

平成 16 年度情報処理軽井沢セミナー レポート

熊本大学 学術研究協力部 図書館サービス課 電子サービス係
森下 和博

1. 作成した（取り組んだ）ポータル名 若しくは 課題

『熊大ポータル金之助（架空）』の一部を作成する。

2. 1の概要

『熊大ポータル金之助（架空）』は「履修登録・確認・業績登録」「履修登録内容や研究内容からの Webcat Plus 連想検索」「熊本大学固有データの横断検索」「手続きオンライン」「マイリンク集」「履歴付き蔵書検索」など様々な機能を学内者に提供するシステムである。今回の研修では、『熊大ポータル金之助（架空）』で提供する熊大固有の情報（古文書情報・水俣病研究・ラフカディオハーン研究・電子ジャーナルなど）を一般公開するため、データプロバイダの構築を目標とする。また、他のリポジトリから OAI-PMH でハーベスティングを行い『熊大ポータル金之助（架空）』にマージすることも検証してみたい。

3. 演習とその成果 何を計画し、実装して、何ができたか。

(1) 演習第 1 日

演習テーマを決定し、情報収集を行った。DOAJ (<http://www.doaj.org>) からフリー電子ジャーナルのメタデータのハーベスティングを、wget を用いて行った。

(2) 演習第 2 日

データプロバイダとして、DSpace の設定を行い、構築した。

(3) 演習第 3 日

1. DSpace に「電子ジャーナル」「阿蘇家文書」等のコレクションを作成した。電子ジャーナルコレクションには熊本大学で利用できる電子ジャーナルのメタデータを、阿蘇家文書コレクションには画像情報を含む文書の情報を登録し、登録データが表示できることを確認した。（別紙図 1 - 3）
2. 構築した自 DSpace からメタデータのハーベスティングができるか確認を行った。（別紙図 4）
3. DOAJ のフリー電子ジャーナルメタデータを Perl で無理やり csv 形式に変換し、熊本大学電子ジャーナルデータベースへ登録するしくみの一部を作成した。（別紙図 5 - 6）

※手間取った点や補足点は「6. 今後の課題」や「別紙」にまとめた。

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）

<http://www.nii.ac.jp/metadata/irp/dande.html>

学術機関リポジトリ構築ソフトウェア実装実験プロジェクトのページ。DSpace のインストールマニュアルやシステム設定等もここからジャンプして調べることができる。

<http://www12.ocn.ne.jp/~zuki/Japanization/index.html>

DSpace の日本語化についてまとめてある。上記サイトからもたどり着ける。

5. 事前準備として必要と思われるもの

パソコンの設定で実習時間が減るのは勿体無いので、あらかじめパソコンの準備をしっかりとしておく必要があると思う。プログラミングやコンパイル方法もある程度経験しておく必要があると思う。

6. 今後の課題（職場で更に調査する必要があるもの等）

- ・ DSpace に登録したデータの日本語での検索。実習では間に合わなかった（別紙図7）。
- ・ DSpace のユーザーの追加（別紙図8）。サーバーのメール設定が正常に終了していないとダメらしい。ただ、設定せずに Admin グループを使いまわしても良かもしれない。
- ・ DSpace への一括登録（インポート）。阿蘇家文書のサンプルデータ数件の一括登録を試みたができなかった（別紙図9）。登録済みデータをエクスポートした後、そのデータをインポートしてフォーマットや方法を確認する。
- ・ Perl のモジュール「OAI-Harvester」の動作検証。DOAJ から listRecords で収集した XML データの変換について、今回は時間的都合によりごり押しで変換した。Perl のモジュール「OAI-Harvester」や「XML-CSV」を用いたスマートな変換法も確認してみる。（XML-CSV は csv を XML に変換できる。）
- ・ GETA。GETA は現在、暇を見つけて構築している。

7. 今後の計画（実際のポータル構築計画等）

- ・ DOAJ からの取得リストを電子ジャーナルデータベースへ登録してみる。（未定）
- ・ GETA に熊大固有のデータを投入して様子を見てみる。（未定）

8. 演習の感想

- ・ 思った以上にコンパイルに時間がかかるので、持参するパソコンはできるだけ高速なものが良いと思う。Perl のあるモジュールを使おうとしたら、これこれというモジュールが無い。そのモジュールをいれようとするときにさらにこれこれが…という感じで、芋づる式に増え、環境設定にかなり時間がかかった。ネットワークにちゃんと接続できたので、自大学に高速なサーバーがあればそちらに構築しても良いとも思った。
- ・ 最初は、データプロバイダにどうやってなるの！？OAI-PMH で要求が来たら応答するプログラムは自分で作る！？といった状態だった。講師や他の受講者の補助もあり、手探りだったが色々な知識を習得できて良かったと思う反面、もっと予習をしていけばよかったとも思う。

