国立国会図書館デジタルアーカイブポータル の現状と今後

—Web2.0世代のデジタルアーカイブポータルの提供を目指して—

久古聡美¹; 吉田 暁¹; 中山正樹²

著者抄録 : 国立国会図書館が目指している「国立国会図書館デジタルアーカイブポータル」は、広く国のデジタルコンテンツ全体へのナビゲーションを行うことを目指すものである。平成19年度以降の稼働に先立ち、プロトタイプを構築し「統合検索」のために必要な機能の検証を行ってきた。本稿ではプロトタイプの概要を示し、さらに今後の方向性について述べる。デジタルアーカイブを保有する各機関にWebサービスによるサービス提供機能の実装を呼びかけるとともに、Web2.0時代にふさわしいポータルの構築を目指し、機械的な連携とデジタル情報の一層の利活用を図る。

キーワード : デジタルアーカイブ, ポータル, ワンストップナビゲーション, ダブリン・コア, 辞書, Web2.0, OAI-PMH, SRW, RSS, CGM, Webサービス, SOA

Current status and future of “NDL Digital Archive Portal”

—Portal service of digital archives in the “Web 2.0” era—

KYUKO Satomi¹; YOSHIDA Satoru¹; NAKAYAMA Masaki²

Author Abstract : The National Diet Library (NDL) is planning to construct “NDL Digital Archive Portal”, which aims at navigating users to any digital information in Japan. Prior to its full-fledged operation scheduled in fiscal year 2007, we build the prototype system and have verified the functions required for integrated retrieval. This paper will show the outline of this prototype system and our portal's possible development in the future. We encourage the institutions who have digital archives to enhance their functions for Web services and make an effort to establish the portal designed for Web 2.0. Through these activities, we will achieve the connected systems among various institutions and improve the usability of digital information in Japan.

Key words : digital archive, portal, one-stop navigation, Dublin Core, dictionary, Web2.0, OAI-PMH, SRW, RSS, CGM, Web service, SOA

原稿受理 (2006-07-28)

(情報管理 49 (6): 313-323)

1. はじめに

国立国会図書館（以下「当館」という）では、平成10年度策定の「電子図書館構想」に基づいて平成16年度に「電子図書館中期計画2004」を策定している。この中期計画は、当館が策定時から5年程度をめどとして達成すべき電子図書館サービスの

具体的方向とその実現に必要な枠組みを示したものである。その中で、当館は国のデジタル情報全体へのナビゲーションを行う総合的なポータルサイトとして、「国立国会図書館デジタルアーカイブポータル」（以下「ポータル」という）の構築を目指すこととした。これは、当館の所蔵資料のみならず、当館が所蔵していない情報に関しても所蔵

¹ 国立国会図書館 総務部企画課電子情報企画室（〒100-8924 東京都千代田区永田町1-10-1）Email: dap-info2@ndl.go.jp

² 国立国会図書館 総務部企画課 課長補佐（〒100-8924 東京都千代田区永田町1-10-1）

¹ Digital Information Planning Office, Planning Division, Administrative Department, National Diet Library (1-10-1 Nagata-cho Chiyoda-ku, Tokyo 100-8924)

² Assistant Director, Planning Division, Administrative Department, National Diet Library (1-10-1 Nagata-cho Chiyoda-ku, Tokyo 100-8924)

元へ案内する役割を果たすものであり、各機関が提供しているデジタルアーカイブ内の有用な著作物を、「いつでも、どこでも、どこにあって」、ワンストップで検索、閲覧できるようにして、資料の利活用が促進されることを目指したものである。

当館では、当館を含む公共図書館の蔵書目録を一元的に検索できるようにする総合目録ネットワーク事業を実施しているが、デジタル情報に関しても同様の仕組みの提供が必要である。当館は、近代デジタルライブラリーや貴重書画像データベース等において、蔵書のデジタル化を進め、インターネットで提供している。また、インターネット情報選択的蓄積事業（WARP）では、各機関がインターネット上で提供している情報を収集し、長期保存している。しかし、リアルタイムに更新されるすべてのインターネット情報を当館が収集し、提供することは現実的ではない。

また、今までの検索エンジンでは、データベース内の著作物を直接検索することは困難であった。ポータルは、その各機関が保有するデジタルアーカイブ内のコンテンツを中心に「統合検索」し閲覧できるようにするものである。

