

6. 電子図書館システムの運用と課題

ここでは現在の電子図書館の運用と現在までに明らかになった課題について述べる。

6.1 電子図書館システムの稼働の現状と課題

電子図書館システムの運用を開始してから、すでに1年余りが経過した。以下では、システムの稼働の現状と課題について述べる。

6.1.1 稼働の現状

稼働状況は参考資料2を参照。

6.1.2 システム稼働における課題

システムのスムーズな運用のためには、組み合わされた機器が相互に最適な状態として稼働することが重要である。そのため、さまざまなパラメータのチューニングが必要であり、現在も運用しながらそれに基づく改良を行っている。現状では、こうしたKnow Howの蓄積が重要であり、必ずしも最適な状況であるとは言い難い面もあるが、これまでの経験から少しずつ最適の状況に近づきつつあると考えている。

運用に当たっては詳細なパラメータの調整を行っている。たとえば、当初システムでは、一次情報入力作業と閲覧用のデータを同一のファイルサーバに格納していた。しかし、これらのデータに対するアクセスの挙動は大きく異なり、これら2つを分離することにより個々のアクセス速度などが向上している。また、複数のファイルサーバで構成されているため、これを検索サーバからアクセスすることで、利用者に対する情報の提供を行っていたが、個々の情報を検索サーバ経由で利用者に提供していることのオーバーヘッドが見られた。そこで、検索作業用情報と閲覧用情報の格納場所を分離し、閲覧データを直接ファイルサーバから提供する形態に変更することで、全体の性能向上が見られている。

6.2 一次情報蓄積システム

これまでの実運用経験を通して、情報を格納する一次情報蓄積システムでは、以下のような問題が明らかになってきている。

6.2.1 入力作業領域の確保

データの入力作業では、データの読み込み、一次情報の編集、書誌情報及び目次情報、補助情報の入力編集など、高速なデータアクセスが要求される。そのため、データの入力作業領域として、十分な容量を持つ高速な記憶装置が必要となり、実際に利用者がアクセスしている情報が格納されるべき領域と競合することになる。

この問題を解決するために、入力作業専用の領域を設け、入力作業と閲覧サービスが相互に干渉しない様になっている。(参考資料5-6参照)

6.2.2 階層構造の記憶配置システムにおける情報の配置

一次情報蓄積システムでは、非常に大きな容量を持つ記憶システムを構成するため、ディスクアレイ、光磁気ディスクジュークボックス、磁気テープジュークボックスによる階層構造の記憶システムを用いている。ここで情報の配置は、その利用頻度などの利用状況、情報の持つ意味(インデックス情報など)などに基づいて自動的に決定されるようになっている。しかし、利用者のアクセスは各利用者個々によって大きく異なり、自動配置のアルゴリズムの選択が必ずしも最適になるとは限らない。

今後は、これまでに集積された情報に基づいて利用者の行動を分析し、最適なアルゴリズム選択を行うとともに、変化する利用者の行動パターンにしたがって動的にアルゴリズムを選択する仕組みの検討を行わなければならない。

6.2.3 情報の格納形態

現在の電子図書館システムにおいては、情報はページ単位の情報とそれを統括する管理情報という形態で格納されている。しかし、こうした形式はPostScript(学位論文等)やPDF(Portable Document Format:一部の電子形式で提供される冊子体情報)といった文献単位、書籍単位で提供される情報との整合性が悪い。現状では、これらの情報はページ単位の情報に分割されて格納されているが、本来これらの情報は文献単位、書籍単位で管理されるべきであり、情報の格納形態として再検討が必要である。

平成8年度において、こうした問題を解決するため、ページ単位での情報だけでなく、文献単位、書籍単位の受け入れも可能なようにデータベースの構造を再編成し、新たなシステムの運用を開始している。(参考資料5-13.15参照)

6.2.4 データの保全

電子図書館において、一次情報蓄積システムに格納されている情報が、通常の図書館に格納される図書に相当し、これらの情報が事故などによって紛失することは避けられなければならない。しかし、現行のシステムにおいては単純なテープによるバックアップ機能しか用意されておらず、個々の情報単位(例えば書籍など)の管理を行うことが困難となっている。そのため、電子図書館の情報の保全システムとしては自由度が少なく、不十分である。

そこで、個々の情報単位での情報保全管理が行われるシステムとしてCD-R(CD-ROMに書き込む装置)などによるバックアップ機能の実現を検討している。(参考資料5-4参照)

6.3 情報検索

現行のシステムにおいて、利用者から見た問題点は検索速度とパーソナライズの問題である。(参考資料5-12.14.16.18参照)

6.3.1 検索速度

電子図書館システムに格納される情報が増えるにしたがって、その検索速度は低下していくことになる。特に本システムでは、全文検索機能を実現しているため、その傾向は顕著である。これは、検索サーバに用いられている計算機の能力不足も一因であるが、データベースの構成、データベースソフトウェアのマルチスレッド化など、現状のシステムにおいても検討すべき点も残されている。

インターネットで広く利用されている検索エンジンに用いられる技術の導入とともに、機能分散、負荷分散を進め、良い快適な反応速度が得られるよう改善を行っている。(参考資料5-5.11参照)

6.3.2 パーソナライズ機能

通常の図書館から借りた書籍では、しおりを挟んだり、付箋などによりメモを添付したり複数の書籍の関係について記録を残したりすることができる。これに対し、現在の電子図書館システムでは、利用者全員に対して画一的なインターフェイスを提供するのみであり、こうした個々の利用者が持つ情報との統合はできない。

しかし、電子図書館が真に有益な情報システムとなるためには、個人が持つさまざまな情報との統合が必要であり、電子図書館はそうしたシステムへの情報提供を行うサーバとなるべきである。そのためには、電子図書館システムのような情報提供システムにおけるプロトコルの標準化が不可欠であると考えられる。(参考資料5-7.10.17参照)

6.4 電子図書館システムの運用

電子図書館システムは、学内の全学情報環境システム(曼陀羅システム)、キャンパス情報ネットワーク(曼陀羅ネットワーク)と密接な関係を持っており、これらシステム・ネットワークとの整合性を維持しながら運用されなければならない。そのため現在の電子図書館システムの運用は以下のように行われている。

・ 電子図書館システム機器の管理運用 (情報科学センター)

ネットワーク設定、利用者情報の管理、システム設定など、電子図書館システム自身の管理は、曼陀羅システム及び曼陀羅ネットワークとの整合性を維持しながら行うため情報科学センターによって行われている。これには、情報のバックアップ作業、ファイルシステム構成の決定及びファイルシステム配置が含まれる。

・電子図書館システムに格納される情報の入力・管理（学術情報課）

電子図書館システムに格納される情報そのものの管理は、学術情報課を中心に行われている。これには、冊子体情報の入力、ビデオ情報の入力、入力された情報の管理、図書等の購入処理、電子化されていない資料の管理、貸出・返却管理などが含まれる。

電子図書館システムの運用には、各機器を扱うための技術的知識と、図書館として情報を収集し管理を行うための司書的知識の双方が必要とされている。特に、こうした先端的なシステムを管理運用するためのスタッフに要求される能力は、単に技術的・司書的なものだけでなく、研究開発及び将来の発展の方向をも見据えた能力が必要である。

現在、2つの組織（情報科学センター及び附属図書館）で運用している電子図書館の運用機能を見直す必要も今後出てくるものと思われる。特に、電子図書館システムの高度化及び運用支援を行っている情報科学センターには、これらの課題に対応できるスタッフの拡充が緊急の課題であると言える。