

Title	機関リポジトリの設立・運営とシステム構築
Author(s)	山地, 一禎
Citation	
Version Type	VoR
URL	https://hdl.handle.net/11094/14164
rights	
Note	

Osaka University Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

Osaka University

DRF-Senri 機関リポジトリの設立・運営とシステム構築

NII学術ネットワーク研究開発センター 山地 一禎
2009-01-19

内容

1 システム概略

- ▶ リポジトリシステムとは？
- ▶ システム構成

2 リポジトリシステム

- ▶ DSpace
- ▶ EPrints
- ▶ WEKO

3 システム構築

- ▶ OS, リポジトリシステムのインストール
- ▶ バックアップ

4 メタデータ流通

- ▶ OAI-PMH
- ▶ OAI-ORE

目標

- リポジトリシステムがやっていることを理解してもらうこと
- 自分でリポジトリを立てようと思う人を増やすこと



システム概略

必要な機能

- ▶ Webインターフェースを提供すること
- ▶ メタデータとファイルが管理できること
- ▶ OAI-PMHに対応していること

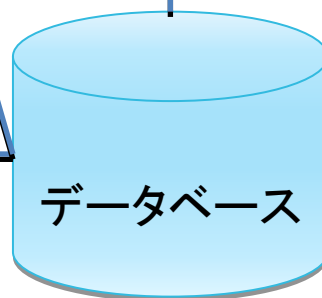
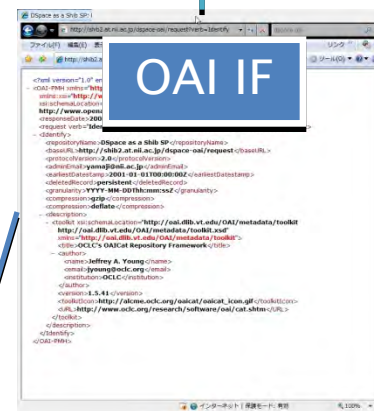