「統合検索」とは、複数のデジタルアーカイブシステムのコンテンツを検索対象とし、次に挙げる点の実現を想定しているものである。

- ・一つのキーワード入力画面（検索窓）から検索できること。

- ・検索結果はデジタルアーカイブの所在や様態によらず統合的に取り出すこと。

- ・可能な限り検索結果から利用者が目的とする情報そのもの（一次情報、または、紙媒体の資料はその入手元）までたどりつけるようにナビゲートすること。

さらに、ポータルサイトの機能として、利用者の情報行動を支援するためのさまざまな機能や仕組みを充実させていく必要がある。一つの方向性として、Web2.0の概念を取り入れる可能性を考えている。

当館では、平成19年度以降の稼働を予定しているポータルの構築に先立って実験システムであるプロトタイプシステム（以下「プロトタイプ」という）を構築し、その試験提供を行ってきた。本稿では、まずプロトタイプ構築に関する概要と主な機能およびそこでの検証成果について紹介する。次に、Web2.0時代にふさわしいポータル機能の提供および今後の方向性に関する考察を行う。

2. プロトタイプシステムの構築

プロトタイプは、ポータルの構築に先立って、サービスの有用性および適用技術の妥当性に関する検証を行うために構築し、試験提供を行うものである^{注1)}。平成16年度から開発に着手し、平成17年7月には一般への試験公開を開始した。

図1 プロトタイプのトップページ

公開後は、試用していただいた方々からご意見をいただくとともに、検索機能等について開発の立場から次々と課題等が洗い出されてきた。平成17年度は、これらの検証成果を踏まえて機能強化を進めた。新たに機能を開発するとともに、問題点の改善を行い、随時プロトタイプ上で追加公開していった。検索対象は、他機関のご協力を得ながら順次追加していき、現在は、次に挙げるデジタルアーカイブシステム全14種類が検索対象となっている。〔 〕内は、所蔵機関またはご協力いただいた機関)

●書籍、出版物、目録等

- ・青空文庫 [青空文庫]
- ・近代デジタルライブラリー [NDL]
- ・NDL所蔵貴重書サンプル (サンプル的にデジタル化したもの) [NDL]
- ・貴重書画像データベース [NDL]
- ・国立公文書館デジタルアーカイブ・システム [国立公文書館]
- ・新書マップ [NPO法人連想出版]
- ・NDL蔵書目録 (和図書・和雑誌) [NDL]
- ・NDLプランゲ文庫 (日本占領期検閲資料コレクション) 雑誌・新聞目録 [NDL]
- ・NDL雑誌記事索引 (2005年分のみ) [NDL]

●有用サイト、データベース情報

- ・府省所管のデジタルアーカイブサイト情報 [内閣官房]
- ・Dnavi (データベース・ナビゲーション・サービス) [NDL]

●郷土情報

- ・デジタル岡山大百科 [岡山県立図書館]

●その他の参考情報

- ・カレントアウェアネス [NDL]
- ・レファレンス協同データベース [NDL]

本格システムの検索対象は、プロトタイプの検索対象に加えて、有用なデジタル情報を保有する機関であれば業種業態を問わず連携し、順次拡大していきたいと考えている。

2.1 プロトタイプの各機能

統合検索では、キーワード検索に加えて自然文から検索する連想検索や分類を通じた検索など、複数の角度からデジタルアーカイブが検索可能となっている。各機能を用途に応じて使い分けることで、情報探索の幅を広げることがねらいである。

ここでは、プロトタイプで提供している統合検索の各機能について紹介する。

2.1.1 統合検索 (キーワード検索)

キーワード検索は、利用者が入力したキーワードに合致するデータを含む資料を検索する。

例えば、キーワード「吾輩 猫」を入力して検索すると、検索結果の一覧表示 (図2参照) には、「青空文庫」と「近代デジタルライブラリー」の「吾輩ハ猫デアル」がヒットする。通常この二つの結果を得るには、おのこのサイトを訪れて別々に探す必要がある。しかし、統合検索では、このような手間を掛けずに、一つの検索窓から複数のデジタルアーカイブシステムのメタデータを一度に検索し、結果を一覧表示する。