9. 備考、その他（場合により省略してください）

- ・ 課題演習中に、杉田係長のパソコンがふしゅーという音と共に落ちていたのが良いアクセントだった。（不謹慎ですみません）
- ・ Google はやっぱり偉大だった。

別紙

手間取った点および補足点

DSpace の設定

・DSpace のインストールマニュアルを Google で探すと nii のページに pdf で見つかった。ただし pdf の 2 ページ目以降を見ようとすると何故か Mozilla が落ちるため、Google の HTML バージョン機能で確認、以下このマニュアルに従いインストールする。ほぼマニュアルどおりに行ったため、エラーが出た点のみ記しておく。

1. PostgreSQL を再インストールすることになった。

あらかじめ PostgreSQL をインストールしていたが、コンパイルに特別なパラメータ指定 (ユニコード & java) が必要とのことで再コンパイルを行う、が失敗する。「Apache-ant」「Java2SDK」が無いとダメらしい。(両方ともどんな仕事をしているかは知らないし、調べなかった)「Apache-ant」と「Java2SDK」を Google で探してきてインストールし、PostgreSQL を再コンパイル・インストールした。

2. servret. jar のありかがわからない。(マニュアルからのリンク切れ)

DSpace のインストール。解凍・展開して、activation. jar serveret. jar mail. jar をそれぞれダウンロードしてコピーしなさいとのことだったが servret. jar のダウンロード元がわからず、PC 内にあった servret. jar をコピーしてインストールを開始。インストールは正常に終了したようだった。実際、DSpace へのアクセスは正常にできたが、OAI での取得命令を正常に処理できず、データプロバイダになれなかった。そのため、serveret. jar を探してきて DSpace を再構築したら正常にアクセスできた。

・DSpace をインストールする際に必要になる各ソフトウェア等のバージョンはマニュアルにあるものと同様バージョンにしたが、リビジョンが少し違うものについては気にせずに最新版をインストールした。
(指定 : Postgres7. 4. 2 実際 : Postgres7. 4. 6 をインストール)

次ページ以降は画面例 (一部合成画面あり)