システム概略

Webインターフェース

図書館員

教員



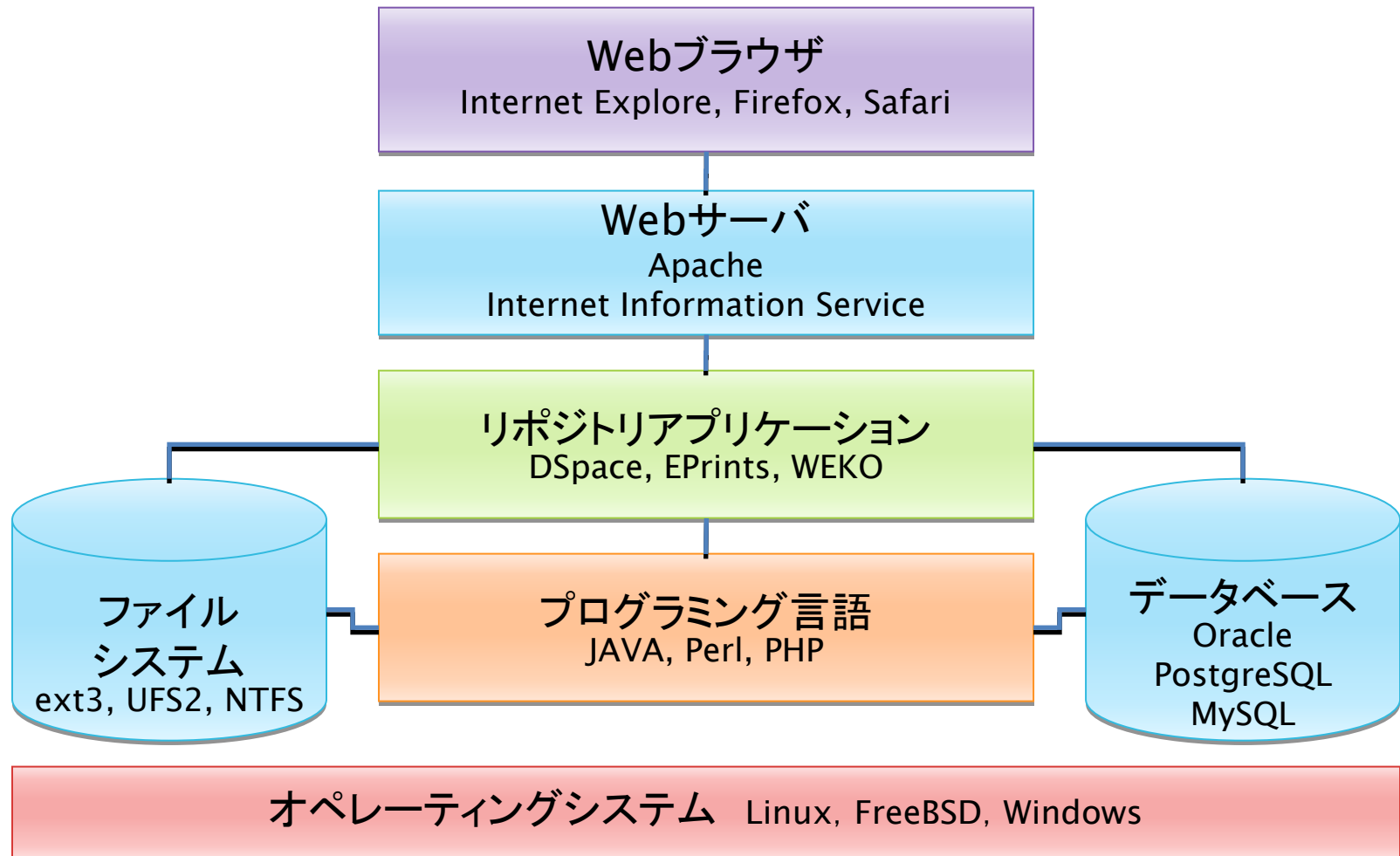
利用者

利用者



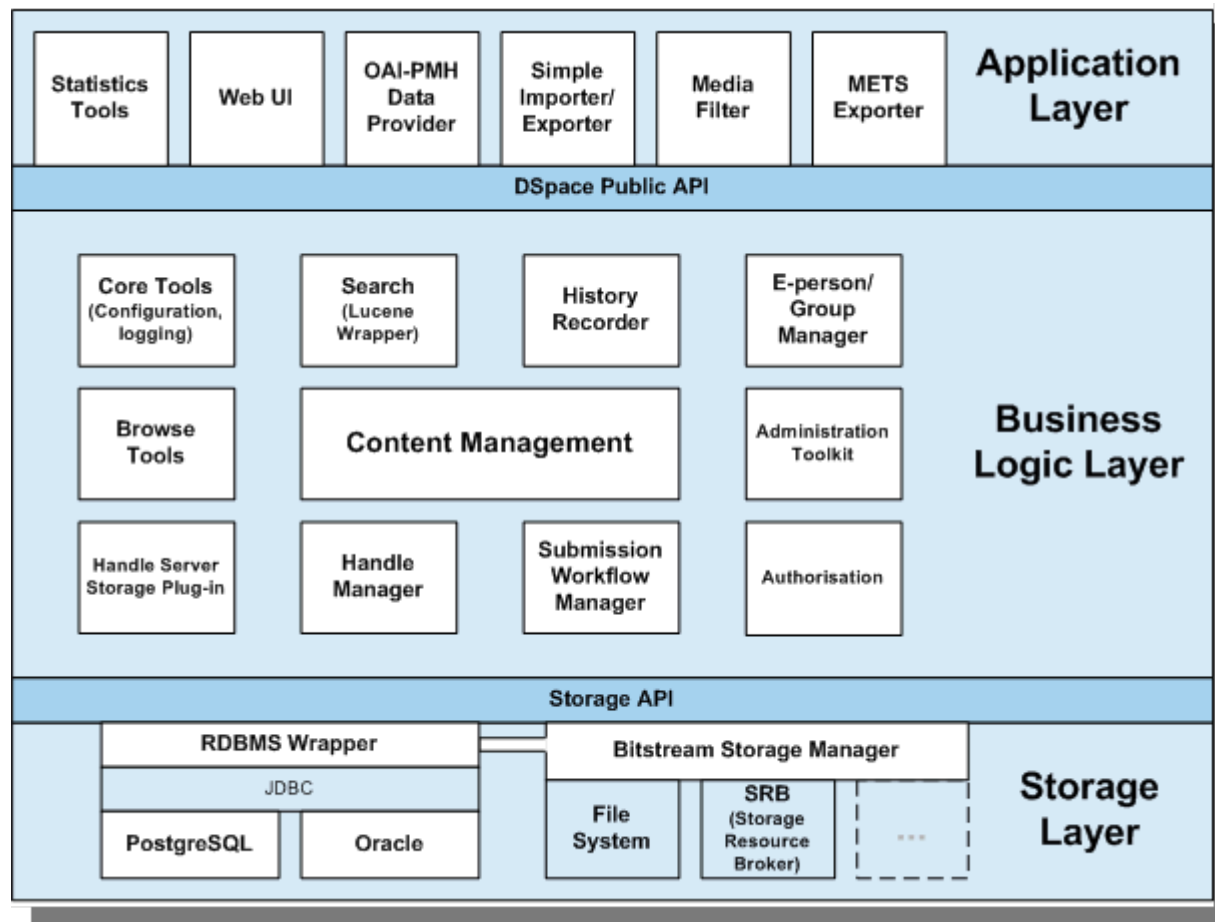
システム概略

システム構成



システム概略

DSpace Architecture



リポジトリシステム

一般的な選択肢

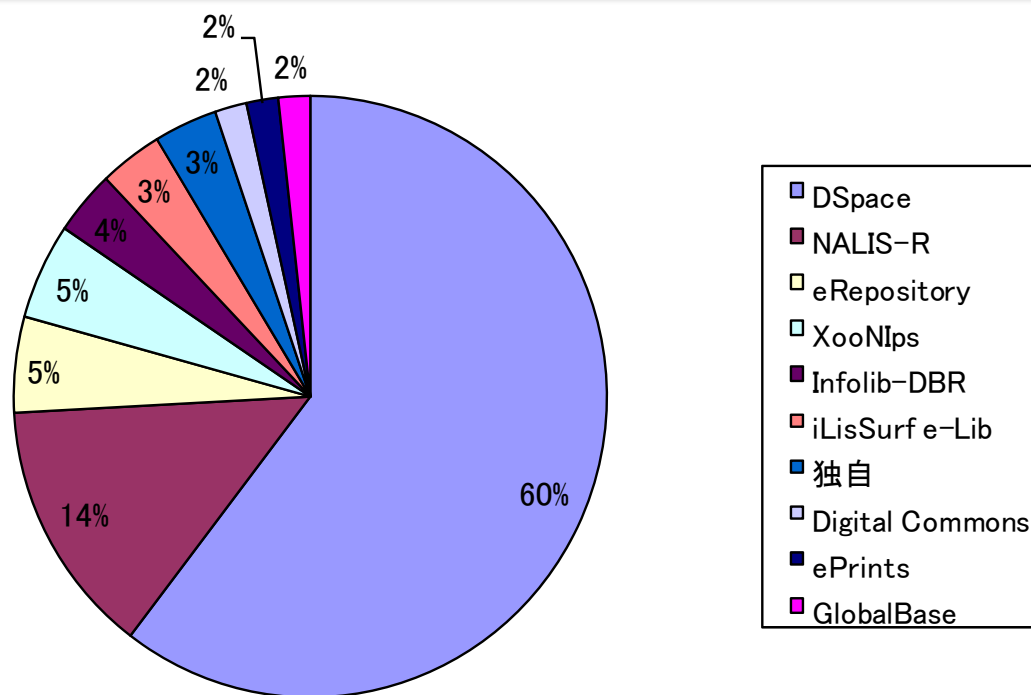
製品種類	構築方式	必要スキル	金銭コスト
オープンソース	自力	高	低
	導入支援	低～中	中
	パッケージ	低～中	高
商業製品	ローカルホスト	低～中	高
	ASP	低～中	中
独自システム	自力	高	低～高
	開発・導入委託	低～中	高