検索結果一覧表示からは、詳細情報表示画面へ遷移可能である。ここでは「近代デジタルライブラリー」を例に見ることにする (図3参照)。

この画面で表示される情報は、デジタルアーカイブシステムからポータルが取得したメタデータで構成される。左側にはDublin Coreをベースにしたメタデータ要素が並ぶ。このうち「URL」は、デジタルアーカイブシステム (この場合「近代デジタルライブラリー」) での書誌情報画面へのリンクとなっている (図4参照)。書誌情報画面には一次情報へのリンクがあり、利用者はこれをたどって一次情報を閲覧することができる (図5参照)。

このように、利用者はキーワードを入力し検索を行うところから、一次情報の閲覧までナビゲートされたことが分かる。「青空文庫」の場合もほぼ同様の画面遷移を経て、テキストファイル化された資料の本文を読むことができる (図6参照)。

キーワード検索には、辞書による検索支援機能

1	吾輩ハ猫デアル 影法師著 近代デジタルライブラリー	作成日:2002-06-19 »»» 所蔵サイトへ	猫 《統制語》 ☐ 猫
2	吾輩ハ猫デアル 夏目 漱石 青空文庫	作成日:1999-09-21 »»» 所蔵サイトへ	《読み寄》 ☐ ネコ ☐ Nekko
3	吾輩は猫である 夏目 漱石 青空文庫	作成日:1999-09-21 »»» 所蔵サイトへ	
4	僕の昔 夏目 漱石 青空文庫	作成日:2005-10-11 »»» 所蔵サイトへ	
5	吾輩ハ猫デアル 夏目漱石著 近代デジタルライブラリー	作成日:2002-06-19 »»» 所蔵サイトへ	
6	漱石全集 第2巻 夏目漱石著 夏目 漱石 (1867-1916)		

図2 検索結果の一覧表示

図3 詳細情報表示画面

連想検索は、文章などの自然文から関連する資料を検索できるのがその特徴の一つである。キーワード検索と比べると、入力された文やフレーズ

図7 辞書による検索支援の表示画面

図8 連想検索の検索窓

図9 連想検索の検索結果一覧

を単語単位に分割して各単語に関連度の高い語を含むデータを連想的に検索するため、キーワード検索では検索されない(該当する単語を必ずしもすべて含まない)ようなデータを検索することが可能である。一方では、指定した単語すべてを含むデータのみを取り出したい場合には、連想検索よりもキーワード検索を利用の方が、より期待する結果一覧が得られる可能性が高くなると考えられる。

2.1.3 統合検索(分類による検索)

分類による検索では、日本十進分類法(NDC)に従い階層的に整理されたカテゴリのリンクをたどることで、資料を検索する。

NDCの最上位の分類が表示されている画面で、探している資料に最も関連があると思われる分類を選択すると下位の第2分類が表示される(図10参照)。同様に続いて第3分類へ進むと、その分類に該当する資料が一覧表示される(図11参照)。

また、第2分類以降は、選択した分類に属するデータに対してキーワードによる絞り込み検索を行うことができるようになっている。

2.1.4 その他

この他、統合検索機能として「新着コンテンツ」「更新コンテンツ」「アクセスランキング」「キーワードランキング」がある。

図10 分類(NDC)による検索 第2分類

図11 分類による検索の検索結果一覧

また、利用者の情報入手全般を支援する機能を実験として提供している。「デジタルアーカイブ関連」をテーマにウェブログ (Blog) 等を含む他サイトが配信する情報を集めて表示し、また「ニュースブックマーク」において「ポータル」「アーカイブ」「情報セキュリティ」などのキーワードに当てはまるニュース情報などを収集してヘッドライン表示を行っている。これらの情報の収集にはRSS^{注3)}を利用するとともに、ポータルを訪れなくても個々の利用者がRSSリーダーで受け取って閲覧可能となるよう、RSSで配信している。プロトタイプからのお知らせや更新情報、新着コンテンツ、更新コンテンツなども、ポータル画面上に表示するだけでなくRSSでの配信も行っている。

