ったプレプリントの受け入れ件数の推移がある。多い年には、6,000 件以上もあった。研究者は、毎日、このプレプリントを閲覧するため、図書室通いが通常であった。研究所が受け入れるプレプリントは、世界で発行されるものすべてが、到着するわけではないため、大きな研究機関同志が協力して、それぞれのデータベースを交換したり、イメージスキャナーにより、デジタル化して、WEB に掲載するなどのサービスが行われた。SLAC では、早くから SPIRES [6] というデータベースが構築され、維持されている。KEK でも、KISS データベース [7] として 1990 年頃から米国 SLAC、ドイツ DESY、スイス CERN、京大基研と協力して行っている。

しかし、図 2 を見ると、1996 年ころから、急速に受け入れプレプリントの数が減少しているのが、はっきりとわかる。これは、次に述べる、論文を電子的に投稿する e-Print が本格的に利用される時期に相当している。このプレプリントのデータベースには、サイテーションの項目もあり、1974 年から構築されて、今では 50 万件に達している。そのため、論文の検索のみならず、研究者の研究業績を見たり、研究経歴を見るのにも利用されている。

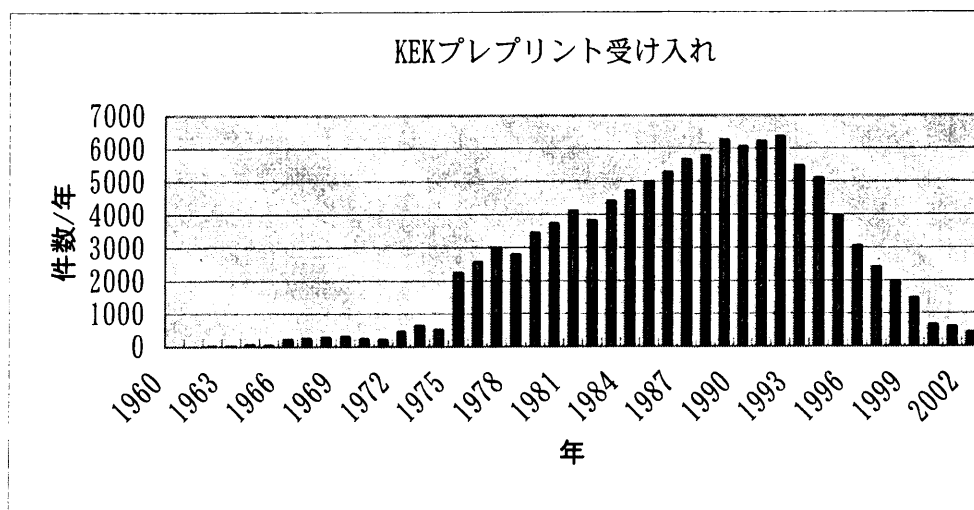


図2 プレプリント受け入れ数の変遷 (KEK 学術情報課)

4.2 e-Print

e-Print [8] は、インターネットが普及するのにもとない、安価で、スピーディな学術情報の交換を促すため、1994 年以来、米国物理学会が主催している。しかし当初は、理論物理のための電子投稿サイトとして、1991 年に、200 名くらいが紐理論と 2 次元重力の研究グループという極く小規模で始められ、立ち上げられている。その後、素粒子理論の全分野と、高エネルギー物理が参加した。1994 年の段階で、25,000 ユーザに提供し、一日、40,000 トランスアクションがあるほどになった。

内容としては、タイトル、アブストラクト、著者、本文などが決められたフォーマットで FTP により転送することで、自動的にデータベースに登録され、他の人は、FTP や Web からアクセスできる。単なるタイトル、アブストラクトだけの投稿は許されていない。論文としての体裁は要求されている。だれでも投稿できることにより、論文の質は問わないことになり、それは読者にゆだねられている。修正もまた、撤回も許されている。どこかのジャーナルに投稿

し、掲載された場合は、その旨、データベースに記入できる。このように e-Print 上の情報は、それぞれの著者によって更新されることが条件である。こうして、e-Print は、紙でのプレプリントに代わると同時に、ダイナミックな学術情報の交換の場になっている。過去に掲載されたすべての論文は、全文データベースとしてアーカイブされて、書誌データベースも備わったものとして、その価値は高まっている。最近の投稿や参照の統計情報も WEB から参照できるように整備されている。その一つの例を図 3 に示した。1991 年の 12 月から、高エネルギー、物性物理、天文の 3 分野について、年ごとの月あたりの平均投稿数を示している。図からわかるように、高エネルギー分野では毎年伸びて、最近では、月平均で、800-900 論文にもなる。