- ▶ サポート(カスタマイズ追加, 不具合対策, バージョンアップ等)も要考慮
 - ▶ 図書館だけではなく, 学内情報関連組織との連携が望まれる
-



リポジトリシステム 国内シェア

「次世代学術コンテンツ基盤共同構築事業(中間まとめ)」
(平成19年3月)



リポジトリシステム 商用製品

▶ ローカルホスト式

- ▶ Infolib-DBR (国産:InfoCom)
 - ▶ <http://www.infocom.co.jp/das/infolib/dbr/index.html>
- ▶ E-repository (国産:CMS/NEC)
 - ▶ <http://www.cmssc.co.jp/modules/office/index.php?id=12>

▶ Application Service Provider式

- ▶ Infolib-ASP (Infolib-DBRのASP版)
 - ▶ <http://www.infocom.co.jp/das/asp/index.html>
- ▶ Digital Commons (米:Berkeley Electronic Press)
 - ▶ <http://www.bepress.com/ir/>



リポジトリシステム オープンソース

▶ DSpace

- ▶ 米国・MIT, HP開発
- ▶ JAVA, PostgreSQL or Oracle
- ▶ デモ <http://dspace.at.nii.ac.jp/>



▶ EPrints

- ▶ 英国・サウサンプトン大学開発
- ▶ Perl, MySQL
- ▶ デモ <http://eprints.at.nii.ac.jp/>



▶ WEKO

- ▶ 日本・国立情報学研究所開発
- ▶ PHP, MySQL
- ▶ デモ <http://weko.at.nii.ac.jp/>



リポジトリシステム DSpace

▶ DSpace系で良くやっている機能強化

- ▶ メッセージの日本語化
- ▶ メタデータ検索の日本語対応化
- ▶ 全文検索の日本語対応化
- ▶ junii2対応
- ▶ 詳細な利用統計
- ▶ ハンドル

やろうと思えば
無料でできる

DSpaceの日本語化について

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

<http://repository.kulib.kyoto-u.ac.jp/dspace-memo/japanization.html>



リポジトリシステム DSpaceメニューの日本語化



リポジトリシステム DSpaceメニューの日本語化



リポジトリシステム DSpaceメタデータ検索の日本語化

The screenshot shows a web browser window with the URL `http://shib2.at.nii.ac.jp/dspace/simple-search?query=%E5%B1%B1%E5%9C%B0`. The page title is "DSpace as a Shibboleth SP Server: ホーム - Windows Internet Explorer". The main content area displays the search results for the query "山崎".

検索結果

検索: リポジトリ全件
検索語: 山崎

検索結果表示: 1-1 / 1.

該当アイテム:

発行日	タイトル	著者
2008	DSpaceのカスタマイズについて	山崎, 一祐

1

登録利用者:

- 検索結果一覧
- マイリポジトリ
- 登録利用者
- 利用状況一覧

ヘルプ

DSpaceについて

WS-XHTML 1.0

DSpace Software Copyright © 2002-2006 MIT and Hewlett-Packard. ご意見をお寄せください。

http://shib2.at.nii.ac.jp/dspace/

インターネット | 保護モード: 有効

100%

リポジトリシステム DSpaceメタデータ検索の日本語化

DSpace as a Shibboleth SP Server: ホーム - Windows Internet Explorer

http://www12.ocn.ne.jp/~zuki/japanization/dspace141/dspace.html

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)

DSpaceで日本語メタデータの検索を実現する

DSpaceで日本語メタデータの検索を実現する

DSpaceは検索に使用しているLuceneが日本語に対応していないので、日本語の検索はできません。ここでは、日本語メタデータの検索を実現します。

Lucene-cjkを使った方法

Nグラム(N=2)による日本語検索を実現します。方法については、http://sourceforge.net/mailarchive/message.php?msg_id=1234567

Lucene-ja + Senを使った方法

形態素解析して日本語キーワードを作成します。方法は以下のとおりです。

1. Senをインストールする

```
wget https://sen.dav.java.net/files/documents/1373/13222/sen-1.2.1.zip
unzip sen-1.2.1.zip # 展開後に作成されたsen-1.2ディレクトリをSEN_HOMEと呼ぶ
cd SEN_HOME/conf
vi sen.xml
<charset>utf-8</charset> に変更
vi sen-processor.xml
<charset>utf-8</charset> に変更
```

2. 変換辞書を作成する

```
cd SEN_HOME/dic
vi dictionary.properties
sen.charset=utf-8 に変更 # dic.charset=EUC-JPは変更しない
ant # da.sen, matrix.sen, posInfo.sen, token.senが作成される
```

3. Lucene-jaをインストールする

```
wget https://sen.dav.java.net/files/documents/1373/35812/lucene-ja-2.0.test2.zip
unzip lucene-ja-2.0.test2.zip # 展開後に作成されたlucene-jaディレクトリをLUCENE_HOMEと呼ぶ
```

4. 検索に使用する品詞を決める

```
cd LUCENE_HOME/lib
```

5. lucene-ja.jar, sen.jar, commons-logging.jarをDSpaceにコピーする

```
cp lucene-ja.jar lucene-ja.jar.org
jar cvf lucene-ja.jar org
```

6. 日本語用のDSAnalyzerを使用するための設定を行う。

```
cd DSPACE/config/
vi dspace.cfg # search.analyzerを変更する。
search.analyzer = org.apache.lucene.analysis.ja.JapaneseAnalyzer
cd DSPACE/bin
vi dsrun # 最後の行を以下のように変更する (SEN_HOME)
java -Xmx256m -classpath $FULLPATH -Dsen.home=[SEN_HOME] "$@"
cd "dspace"
vi .bashrc # 環境変数JAVA_OPTSを以下のように変更する
JAVA_OPTS="-Xmx512M -Xms64M -Dfile.encoding=UTF-8 -Dsen.home=[SEN_HOME]"
. .bashrc # 環境変数の変更を有効にする。
```

7. DSpaceを更新する

```
cd DSPACE_SOURCE
ant -Dconfig=DSPACE_HOME/config/dspace.cfg update
cp build/dspace.war build/dspace-oal.war TOMCAT_HOME/webapps
rm -r TOMCAT_HOME/webapps/dspace
rm -r TOMCAT_HOME/webapps/dspace-oal
```

