

ネットワーク情報源の「組織化」について

藤岡 豊

A. 現状について

図書館ホームページにおいて「ネットワーク情報源」が公開されたのは、図書館ホームページがリニューアル（平成11年11月11日）した後の平成12年2月のことである。それまでの図書館ホームページにも、情報検索サイトや文献情報サイト・図書館関連サイトへのハイパーリンクと、それらについての簡単な紹介文を収録する「リンク集」があったが、それをさらに発展させて、図書館が提供するオンライン方式（ネットワーク方式のCD-ROMをも含む）の各種データベースやオンラインジャーナルといったネットワークで学内共有される各種データベースの紹介を織りまぜ、ネットワーク上のさまざまな学術情報が簡易なナビゲーションに導かれてアクセスできる総合的なポータルサイトとして位置づけられるコーナーとなった。また、スタートページには「Net News」と題するコラム欄において、新規に登録される情報サイトのあらましが並ぶことで、紹介記事の掲載が追加されていくのがわかるようになって

ている。

中身となるそれぞれのカテゴリーの情報源リストのページにおいては、単に紹介しているサイトへのリンクだけでなく、その紹介を行う文中にも情報サイトに関する有用な関連情報へのリンク（例えば、情報を提供している組織・機関のウェブページとか検索機能の使い方の説明があるウェブページへのリンク等）や類似する他の情報源へのリンクを併せて盛り込むことで、情報探索を行う前に知っておきたい予備知識的な内容をも、まとめて多角的に入手できるように配慮している。

このように各アイテムのリンクからそのまま探索ツールや一次情報がアクセスできる情報ランチャーを通じて、コンピュータとそのネットワークを用いてその場で手軽に文献や情報の探索を展開するという調査活動の新しい在り方を提案することは、今後図書館が提供するサービスとして重要な役割を持つ内容だと思われる。



図：ネットワーク情報源のスタートページ
 (<http://www.kansai-u.ac.jp/Library/netresource/nwlink.html>)

ところで、当初は千里山と高槻両キャンパスに在るCD-ROMサーバーから提供されるネットワーク方式のCD-ROMデータベースしかなかった学内共有型データベースも、2000年秋季より次々にウェブによるインターネットを用いたオンラインデータベースの利用提供を開始したため、急速に収録数が増加し、有力な検索ツールが揃うようになった。しかし同時に多くの情報源の紹介文を羅列することになり、個々の情報源が埋没してしまつて目的とする検索ツールが探し出しにくくなりつつあり、より効果的なナビゲーションを持つインターフェイスへの改善や内容の見直しが求められているのも事実である。

B．今後の課題としての「組織化」について

今後「ネットワーク情報源」では、これまで棚上げにしてきたシステムティックな論理性や形式性に重きをおいた構造的な意味での「組織化」に取り組むことで、分かりにくさや不便さを改善していく必要があると思われる。「ネットワーク情報源」の編集作業全体を指して「ネットワーク情報源を組織化する」と言うことがあるが、これからは後述する構造的な「組織化」という意味に限定してこの言葉を使うことにする。以下は、これまで検討されてきた「ネットワーク情報源」の今後の在り方についての主要なポイントをとりあげることで、目指している「組織化」とその問題点を概説したいと思う。

1．学術情報ポータルとして最適化した主題分類の細分化

現在の「ネットワーク情報源」の左側フレームにはカテゴリーが2階層で並んでいる。この分類体系では「データベース」「オンラインジャーナル」「機関・団体」「各種情報」の4つに分けられ、その中にサブカテゴリーとして「データベース」に主題分類的なカテゴリー名が多く並んでいる。これは、当初インターネット上の情報源で優先的にとりあげてきたものがほとんど文献データベースであったため、そこからカテゴリーを分化したことからこのような偏りが生じている。

近い将来このような分類から、本来的な主題分類によるカテゴリー群があらゆる分野を網羅するように全体的な布置を組み替えて配列すべきだと考えている。こうした機能を持つものの例として、「サブジェクトゲートウェイ」と呼ばれているような一般の方に親しみやすい主題分類による総合型リンク集

