

1. 作成した(取り組んだ)ポータル名 若しくは 課題

ポータル名: 学術論文ポータル

取り組んだ課題:

- (1) Dspaceへのデータ一括搭載およびOAI-PMHの動作テスト
- (2) GETAの実装(インストールおよびデータの格納)
- (3) GETAによる検索結果からDspaceへのリンク

2. 1の概要

今回の演習では「学術論文ポータル」を大まかなテーマとして掲げた。その背景には、現在、業務で学位論文データベースを構築しているため、これを基にしたポータルサイトを設けられないか、ということがあった。ただ演習中にこれを完成させることは不可能であるため、今回はその一部分として、収集したデータをOAI-PMHで広く外部に公開するためのシステム構築を中心課題に据えることとした。構築するシステムのベースにはDspaceを用いることにしたが、これまでDspaceについては文献で知るだけであったので、その機能や管理方法を学ぶことも演習の効果として期待していた。

また今回のセミナーで取り上げられた汎用連想計算エンジン(GETA)にも興味を持ったため、その実装にも挑戦してみることにした。これを通じて従来型のキーワード一致とは異なる発想の検索エンジンの運用ノウハウを学ぶとともに、学術論文ポータルでGETAを用いる場合の評価を行うこととした。

課題の(3)に掲げたGETAとDspaceの連携は、GETAによる検索結果から詳細情報へのリンク実装の一環として取り組んだ。Dspace自体に検索機能があるため、本来このような使い方をする必要はないのだが、結果的にはDspaceに格納された個別データへの直接的なアクセス手法を学ぶことができ、そういう意味でもよい勉強になった。

3. 演習とその成果 何を計画し、実装して、何ができたか。 ※成果物のスクリーンショットも

(1) 演習第1日

Windows2000 professional+Virtual PC 上に Redhat Linux9 (Linux) をあらかじめ構築して演習に望んだが、当日になってLinuxからのネットワーク利用ができず、その環境設定変更にかかなりの時間を費やした。結果的にはWindows2000上に搭載したファイヤーウォールソフトによってLinux側の通信が遮られていることが判明し、ファイヤーウォールの設定変更によりネットワーク利用が行えるようになった。

その後は外部サイトへのOAI-PMHによるハーベスティングのテストや、OAI-PMHサイトの調査などを通じて今回のセミナーにおける演習計画を検討し、翌日に行われる着手発表の準備を行った。

(2) 演習第2日

着手発表後、午前中に行われたGETAの説明に触発され、着手発表の内容を修正し、まず最初にGETAの実装に取り組んでみることにした。使用したGETAのバージョンは第2版(単独PC用)である。インストールはスムーズに進めることができたが、そこに格納するデータのインデックス作成に苦しみ、これに1日を費やした。

インデックス作成を困難にした大きな問題の一つは文字コードの問題であった。インデックス作成に使うデータはすべてEUCに統一されている必要があったが、このデータ作成をWindows側で行っていたため、文字コードはShift-JIS、改行コードはCR+LF (UNIXはLFのみ)となっていた。文字コードの問題は知っていたので意識しながら作業していたつもりだったが、それでもうまくいかないことが多々あり、その都度変換作業に頭を悩ませた。

もう一つの大きな問題は、インデックス作成の際には、対象データから freqfile というファイルを作成する必要があったが、それを自力で行わなければならないことであった (freqfileは記述形式が厳格に決まっているが、これを自動的に作成してくれるツールは用意されていない)。データからfreqfileを作成するにはシェルのコマンドを活用したバッチファイルを作成しなければならなかったが、私自身がその知識に乏しく、まわり道をしながら freqfile の作成をすることになった。それだけでかなりの時間を浪費することになった。

(3) 演習第3日

昨日苦しんだインデックス作成もようやくメドが立ち、無事インデックス作成が終了したため、次の段階として検索結果一覧にタイトル表示を出すためのタイトルファイルの作成に着手した。丸山講師によれば、その作成方法はマニュアルに書かれていないとのことで、講師から詳しい手順を丁寧に教えてもらった。これにより午前中にはGETAの実装が完了した(図①)。このような機会であればGETAを実装することは簡単ではなかったと思われる、今回の演習で実装に挑戦したことは、大変有益な経験となった。なお格納したデータには、学位論文データのサンプルを使用した。