8. DSpaceの索引を再作成する

```
cd DSPACE_HOME/bin
./index-all
```

9. Tomcatを立ち上げる

```
cd TOMCAT_HOME/bin
./startup.sh
```


リポジトリシステム DSpace お茶の水女子大学TeaPotサイト



DSpace as a Shibboleth SP Server ホーム - Windows Internet Explorer

http://teapot.lib.ocha.ac.jp/ocha/

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) リンク

TeaPot - お茶の水女子大学 教育・研究成果

お茶の水女子大学 教育・研究成果

TeaPot - Ochanomizu University Web Library Institutional Repository

TeaPotへようこそ！

お茶の水女子大学教育・研究成果コレクション「TeaPot」はお茶の水女子大学の教育・研究成果を世界に発信します。

2007年4月18日 正式公開しました。
2008年2月21日 登録コンテンツが1,000件を超えました！
2008年5月02日 登録コンテンツが2,000件を超えました！
2008年5月23日 登録コンテンツが3,000件を超えました！
2008年6月16日 登録コンテンツが4,000件、5,000件、6,000件、7,000件、8,000件を超えました！
2008年6月17日 登録コンテンツが9,000件、10,000件を超えました！
2008年6月19日 登録コンテンツが11,000件、12,000件、13,000件、14,000件、15,000件、16,000件を超えました！

現在 16691 件の資料が保管されています。

0 総記 General [13]
1 哲学 Philosophy [2]
2 歴史 History [5]
3 社会科学 Social sciences [76]
4 自然科学 Natural sciences [103]
5 工学技術 Technology and engineering [20]
6 産業 Industry and commerce [3]
7 芸術 Arts [3]
8 語学 Language [0]
9 文学 Literature [0]
A 特別コレクション [12206]
B 紀要等 [4261]
C 科学研究費補助金研究業績報告書 [1]

最新登録資料

熊野, 隆 (1997). 学校教育の課題と教育者養成への期待. 日本教育学会大会発表要録集, 56: 80-81. [View](#)

鈴木, 智通 ... [ほか] (2000). 緑の学校の存立構造と地域社会：新潟県緑学校の社会的分析. 日本教育学会大会発表要録集, 59: 86-87. [View](#)

Annual Report 2008

新緑教育フォーラム
新緑エッセイコンテスト

「新緑と未来」の未来を
支える「新緑」の未来

本学研究紀要等

- ・ 緑の教育
- ・ F-GNISジャーナル
- ・ お茶の水音楽論
- ・ お茶の水文学
- ・ お茶の水女子大学
子ども発達教育研
究センター紀要
- ・ お茶の水女子大学
人文科学研究
- ・ お茶の水女子大学
中国文学会報
- ・ お茶の水女子大学
自然科情報
- ・ お茶の水地理
- ・ 研究開発業績報告
書
- ・ 研究紀要 / お茶の
水女子大学附属高
等学校
- ・ 研究紀要 / お茶の
水女子大学附属中
学校
- ・ 人間文化調査
- ・ 生活工学研究
- ・ 日韓大学生国際交
流センター報告書
- ・ 比較日文学研究セ
ンター研究年報

RSS Feeds

インターネット | 保護モード | 有効

100%

リポジトリシステム

DSpace 1.5

DSpace 1.5 Overview

- Maven Build Architecture



- DSpace "modules"

- Application "Overlays"

- XMLUI (aka. Manakin)



- Configurable Submission

- Streamlined Browse System

- SWORD Interface



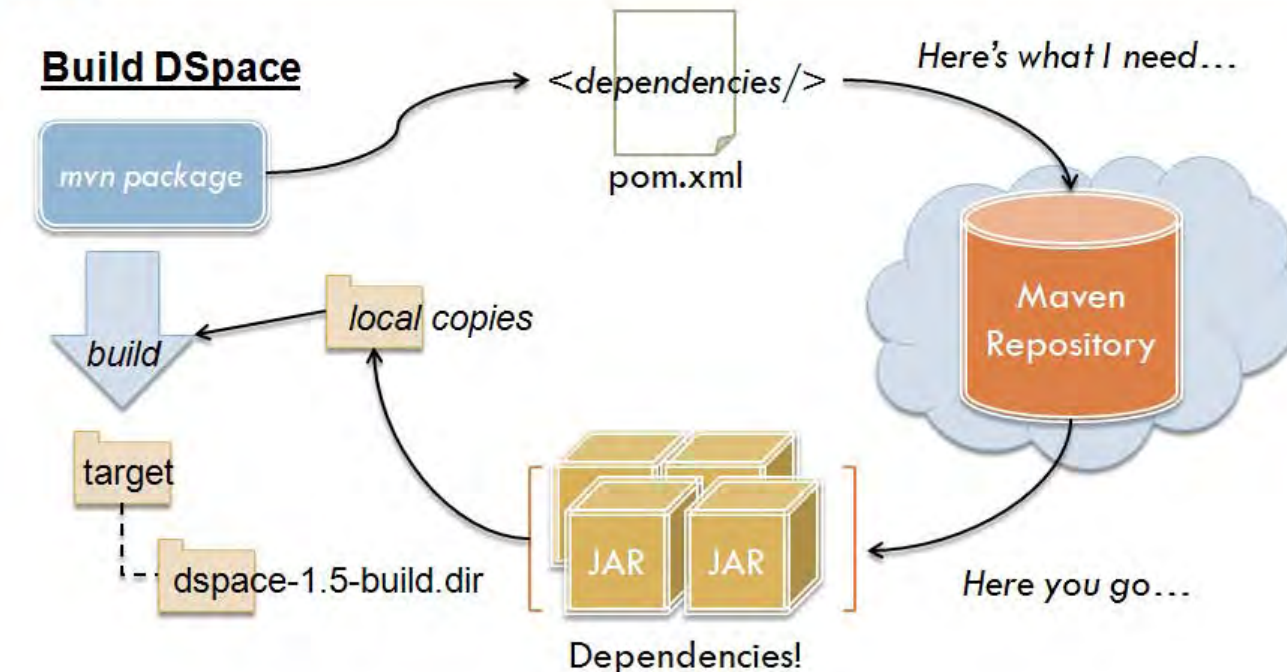
- LNI (Lightweight Network Interface)