図 1 ~ 6 は成果で、図 7 ~ 9 は失敗例・課題である。

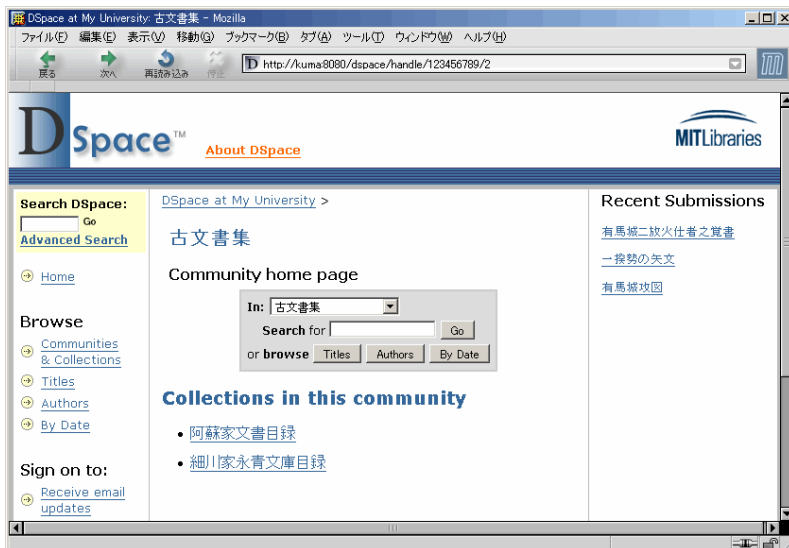


図 1
設定した DSpace にコミュニティとコレクションを作成

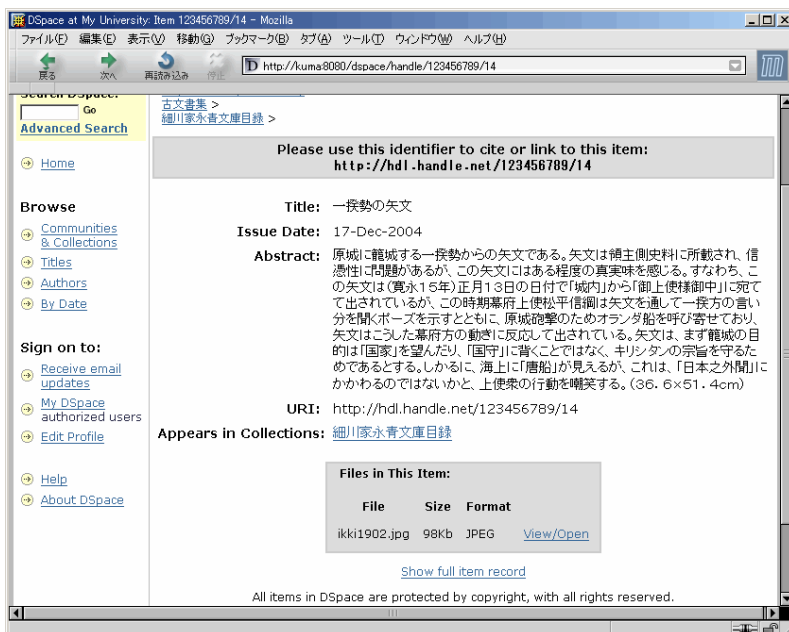


図 2
コレクションに登録した古文書の情報を表示（リストからたどる）

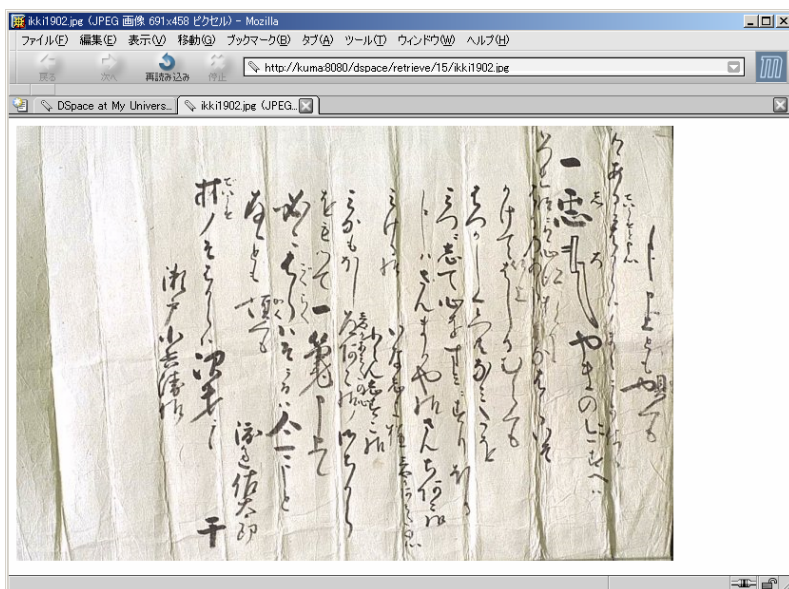


図 3
登録した画像情報を表示