午後からはDspaceに着手した。Dspace(ver.1.1)は今回の演習に備えてインストール済みであったため、まずデータの一括登録(インポート)に取り組んだ。インポートの方法についてはDspaceシステム説明書の「アイテム・インポータ/エクスポータ」に記述されているが、鈴木講師から詳しい手順について教えていただきながらインポートを行った。Dspaceに搭載するデータはもともとXMLで用意していたが、Dspaceの構造にあうよう、XSLTなどを用いて適宜加工した。インポート終了後、このDspaceに対しOAI-PMHによるハーベスティングを行い、問題なく動作することを確認した。(図②)

最後に、GETAの検索結果からDspaceへのリンク作業を行った。このためには、Dspaceで個別のデータを検索するためのURLの形式の調査と、そのURLをGETAのCGI(assoc.cgi)中に反映させることの2つの作業が必要であった。本来、GETAで用いるキー(ID)とDspaceのそれとが同じものであれば難しい問題ではなかったのだが、これを異なる体系にしてしまったため、GETAのIDを元にしてDspaceのデータを検索する手法の調査にかなり時間を費やした。Dspaceのマニュアルも参照したが、最後はDspaceの検索の際に組み立てられるURLを何度も試しながら調査することで、目的となるURL形式を得ることができた。(図③)

4. 研修で学んだ技術及び知識 ※参考になったURLとその簡単な内容紹介(1行程度で)

- ・ 汎用連想計算エンジン(GETA)
URL <http://geta.ex.nii.ac.jp/>
GETAに関するWebサイト。GETAはオープンソースであるため、ここからダウンロードも可能。
ただしドキュメント類はわかりにくく、それだけで実装を進めるのは難しいと思われる。
- ・ DSpace システム説明書(ver.1.1.1) アイテム・インポータ/エクスポータ
URL <http://www.nii.ac.jp/metadata/irp/dspace-docs-jp/application.html#itemimporter>
NIIが公開しているDspaceシステム説明書の日本語訳。
- ・ Registered Data Providers
<http://www.openarchives.org/Register/BrowseSites>

Open Archives Initiativeが確認済みの、OAI-PMHに対応したデータプロバイダのリスト。OAI-PMHを用いたハーベスティングのテストを行う場合、そのテスト先を選ぶのに便利。

5. 事前準備として必要と思われるもの（不足していたソフトウェアや予習事項等）

UNIXのシェルコマンド、および PerlやPHPなどプログラミング言語に関する知識

→ GETAのインデックス作成や、Dspaceへの一括登録など、データを操作することが多かった。自分でも操作のためのアルゴリズムは浮かんでも、それを実際に行うためのプログラミングの知識が圧倒的に不足していた。

6. 今後の課題（職場で更に調査する必要のあるもの等）

- ・ Dspaceの管理機能の調査
- ・ Dspaceの日本語対応（メニューの日本語化、検索機能の日本語対応）

7. 今後の計画（実際のポータル構築計画等）

- ・ まずはDspaceをベースとした、機関リポジトリのプロトタイプを構築。
- ・ 学外（国内外）の学位論文データベースのうち、OAI-PMH対応のリポジトリを調査。
- ・ また学外の学位論文データベースの情報をGoogle APIなどを活用して収集し、リンク集を構築。