リポジトリシステム

DSpace 1.5 Maven

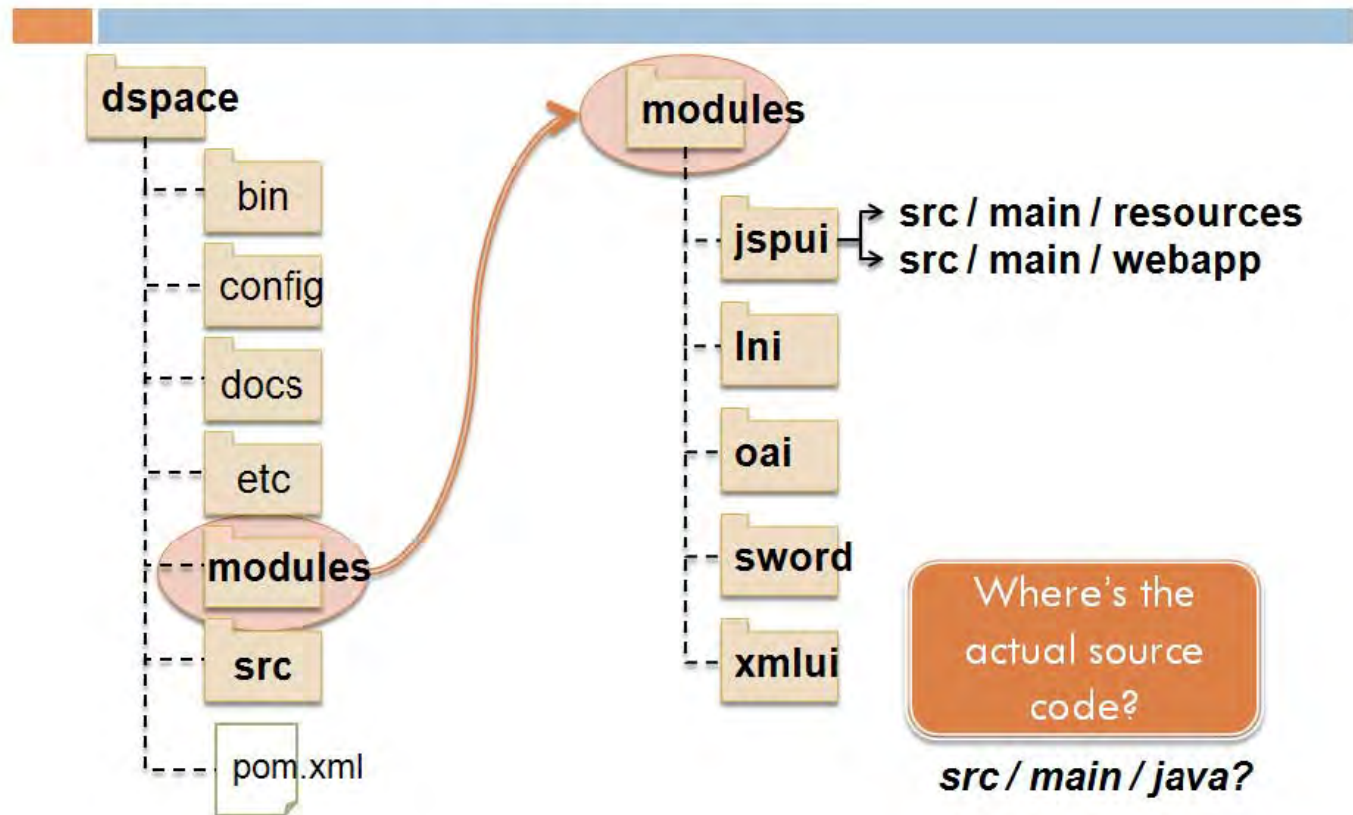
Ok, so what really is Maven?