のページがある。データベースによる検索では対応できないようなテーマによる探索をアシストするディレクトリ型索引(例として、Yahoo[<http://www.yahoo.co.jp/>]などサーチエンジンのトップページに見られるような一目で全体を見渡せるくらいのカテゴリー群のリストを作り、そこから分化して2階層目のカテゴリーへ入っていき、目的のテーマに該当するサイトリストに辿り着くというアプローチを行うもの)が理想として考えられる。しかし、その分類に用いるカテゴリーには図書館で一般的に用いられている日本十進分類法などの図書の分類をそのまま適用するか、もしくは独自の学術分野カテゴリー体系を構築するのかといった基本的な選択で苦悩するところだ。また、そのようなカテゴリー体系で全分野を網羅しても、さしあたってはカテゴリーに合致するサイトが存在しないような領域がまだまだ残っており、全体の均衡を欠いてしまうことが予想される。そこで、そういう偏りがなくなるよう、エントリーされた情報源が少ないカテゴリーを重点的に調査してアイテム数をてこ入れしたり、他の似た内容のカテゴリーと併合するといった調整をしなければならないなど、準備のための時間が必要となる。

2．データベース構築によるサーチ機能の導入

収録する情報源のアイテム数が一方的に増加していく状況にあるため、いくら前述したようにテーマを整理しなおしても、アイテムが多量になってくるとカテゴリーを階層化したディレクトリを辿って目的の情報源を見つけるのは困難になる。そこでキーワードで検索し、多くの中から必要なものだけを抽出するデータベースによるサーチ機能が求められるようになる。データベースの構築のためには、

図書について一定の書誌記述のもとで目録データを作成するように、インターネット上のサイトの情報源についても個々の情報源に関する索引データを作成すること。その索引データを対象に検索を行うサーチエンジンを用いて利用者が求めるキーワードとマッチングしたレコードを抽出して出力し、目的のリンクとその紹介文等を提示できるようなシステムを開発すること。といった取り組みが必要である。の索引データ作成については、従来から図書館が行ってきたような図書の書誌記述がウェブ上の電子情報に対して適用しにくいと、ウェブページのメタデータ(データについてのデータ=個別の情報源についての索引)の表記において

| | |
|--------------------|--------------------------------------|
| Title | 情報源の名称 |
| Creator(Author) | 情報作成者(情報を記述している人・団体名) |
| Subject(Keyword) | 情報内容の主題(件名・分類コード等の統制された表記) |
| Description | 情報内容の記述(アブストラクトや目次などの自由記述) |
| Publisher | 情報提供者(情報を利用公開している人・団体名) |
| Contributor | 情報寄与者(情報内容の制作に関与した人・団体名) |
| Date | 情報公開日 |
| Type | 情報資源のタイプ(ジャンル・機能を示すカテゴリー) |
| Format | 情報の形式(物理的な特性とそれを表示するのに必要なソフトやハード) |
| Identifier | 情報内容を一意に特定できる識別子(URL や ISBN/ISSN) |
| Source | 情報を作り出す元になった別の情報資源に関する情報(参照文献) |
| Language | 言語(言語コード・国名コード) |
| Relation | 関係する他の情報資源の識別子とその関係(版、章、翻訳といった関連性) |
| Coverage | 情報が表わしている内容の特性としての空間的な範囲や時間的な範囲 |
| Rights Management | 情報資源に係る権利管理に関する記述(著作権や知的財産権) |

表：Dublin Coreの基本15エレメント

有力視されている、米国OCLCで研究されたネットワーク上での流通に耐える汎用性を持ったスタンダードな記述規則「Dublin Core(Element Set)」¹を用いることが検討されている。

しかし、冊子体の資料と異なり、特にウェブ上の資料は資料単位の曖昧性や著作権の曖昧さ、更新内容や頻度の曖昧性といった性質により、対象となる項目を確認するのに大変な困難を伴うことが予想される。例えば、有意な検索結果が得られるためには、上表のエレメントのうちの「Description(情報内容の記述)」において、ページ内の内容をどれだけ書き出しておかなければならないのか判断に苦しむところで、それが膨大なデータ量になる場合も考えられる。また、せっかくデータをこしらえてもすぐに新しい内容に更新されて、データのメンテナンスが大変煩雑になることも考えられる。