※公開できるものになるかは未定

8. 演習の感想

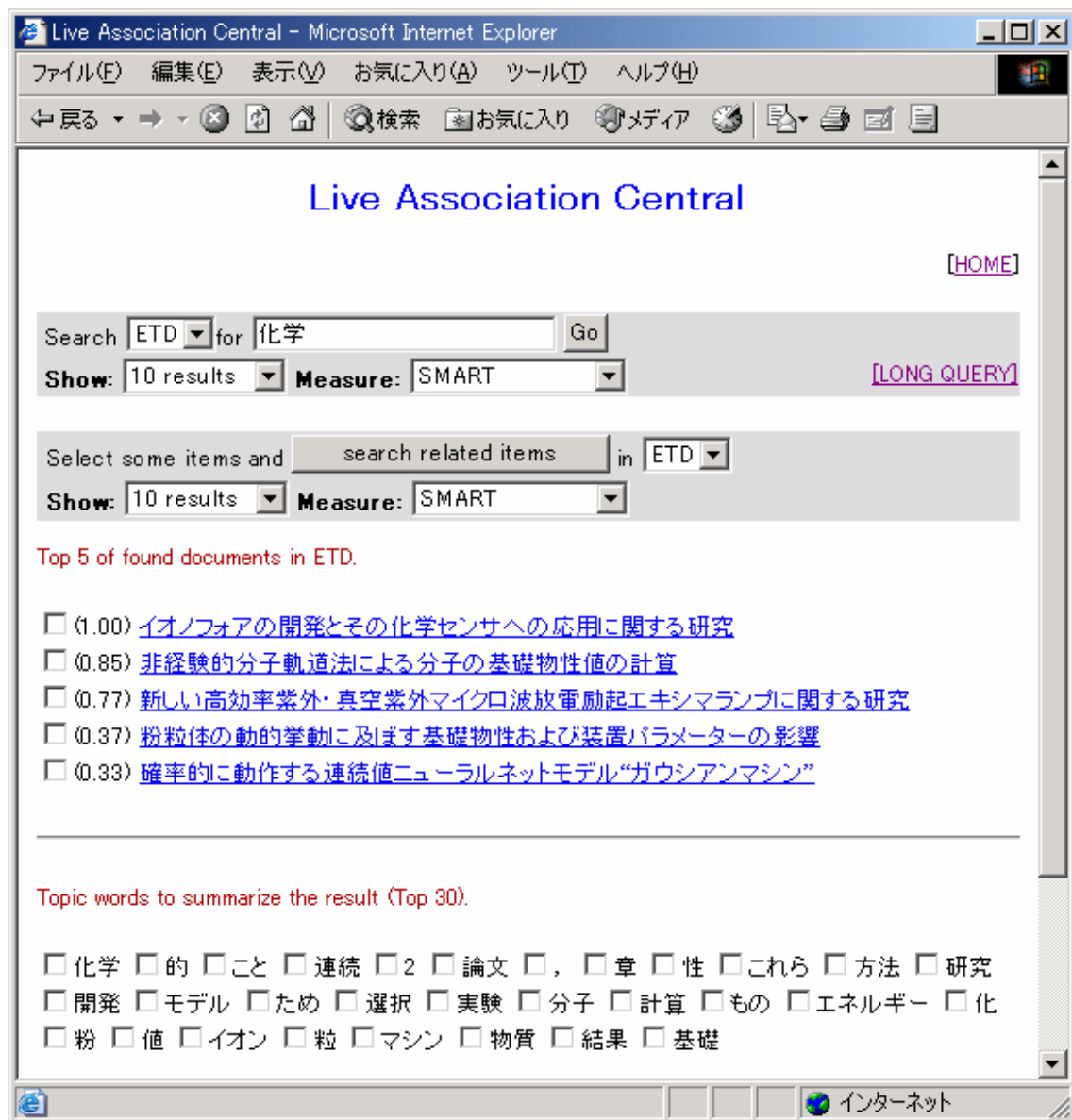
演習は、基本的に全て自力でこなしていくスタイルであったため、プログラミングに関する知識はあったつもりだったが、何をするにも一からプログラムを組まなければならなかったのは予想外で、そのためかなりの時間を浪費してしまうことになった。また、テーマは各自で設定したため、演習開始後は参加者間の共通の話題が得にくく、意見交換や情報交換をする機会がほとんどなかったのが残念であった。大まかにでも、演習で用いる技術の範囲が決まっていれば、意見交換などももう少し活発に行えたのではないかと思う。

新しい技術や知識を習得するためには、ある程度まとまった時間が必要だが、今回の演習では3日間という長い時間を得られたため、失敗を恐れずに新しい技術に存分に取り組む事ができた。また講師にもめぐまれ、壁に行き当たったときに、解決策をすぐその場で指導してもらえたことは、大変ありがたいことであった。