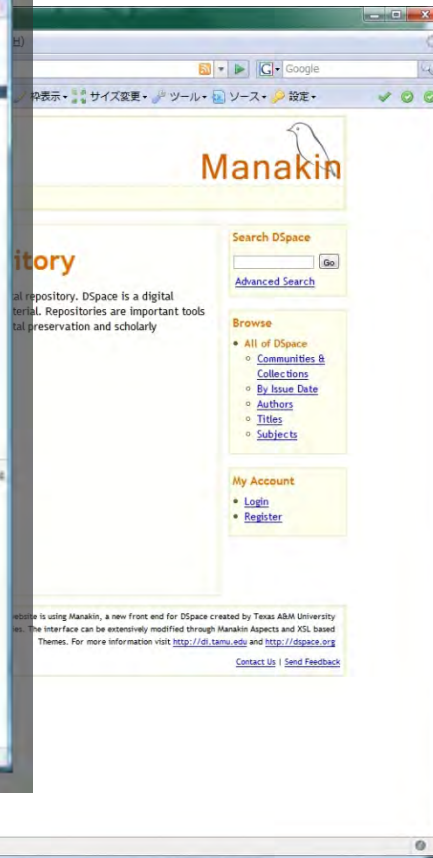
リポジトリシステム

DSpace 1.5 Maven

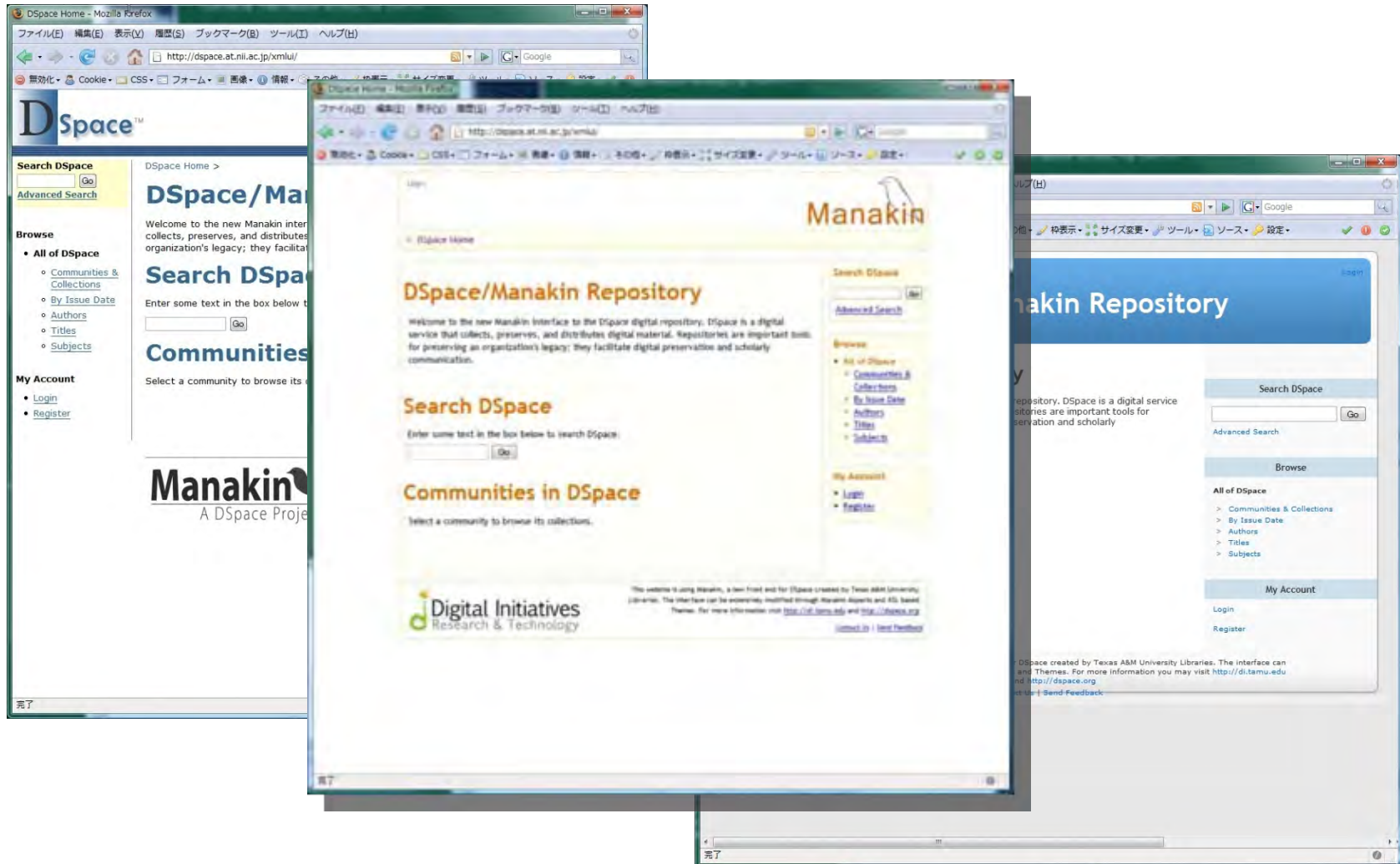
DSpace 1.5 “Modules”