閲覧は、WEB からでもできるが、登録することで、毎日、新しい論文が投稿された情報を電子メールで通知するサービスが受けられる。こうして、常に新しい情報を手に入れることができる仕組みになっている。毎日の接続回数は、16 万件にもなっている。

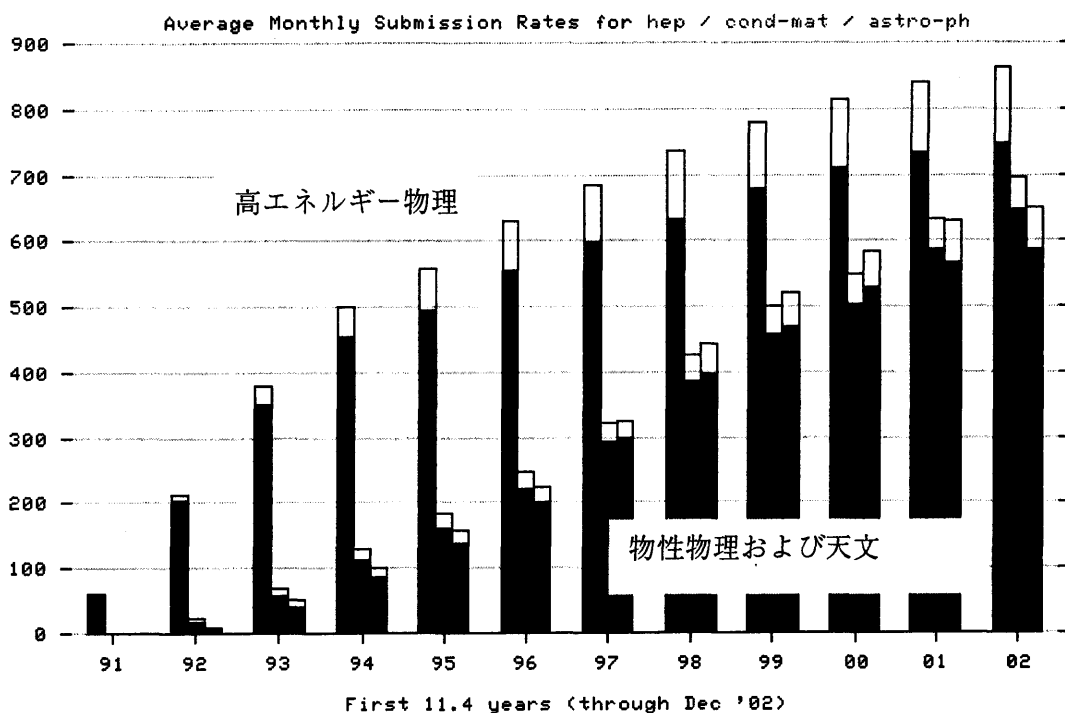


図 3 e-Print の月平均の投稿数の推移

4.3 学術雑誌と電子ジャーナル

e-Print に投稿したとしても、レフリーの査読がないので、そのままでは、客観的な論文の価値は判定されていない。ジャーナルへの投稿によって、レフリーの判断が下されて始めて学会としての評価が決まる。従ってジャーナルのもつ意義は失われている。しかし、ジャーナルも紙での出版のみならず、電子的な情報としての出版が行われている。現在、KEK で利用できる電子ジャーナルには、以下のようなものがある。この中で、*印のものは、国立大学図書館協議会に加盟している総合研究大学院大学の基盤研究機関として利用できるものである。

サービス名 : KEK で関心のある分野

ScienceDirect*: 物理、素粒子、核物理
 LINK*: ヨーロッパ応用物理、数理物理
 Wiley-Interscience: 固体物理

Synergy*: 天文

Kluwer: 低温物理

J-STORE*: サイエンス (1996 年以前)

OJPS: 米国物理学会ジャーナル

IOP: 物理、固体物理

Nature*: ネイチャー

Science: サイエンス

これ以外に、シリーズもので、オンラインで提供されている出版物がいくつかあり、今後は、さらにペーパーレス化に拍車がかかると思われる。したがって、最近では、研究所の図書室に、研究者があまり行かなくなっている。

4.4 国際会議

国際会議に関しては、他分野の会議と同様である。従来から、2 年に一度のペースの大きな会議の他に、特定分野での会議やワークショップが、頻繁に開かれている。この国際会議などは、まさにナマ情報の宝庫であり、新しい共同研究の始まる場所でもある。国際会議のプロシーディングに掲載

される論文も、プレプリントとして、印刷したり、また、e-Print に投稿することもある。

以上が、研究の現場に関連した情報の流れと言える。

5 物理学成果の情報

5. 1 Particle Data Group (PDG)