9. 備考、その他

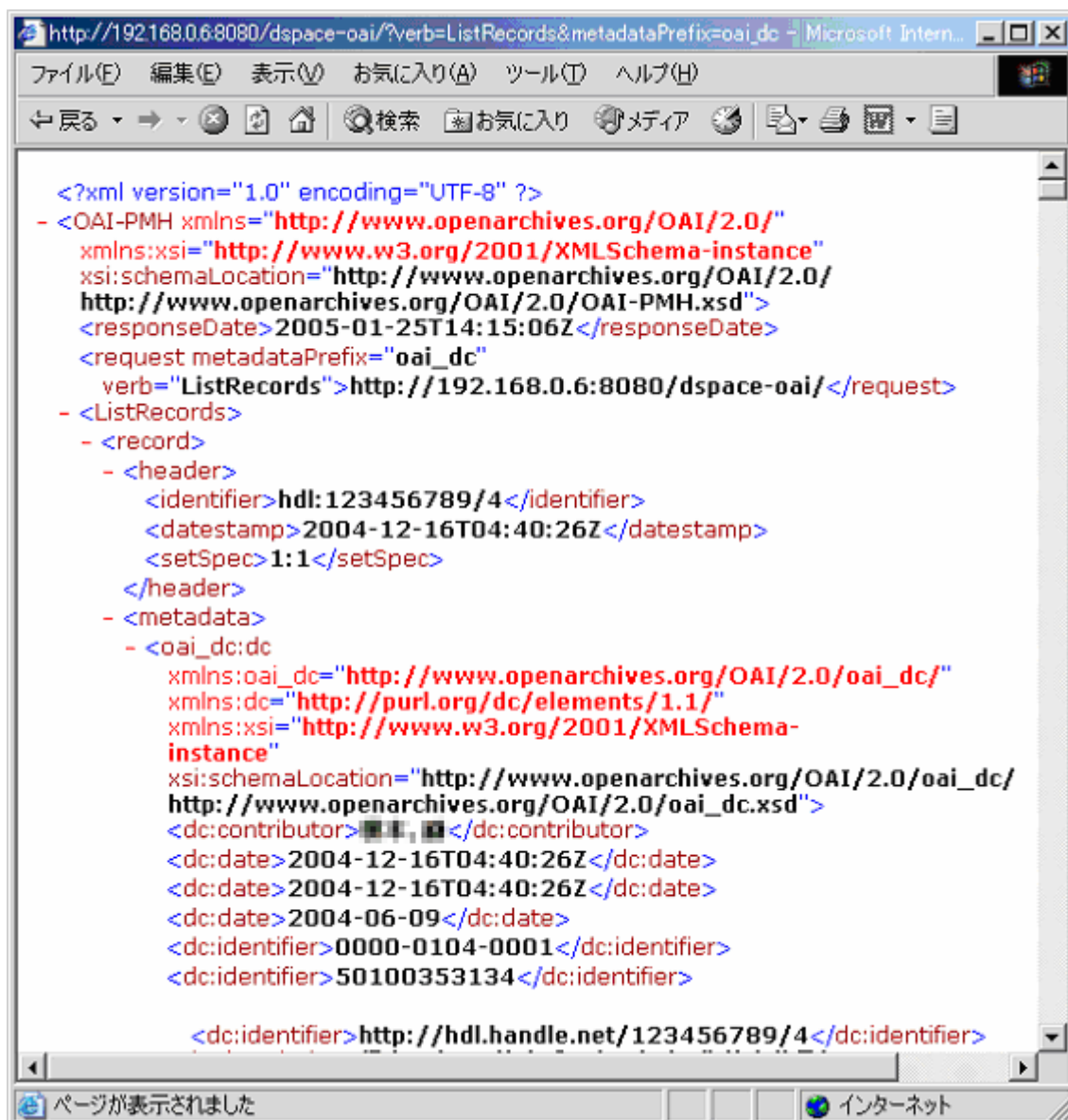
このような機会を与えてくださったNIIと、一週間に渡るセミナー期間中、参加者を陰からサポートしてくださいました羽田課長補佐と成沢さん、そして杉田さんはじめ講師の皆様にご礼申し上げます。

図① GETAが動作したところ



図② 自作のDspaceに対してOAI-PMHコマンドを実行したところ

[http://192.168.0.6:8080/dspace-oai/?verb=ListRecords&metadataPrefix=oai_dc]



図③ GETAの検索結果から、直接DSpaceのデータを検索したところ

[[http://192.168.0.6:8080/dspace/simple-search?query=\(\(identifier:"%200000-0189-0047%20"\)\)](http://192.168.0.6:8080/dspace/simple-search?query=((identifier:)]

DSpace at My University: Search Results - Microsoft Internet Explorer

ファイル(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)

← 戻る → → 検索 お気に入り メディア

DSpace™ About DSpace MITLibraries

Search DSpace: [DSpace at My University >](#)

Go
Advanced Search

Home

Browse

- Communities & Collections
- Titles
- Authors
- By Date

Sign on to:

- Receive email updates
- My DSpace authorized users

Search: All of DSpace
for Go

Results 1-1 of 1.

Item hits:

Date of Issue	Title	Authors
	イオノフォアの開発とその化学センサへの応用に関する研究	藤岡 博司

1

インターネット