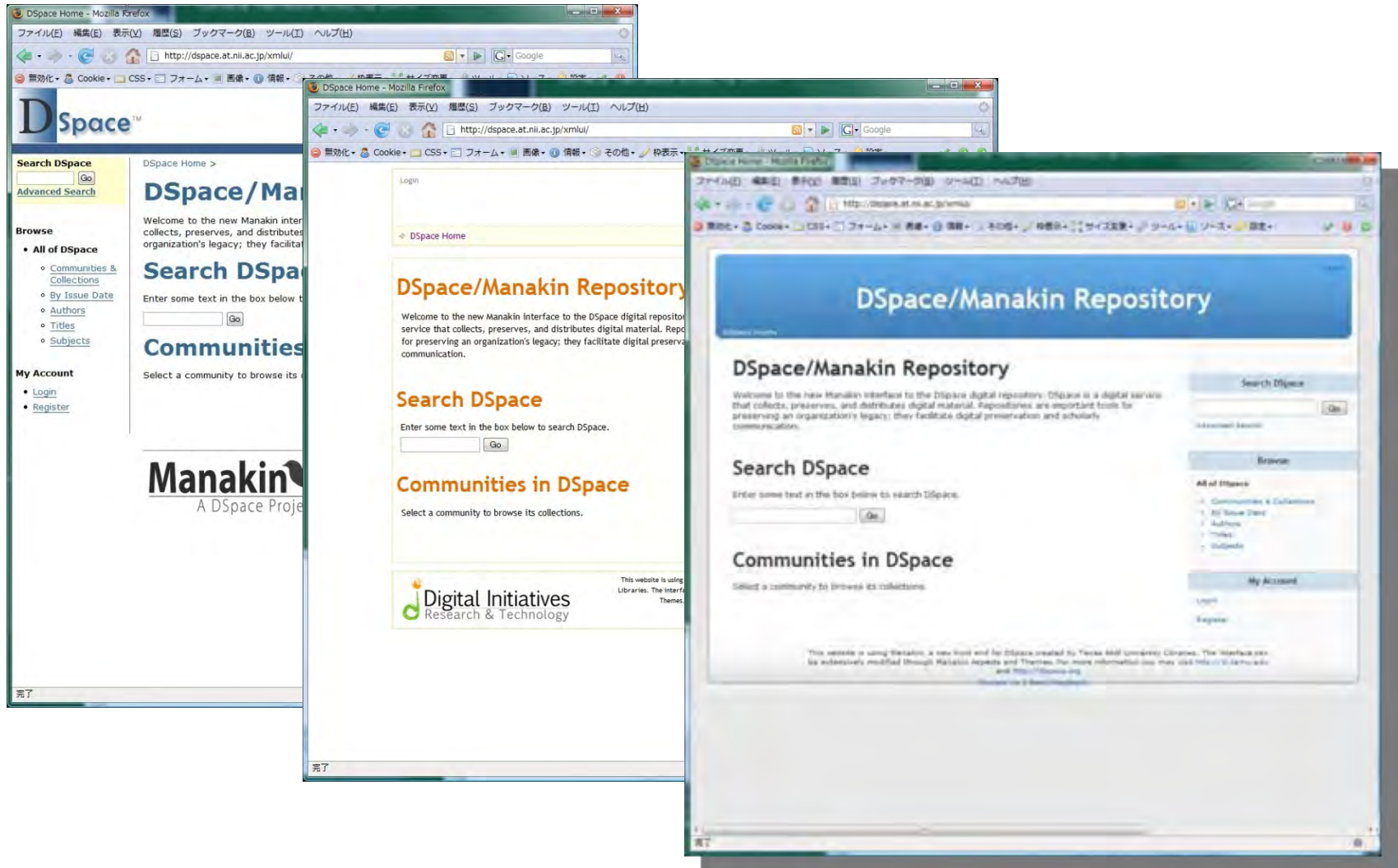
リポジトリシステム DSpace 1.5 Manakin



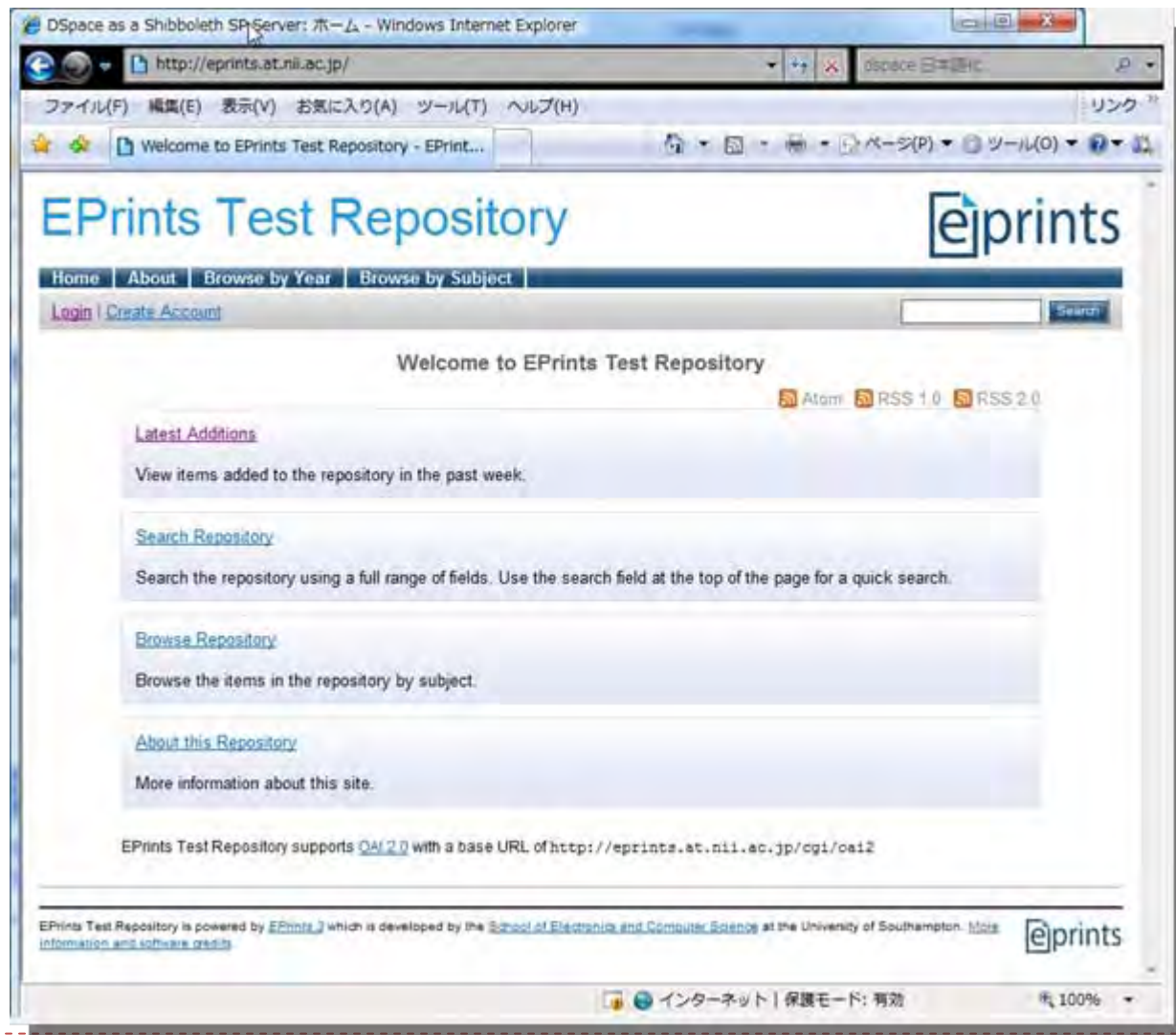
リポジトリシステム DSpace 1.5 Manakin



リポジトリシステム DSpace 1.5 Manakin



リポジトリシステム EPrints デモサイト



► <http://eprints.at.nii.ac.jp/>

リポジトリシステム EPrints OR08サイト



<http://pubs.or08.ecs.soton.ac.uk/>

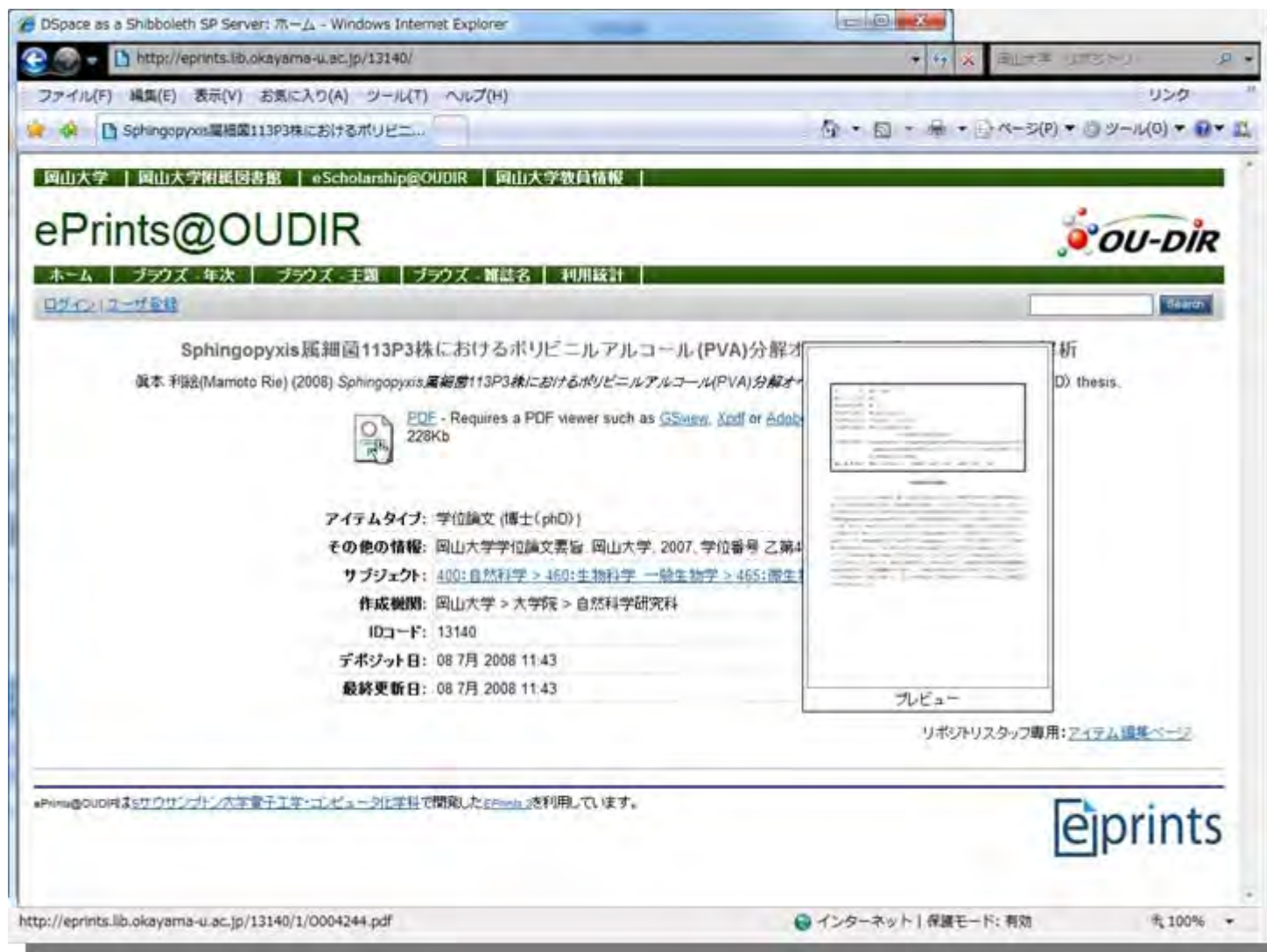
リポジトリシステム EPrints 岡山大学OU-DIRサイト

The screenshot shows the ePrints@OUDIR website in a Windows Internet Explorer browser window. The address bar shows the URL <http://eprints.lib.okayama-u.ac.jp/>. The page features a green header with navigation links: 岡山大学 | 岡山大学附属図書館 | eScholarship@OUDIR | 岡山大学教員情報. Below the header, the main title "ePrints@OUDIR" is displayed, followed by a sub-header with links: ホーム | ブラウズ・年次 | ブラウズ・主題 | ブラウズ・雑誌名 | 利用統計. A search bar is located on the right side of the sub-header.

The main content area is divided into two columns. The left column, titled "最新情報" (Latest Information), lists recent publications with their titles, dates, and authors. The right column, titled "最新のアイテム" (Latest Items), lists specific publications with their titles, authors, and publication details.

At the bottom of the page, there are logos for OALIST, Junii, and Digital Library of Okayama University. The footer contains the text "ePrints@OUDIRは岡山大学電子図書館によって開発されたePrintsを利用しています。" and the ePrints logo.

リポジトリシステム EPrints 岡山大学OU-DIRサイト



リポジトリシステム WEKOとは