2.2 統合検索機能の実装と「共通仕様」

図12はプロトタイプの実装イメージ図である。

上部の楕円は、検索対象またはその候補となる個々のデジタルアーカイブシステムを表している。

デジタルアーカイブシステムは、自身が保有するデジタル情報などを管理するためにそのメタデータを保有している。これらのメタデータを検索可能とするため、プロトタイプのサービスプロバイダでは、機械的に連携し、メタデータを事前に収集してサービスプロバイダ内のメタデータデータベースに格納しておく方法 (以下「ハーベスト」と呼ぶ) または利用者からの検索要求の都度にデジタルアーカイブシステムへ検索をかけて結果を受け取る方法 (以下「横断検索」と呼ぶ) のどちらかを採っている。ハーベストでは、メタデータデータベースに格納されたデータのインデックスをサービスプロバイダ内であらかじめ作成しておき、検索時にはこのインデックスを参照して該当するデータの取り出しを行う。横断検索では、データの取り込みやインデックス作成はポータル側では行わず、検索時にはデジタルアーカイブシステム側へ検索要求を送信し、そこで生成された結果をポータルでそのまま受け取ることになる。

図12 平成17年度プロトタイプの実装イメージ

さらにポータルでは、ハーベストと複数の横断検索の結果を統合し表示している(図13参照)。これにより、データの所在や検索方法の差異を意識することなく、一元的な結果の閲覧が可能となる。

現在、デジタルアーカイブシステムが保有するメタデータ要素は、おのおののシステムに依存した定義により作成され、プロトコルも統一されていない状態である。ポータルでそのようなメタデータを統合的に扱うためには、個別に対応して他システムのメタデータ要素との差異を調整しなければならず、統合的に扱うことが難しいものとなっている。この問題は連携対象が増加するほど複雑になる。そこで、メタデータをハーベストもしくは横断検索するためのメタデータ要素、記述規則および通信プロトコルの「共通仕様」をあらかじめ決めておくことが必要となる。各デジタルアーカイブシステム側でその仕様に沿ったインターフェイスを実装してもらうことで、ポータルとしては連携対象をよりスムーズに拡大することができ、保有コンテンツが一層利活用されることになる。それによって個々のシステム間の連携も容易になり、また、各検索エンジンや各種ポータルがシステム内の情報を自由に検索でき、副次的に有用な連携が生まれることが期待される(このような連携を行える機能を備えたデジタルアーカイブシステムを「データプロバイダ」と呼ぶことにする)。

プロトタイプ構築に当たっては、暫定的ながら「共通仕様」の案を定めた。今後は他機関と調整を行いながら、案を確定し普及していく予定である。

2.2.1 「共通仕様」(メタデータ要素)

データプロバイダとポータルとのデータ交換に使用するメタデータ要素は、最低限Dublin Coreに

定義されたものは可能としている。Dublin Coreは業種を超えてある程度普及しているという利点があるが、検索やその他実現したい機能のために必要と思われるメタデータ要素が定義されていない場合があった。例えばタイトルおよび作成者の日本語読み、全国書誌番号、国立国会図書館分類表(NDLC)の分類記号や、サムネイル表示のための要素等である。そこでプロトタイプでは、Dublin Coreにこれらの要素や独自の限定子を追加して、「共通仕様」の案に取り入れている。

2.2.2 「共通仕様」(通信プロトコル)

プロトタイプでは、次の通信プロトコルに関しても実験を行った(図14参照)。

- ・ハーベスト：OAI-PMH^{注4)}、SOAP^{注5)}、RSS
- ・横断検索：SRW^{注6)}、Z39.50^{注7)}

「共通仕様」では上記のほか、RSSやその他Webサービスによる横断検索も想定した。

2.3 プロトタイプでの検証結果

検証の結果、複数のコンテンツを一元的に検索し、利用者を一次情報までワンストップでナビゲートすることが可能であり、有用だと確認できた。ただし、プロトタイプの機能やサービスはまだ実用としては十分でない点もあると認識しており、検証結果として問題点も多数洗い出されている。例えば、次のような点である。

- ・検索機能の面では、ハーベストと横断検索の結果のマージに問題があり、結果の並べ替えができない場合や結果のページ送りが1ページずつのみになってしまう場合がある。
- ・ハーベストしたメタデータの分かち書き(単語をインデックス作成のために最小単位に分割

図13 統合検索のイメージ図

図14 適用した通信プロトコルの概念

すること)に問題があり、入力されたキーワードから適切に検索できないケースがある。

- ・横断検索の手法では、横断検索先の検索機能に依存するので、他のデータプロバイダと結果の取り出しに相違が出てしまう可能性がある。
- ・適用技術の検証部分に特に注力したため、利用者にとって大切な、操作性やサイトの見た目等への配慮が欠けている。