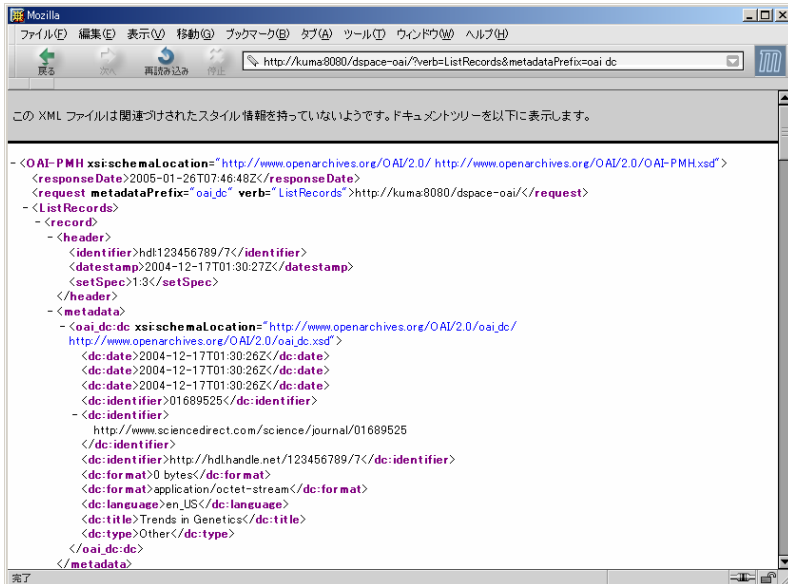


図 4

自 DSpace のデータプロバイダ機能。OAI での取得命令に応えることができるようになった。

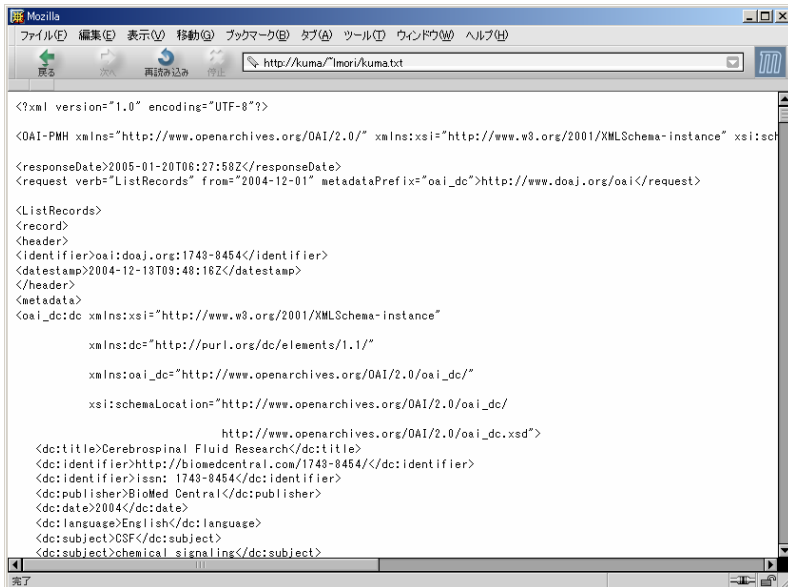


図 5

DOAJ からハーベストしたデータをテキスト出力。

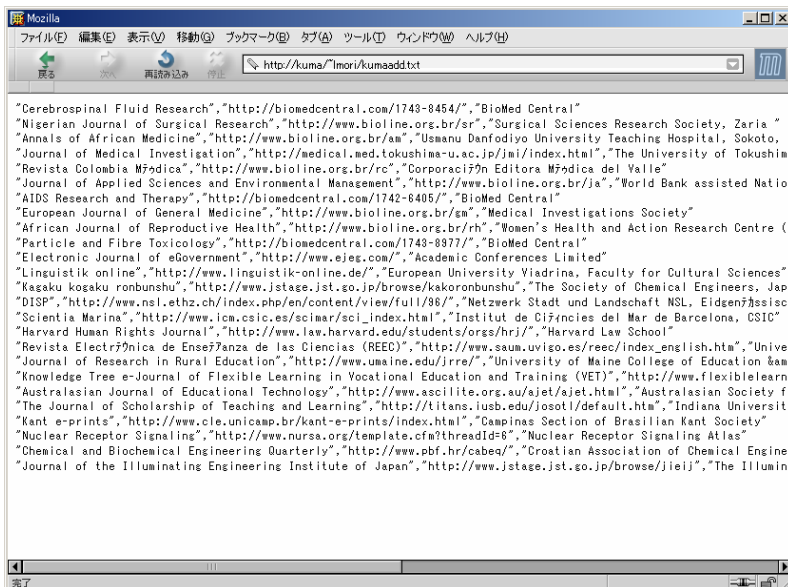


図 6

Perl で無理やり csv へ変換

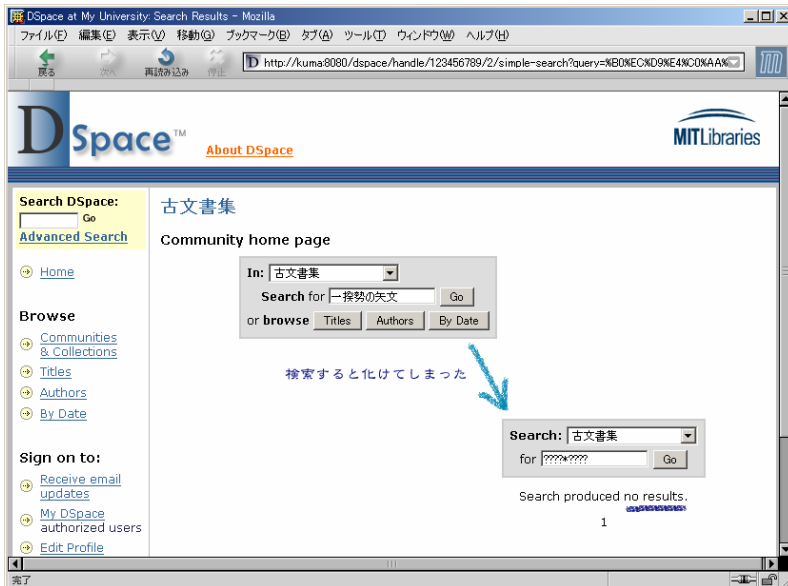


図 7