なかにはこのように言うと、なにもこのような手間をかけなくても、ウェブ上の情報ならYahooやGoogle〔<http://www.google.com/>〕といったインターネットサイトのサーチエンジンで検索できるのではないと思われる向きがあると思う。しかし、「ネットワーク情報源」で目指すサーチ機能は異なるところにポイントがある。これらサーチエンジンの検索の中身であるが、Yahooは独自の整理された索引データを使っているが、それ以外の多くのインターネットサーチエンジンは、対象となるサイトについての情報の索引データを整理しないまま(実際は「ロ

ボット」と呼ばれるシステムで全文取り込みを行い索引を自動作成している)、いわばまるごと全世界のウェブ上に公開されているありとあらゆる情報を検索対象にして、入力されたキーワードとマッチングする情報を抽出していると言ってよい。しかし、そのようなやり方では検索したい内容に深く関連しているものも関連性の低いデータも十把一絡げに取り扱われるため、検索結果件数がひどく膨大な数になることが多く、その中から必要とする情報をさらに吟味しなければならない。際限なく膨張する混沌としたネットワーク上の情報源のドラスティックな状況の中で、「Dublin Core」に見られるような索引データ作成を行う根底には、情報に人為的なメタデータ付与を行うことで、中性的・匿名的になりがちなネット上の情報の在り方から、図書資料のように主題や作者といった固有的な特性が優位性を取り戻すようにデータを再編したいという指向性がある。「ネットワーク情報源」のデータベースの構築は、こうした索引データのサーチに特化することで、検索結果内容の実効性において学術情報が理想的な状態でヒットするように最適化しようとする取り組みとして理解していただきたい。

3. サイト情報収集の体制

現在のところ館員で構成されるウェブサイト運営委員会の5名でウェブサイトの情報源を探索し登録を行っている。前述したように現行の「ネットワー

ク情報源」がレファレンスツールにスポットを当てた構成になっているため、委員会の5名以外ではレファレンスツールを扱うことの多いレファレンスカウンターの係員からもその方面のサイト情報をもたらうこともある。しかし、雑誌出版情報のサイトやそれらのカテゴリーに属さない一般的な各種学術情報のサイトもかなり増加している状況において、それらの収録にも積極的にフォローしてゆきたいところだが、なかなか着手できないままである。また、「ネットワーク情報源」のスタートページにおいて図書館員以外の学生・研究者にも収録を希望するサイトについての情報提供の協力を呼びかけているが、それらの情報が寄せられることがあまりなかったことも事実である。また、特にこれからデータベースへ発展させるためにはデータ作成などの構造的な「組織化」の作業も必要になってくるので、サイト情報の登録のプロセスが複雑になる。これについては図書や雑誌の目録データで既に行われているように、レディメイドの索引データを導入するという選択肢もある（OCLCのCORC²など）が、そこで用いられているデータ形式にまだ検討すべき余地が残るため採用には時期尚早の感がある。当面は実験的な試行錯誤をしながら、利用者にも自信を持って情報探索のポータルサービスを提供できるクオリティを保つよう、館員が手分けして取り組んでいくべき課題だと思われる。

×

×

インターネットの普及でデータベースによる情報検索は、単に学術的用途やビジネスの範疇にとどまらずに身近な日常生活の中にも浸透してきている。そうした状況をふまえ、ネットワーク上の情報源においてこういった内容の情報の入手が可能なのかというアベイラビリティを整理し、その進展していく状況について絶えず確認し、それらを「ネットワーク情報源」というひとつの情報ナビゲーションのタブローに登録＝定着して利用公開することで、図書館が担うべき学術情報のディストリビューターとしての役割の一端として貢献することの意義を再確認して、これを支えるより充実したサービス提供の体制への積極的な協力と理解を呼びかけたいと思う。

- 1 Dublin Coreと図書館におけるメタデータ制作については図書館情報大学・杉本重雄氏の論考に詳しい。
http://www.dl.ulis.ac.jp/DLjournal/No_14/1-sugimoto/1-sugimoto.html はそのうちのひとつ。
- 2 OCLC CORC(Cooperative Online Resource Catalog) 共同で作成するインターネット上の情報源の目録データとその制作体制。 <http://www.oclc.org/corc/>

[参考文献]

電子資料の組織化 / 日本図書館協会目録委員会編 2000 5
ISBN:4-8204-0003-7

(高槻キャンパス事務室 ふじおか ゆたか)