以上は、検証結果のほんの一部であるが、本格システムの構築においてはこれらの成果を十分活かしていく所存である。

3. Web2.0世代のポータルを目指して

ポータルにおいて現在注目しているのは「Web2.0」と呼ばれる概念である。

Web2.0¹⁾とは、このような名称の技術があるわけではなく、既存のWebを「Web1.0」と総称し、これの対比として次世代のWebに見られる特徴的な姿勢を示したものだといえることができる。Web2.0の性質を端的に表現する言葉として、次のようなものがあると言われている²⁾。

「ユーザの手で自由に分類」「リッチなユーザ経験」「ユーザ体験の蓄積をサービスに転化」「ロングテール」「ユーザ参加型開発」「ユーザ生成型コンテンツ」「知のオープンソース」「進歩的分散志向」

Web2.0で論じられる世界を(部分的に)実現する一形態として、既存サービスでは例えば、GoogleMap, Google AdSense, AmazonのユーザーレビューやBlog, mixi, Wikipediaなどがよく言及されている。既存の技術も見直され、Ajax^{注8)}が注目されている。

Web2.0の概念は、利用者の利便性を向上させるものであり、ポータルでも積極的に取り入れていくべきであると考え。本章では、Web2.0の概念に適合する要素を取り入れる可能性への考察を交え、デジタルアーカイブポータルの今後の展望を紹介する。

3.1 メタデータの取得

ポータルでのナビゲーションは、各デジタルアーカイブシステムが持っているメタデータを収集、もしくは、ポータル側から横断検索できるようにしていただくことにより実現する。前述した「共通仕様」に従った提供機能があるデータプロバイダの場合、ポータルは円滑にメタデータを取得することが可能である。

このような提供機能がないシステムの場合、ポータルを含む他システムからアクセスするため、「共通仕様」に準拠した仕組みを提供してもらう。実装を促すために、当館が汎用的な提供機能のソフトウェアを開発して配布し、それを他システムに適用してもらうことも一法として考えている。

3.1.1 RSSの活用

RSSはWeb2.0の「ユーザ生成型コンテンツ」の促進や「進歩的分散志向」を実現する技術と言えるだろう。BlogにおけるRSSを用いた更新情報の自動作成・配信機能、また、配信されたRSSデータの「RSSリーダー」を用いた受信はよく見られる。

またRSSはデータ構造が単純でありかつ拡張性があるという特徴を持つ^{注9)}。この適用容易性ゆえに、ポータルではRSSの可能性に注目している。

ポータルのメタデータ収集場面での適用を考えると、他のサイトから配信されたRSSデータをサイト側の改修負担をかけることなく収集し、統合検索対象のコンテンツとして用いることもできるだろう。また、通信するデータフォーマットにRSS2.0を用いた検索「OpenSearch」³⁾を用いることも考えられる。これにより、一層RSS配信元のコンテンツが活用されることが期待できる。

3.2 ソーシャルブックマーク

Web2.0の考え方では、利用者がコンテンツ作成に「参加」したり、「ユーザ体験の蓄積をサービスに転化」し他の利用者が利用可能としたりすることが行われている。

ポータルでは、データプロバイダおよびポータルでのメタデータ作成機能の一つとして、利用者が興味を持ったコンテンツやサイトをブックマークする機能を想定している。ブックマーク登録時に該当コンテンツのメタデータを取り出してデータベースに蓄積するが、その際利用者はフリーキーワードを含めたタグ(カテゴリ)を付与可能としておく。利用者が推薦したサイトを検索対象とし、さらに利用者の手で付与したカテゴリを用いた検索を行うことで、利用者の知識と体験を取り込むことができる。

3.3 辞書の活用

プロトタイプでは、当館の典拠データを辞書として用いたが、他の有用な辞書データを追加してともに検索可能とすることで、より幅広い支援を実現したいと考えている。辞書として有用なデータを蓄積したサーバ機能を用意し、外部システムと辞書検索のWebサービスで連携することも想定している。例えば、WikipediaなどはWeb2.0の「ユーザ参加」により編集された、評価の高い「知のオープンソース」の一つの実現形態であり、このような辞書を取り入れることも考えられる。

辞書はそれ自体の検索サービスのみならず、ブ

ロタイプ の辞書による検索支援と同様の機能の実現やメタデータの自動付与および統制等のさまざまな場面において活用が可能である。さらに、OWL(Web Ontology Language)^{注10)}を用いて辞書データを記述するなど、セマンティックWeb技術^{注11)}を活用すれば、機械的に意味解析可能となり、さらに活用の幅が広がると考えられる。