DSpace の日本語検索。解消するには 4. 参考になった URL 内の DSpace 日本語化を実行すればよいらしい

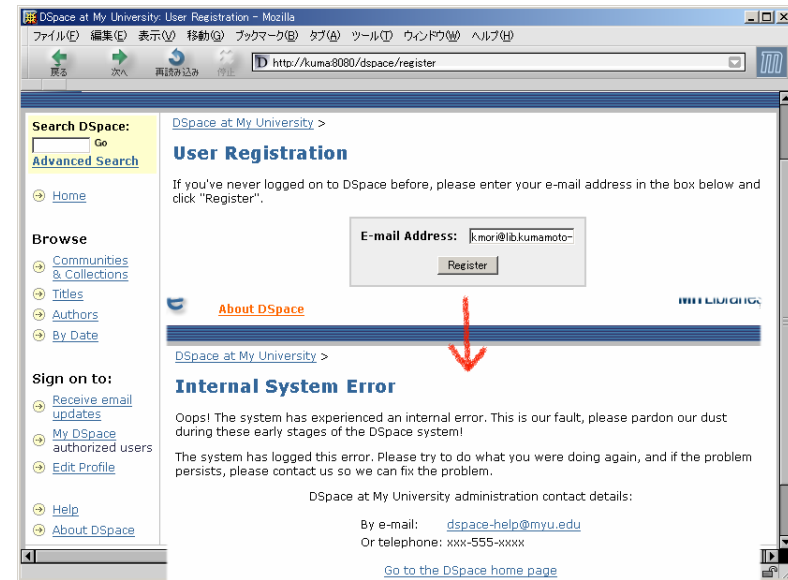


図 8

ユーザー登録の画面。(失敗画面) Internal System Error が出ている。おそらくメール (または mail.jar?) の設定が正しくできていなかったためだと思われる。

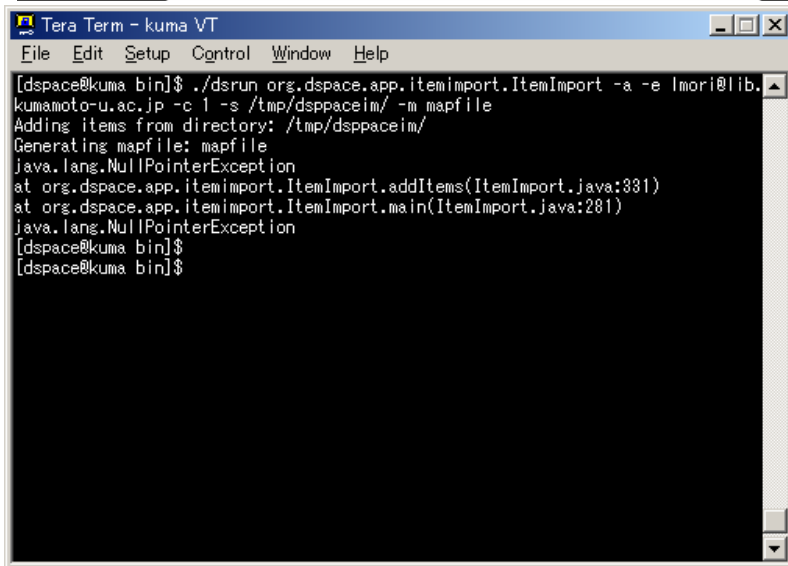


図 9

一括インポートを試してみた時の失敗例である。