先にも出てきたが、素粒子物理学での、数値データベースは、Particle Data Group がコンパイルしている。実験からの一次データではなく、物理量の観測値とその誤差を、発表された論文から引用して、総合的に評価している。この素粒子データベースは、世界の主要機関の共同出資により、また、研究者も順次交代しつつ、実験や理論の成果の収集と評価、まとめを行ってきた。

1960 年代後半に、最初に米国ローレンスバークレイ研究所で始められた実験データのコンパイルがもとになり、現在では、米国 DOE, NSF、日本 文部科学省、CERN、イタリア INFN、スペイン、ロシアが資金を出している。

実際の作業は、世界の 80 近い機関から、100 名あまりの研究者がそれぞれの専門の領域について、最近の発表論文を読み、データを検討して、これまでのデータに付け加えていく作業を繰り返している。2002 年版については、667 編の論文から、2,200 点の測定結果を詳細に分析して、データベースに入力している。成果は、米国物理学会論文誌の特別版として、隔年に出版されると同時に、WWW 上でも随時更新された情報を提供している。

PDG では、さらに、縮刷、簡略した手帳版も原則無料で、研究者に配布している。その数は、30,000 冊に上っている。最近では、PDF の Palm 版もある。その内容は、

- 物理量の定義、単位
- 原子、原子核の特性
- 素粒子標準理論
- 素粒子の特性
- 仮定されている粒子に関するデータ

天文、宇宙関連の定数

実験方法

数学の公式

反応断面積の公式

など、多岐にわたり、研究者にとっては、大変便利な手帳である。

5. 2 成果の広報

素粒子物理学は、先端的な学問ではあるが、それがそのまま日常生活に直結するような研究分野ではなく、基礎物理学の一部である。しかしながら、その実験的研究には、巨大な加速器施設と実験装置が必要である。エネルギーフロンティアを目指す、益々大きな施設の建設および運転には膨大な資金が必要である。このような研究を進めるには、各国政府も当然であるが、納税者、国民の支援がないかぎり、不可能である。巨大加速器のような先端的な研究にとっては、その建設には技術的な開発が必須で、それを通して、社会への還元はあるが、一般国民への広報活動は欠かせない。そこで、教育的な広報と一般広報が重要である。そのため、どこの国の研究機関でも、この辺には大変力をいれている。学会誌ではもちろんの事、研究所の広報誌、公開講座、一般公開、最近では、WWW による情報発信にも多様な工夫がなされている [9]。

6 まとめ

—これからの図書館に期待すること—

日常的な情報の検索は、インターネットでの検索で多くのことが得られるが、情報の質は受け取り側の判断に委ねられている。この大量の情報からは期待できない整理された情報は、特別の価値のあるものとして位置付けられる。そのような情報は、組織として構築できる図書館機能に期待される。大学に於ける図書館と研究所の図書館部門とでは、その期待される機能が同じではないけれども、先端的な研究を行っている組織の中での図書館機能は、受身の資料の蓄積でなく、積極的な情報の蓄積と発信がも

とめられる。情報を持っている研究者は系統的にそれを整理、加工してデータベース化することには、必ずしも積極的でない。しかし、各研究所での図書館機能には、内外との系統的情報のセンターとしての役割が期待されている。そのためには、研究者とのタイアップが重要であるし、研究者を通じた国際的な連携も行うことが必要である。来年度からは、大学も含めて法人化が行われるのであれば、それぞれの組織に、ふさわしい図書館機能の運営ができるはずである。外国の研究所などでの図書館の運営を見ると、その方面の研究者ではないけれども、比較的長期の在職の職員は、その間に多くの情報を手がけ、成果をあげているのが見られる。このような人を中心に図書機能が充実されることが望まれる。

図書館は、蓄積、整理された情報の発信、交流の拠点となることが期待される。

参考文献/ URL

- 1) CERN, <http://www.cern.ch/>
- 2) <http://www.ibarakiken.gr.jp/www/main.html>
- 3) Internet Archive,
<http://www.archive.org/>
- 4) The GRID: Blueprint for a new
Computing Infrastructure; I. Foster
and C. Kesselman, ISBN-1-5586-475-8
- 5) Particle Data Group,
<http://www.pdg.org/>
<http://ccwww.kek.jp/pdg/>
- 6) SPIRES.
<http://www.slac.stanford.edu/spires/>
- 7) KEK Information Service System.
<http://www-lib.kek.jp/top.html>
- 8) e-Print Archive, <http://arxiv.org/>
- 9) Kid's Scientist,
<http://www.kek.jp/kids/>