3.4 メタデータの自動生成、組織化、保存

ポータルで扱うメタデータは、先に示した「共通仕様」におけるメタデータ記述要素を用いることを基本としている。しかし、一般の組織では、図書館のように統制された書誌を作成することは難しいだろう。そもそも、多くのデジタルコンテンツに手作業できちんとしたメタデータを付与することは困難である。

メタデータ収集時や検索インデクス作成時には、辞書や日本語構文解析技術(日本語の文節の構造を文法に基づいて解析する技術)等を活用して、他のメタデータ要素の値や一次情報のデータを解析し、メタデータ要素を推測して可能な限り機械的に付与することが考えられる。ただし、期待される精度が得られない場合のため、人手による一部要素の入力や、自動作成した結果の修正等(メタデータの半自動生成)を取ることを考慮する必要がある。

3.5 メタデータ提供とインターフェイス

ポータルがサービスを提供する際、そのインターフェイスは大きく二つに分けられる。利用者に対するGUI(Graphical User Interface)、システム間連携のためのAPI(Application Program Interface)である。

GUIについては、従来のLook&Feelに優れたレイアウト・デザインの考慮はもちろんだが、Web2.0に基づく「リッチなユーザ経験」を実現するために、特段の説明がなくとも利用者が直感的な操作を行えるGUIを備えていることが望まれる。Ajaxを用いることで、マウスポインタを当てた部分のヘルプを半透過表示したり、統合検索の結果表示において、メタデータの一部を隠しておいて必要な時に展開して表示したりすることが可能となる。また、利用者自身により、表示するブロックを選択可能としたり、表示ブロックの配置をマウスのドラッグ操作により動かすなど、利用者ごとに使い勝手の良いポータル画面を容易にカスタマイズできるパーソナライズ機能が実現可能だと考える。

統合検索機能においては、収集したメタデータに対して辞書を活用することで、入力キーワードの推定表示等が可能と思われる。また、パーソナライズ機能として、利用者の志向に応じて検索対象を自動的に絞り込む機能を備えることも考えられる。

3.6 Webサービスによる連携

Web2.0の「進歩的分散志向」に従うと、今後はそれぞれのデジタルアーカイブシステムが、Web上で他システムからサービスとして利用可能な形での連携に対応が求められる。このようなWebサービスによる連携機能が提供されると、コンテンツの検索サービス、閲覧サービス、貸出・複写申込みサービス、Amazonのように注文サービス等を自由に利用したサービスを提供することが可能になる。

ポータルでは、利用者がGUIを通じてメタデータ検索を行う機能だけでなく、統合検索などの各提供サービスについて、WebサービスのAPIとして提供することが好ましいと考えている。国等の公共機関に関しても、同様のサービスの提供を期待したい。例えば、公共図書館が「情報ハブ」として各地域のデジタルコンテンツ等を集約する枠組みが報告されているが⁴⁾、この集約した情報を提供するWebサービスを提供する、すなわち蔵書目録検索サービス、複写申し込みサービス等を独立したWebサービス機能として構築しAPIを公開すれば、その情報が広く利活用されることになる。

各種ポータルサイト間でもシステムの連携して相互補完しあい、全体で相乗効果を発揮できるようになることが望ましいと考える。

3.7 Blog, Wiki

ポータルの説明文書や開発運用等の最新情報は、BlogやWebブラウザから容易にWebページを編集できるWikiといったソフトウェアを用いて編集し、利用者からもコメント付与や編集を可能とすることが考えられる。Web2.0の「ユーザ生成コンテンツ」や利用者が情報を作成して発信していくCGM (Consumer Generated Media) を実現する一つの方法である。Wikiはプロトタイプでも導入したが、Wikipediaのように、知識豊富な(不特定な)利用者の力を借りることで想像以上に優れた文書が完成することもあると考えられる。

3.8 システム基盤

ポータルでは、多くの利用者およびシステムからアクセスがあると想定される。また、ポータルでの各種サービスに必要なデータは増加していくと予想できる。そこで、これらを考慮したシステム基盤の構築が不可欠である。

現在主流のアーキテクチャとして、システムをサービス単位で構築し、サービス間は疎結合な関係としてネットワーク上で連携させるSOA (Service Oriented Architecture) が提案されている。SOAはサービスの改修の影響が他のサービスに及びにくい。また、ネットワーク上で複数のコンピュータを連携させ、仮想的に一つの大規模なシステムとして用いることが可能なグリッドコンピューティング技術を適用し、スケーラビリティ (拡張性) に富むシステムを構築することが可能である。

ポータルでもこのような技術を適用し、改修の負担の最小化、システム管理の省力化を図る想定である。

4. 終わりに

インターネット上にある膨大な情報は、「意味ある情報資源」として活用するための技術開発を進めるとともに、情報の提供者はその技術を適用し提供することにより、巨大な知識ベースとなっていく。デジタル化されていく情報を日本の文化遺産として後世に残し、新たな知識を創出するための知識として現在および将来にわたって活用できるようにするため、データ提供者、サービス提供者おのこの機関が互いにインセンティブをもち、相互補完しあいながら発展していけることが重要である。当館は国のデジタルアーカイブポータルを構築し提供することにより、各機関のデジタルアーカイブが、他システムからの利用を意識したサービス構築のための仕組みを実装することの普及促進の役割を果たしたいと考えている。

利用者の日常の情報探索行動を支援するサービスとして、利用者のニーズに応えられる国のポータルを構築するためには、関係機関が持つコンテンツ、Web上で提供しているサービス、および今後普及が見込まれる次世代Webに関連した有用な技術を活用させてもらうことが重要であり、関係機関と協力して実現していきたい。

参考文献

- 1) O'Reilly, Tim. "What Is Web 2.0". (online), available from <<http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>>, (accessed 2006-07-07).
- 2) 近江商人JINBLOG. "Web2.0とはー7つの分類と要素MAP". (オンライン), 入手先<http://ceonews.jp/archives/2005/10/web20_7map.html>, (参照2006-07-07).
- 3) "OpenSearch". (online), available from <<http://opensearch.a9.com/>>, (accessed 2006-07-07).
- 4) 図書館をハブとしたネットワークの在り方に関する研究会. "地域の情報ハブとしての図書館ー課題解決型の図書館を目指してー". (オンライン), 入手先<http://www.mext.go.jp/a_menu/shougai/tosho/houkoku/05091401.htm>, (参照2006-07-07).

本文の注

- 注1) プロトタイプのURL: <http://www.dap.ndl.go.jp/home/>
- 注2) 汎用連想計算エンジン (GETA) 公開HP: <http://geta.ex.nii.ac.jp/>
連想検索の仕組みについては次のページを参照されたい。 http://webcatplus.nii.ac.jp/about_plus/top.html
- 注3) RDF Site Summary または Really Simple Syndicationの略。Webの見出しや要約を構造化して記述するフォーマットで、ニュースサイトやBlog等の更新情報の配信の用途によく用いられている。
- 注4) OAI (Open Archive Initiative) が定めた、サーバ間でメタデータを機械的に収集するためのプロトコル。
- 注5) WWW 関連技術の標準化を進める W3C により策定された、XMLベースで他のシステムとの情報の交換を行うためのプロトコル。SOAPを用いることにより、プログラムは異なるサーバの機能を自動的に呼び出し、活用することが可能となる。
- 注6) 米国議会図書館のZ39.50 International: Next Generation (ZING) によって策定された、SOAPを利用してデータの検索および取得を行うプロトコル。Z39.50の次世代として開発された。
- 注7) 米国議会図書館によって策定され、従来から複数の目録情報の横断検索等に用いられてきたプロトコル。
- 注8) Asynchronous JavaScript + XMLの略。サーバと非同期に通信を行うことで、操作の結果をリロードすることなく随時に取得可能とする。これにより、サーバを意識しないスムーズな操作性が実現する。
- 注9) RSS の活用可能性を論じたものとして、次の文献がある。
林賢紀; 宮坂和孝. RSS(RDF Site Summary)を活用した新たな図書館サービスの展開. 情報管理. Vol.49, No.1, 2006, p.11-23.
- 注10) Web上の各種サーバ等で独自の定義で用いられているメタデータスキーマ等の語彙について、語彙同士を意味内容から関連付けて定義するための言語。OWLにより相互の語彙の関係を示すことで、異なったシステム間の相互のデータ連携を促すことができる。
- 注11) Webサイト等の情報資源にメタデータを付与し、その意味をコンピュータが解釈して自動的に必要な処理を行うことを可能とする概念。