- ▶ NIIが開発・公開中のCMSネットコモンズ2.0で動作するリポジトリシステム
- ▶ DSpace, EPrintsなどと等価な機能を装備



リポジトリシステム

WEKO 機能

▶ ゲスト

- ▶ ディレクトリサーチ
- ▶ キーワードサーチ
- ▶ メタデータ&全文
- ▶ ランキング

▶ 登録ユーザ

- ▶ アイテム登録
- ▶ ワークフロー

▶ 管理者

- ▶ アイテムタイプ管理
 - ▶ 作成／編集
 - ▶ OAI-PMHマップ
- ▶ アイテム管理
- ▶ ツリー編集
- ▶ 査読／承認
- ▶ インポート
- ▶ ログ解析
- ▶ WEKOカスタム



リポジトリシステム WEKOサイト



リポジトリシステム WEKOデモサイト

The screenshot shows the WEKO demo site at <http://weko.at.nii.ac.jp/demo>. The interface includes a header with the WEKO logo and a brief description of the system. Below the header, there are navigation tabs for 'トップ' (Top) and 'ランキング' (Ranking). A search bar is present, and a sidebar on the left lists various categories under 'インデックスツリー' (Index Tree). The main content area displays a list of items under 'インデックスリスト' (Index List). A blue overlay box in the foreground contains the following text:

アイテム

- 登録するコンテンツ

アイテムタイプ

- アイテムの大まかな種類

インデックスツリー

- アイテムを格納する場所

The background interface also shows a calendar for November 2009, a login status section, and a footer with 'インターネット | 保護モード | 有効' and '100%' zoom level.

リポジトリシステム WEKOサイト



リポジットリシステム Windows 簡単WEKO



内容

1 システム概略

- ▶ リポジトリシステムとは？
- ▶ システム構成

2 リポジトリシステム

- ▶ DSpace
- ▶ EPrints
- ▶ WEKO

3 システム構築

- ▶ OS, リポジトリシステムのインストール
- ▶ バックアップ

4 メタデータ流通

- ▶ OAI-PMH
- ▶ OAI-ORE

目標

- リポジトリシステムがやっていることを理解してもらうこと
- 自分でリポジトリを立てようと思う人を増やすこと



システム構築

OS, リポジトリシステムのインストール

明日午後からのインストールフェストに参加して下さい



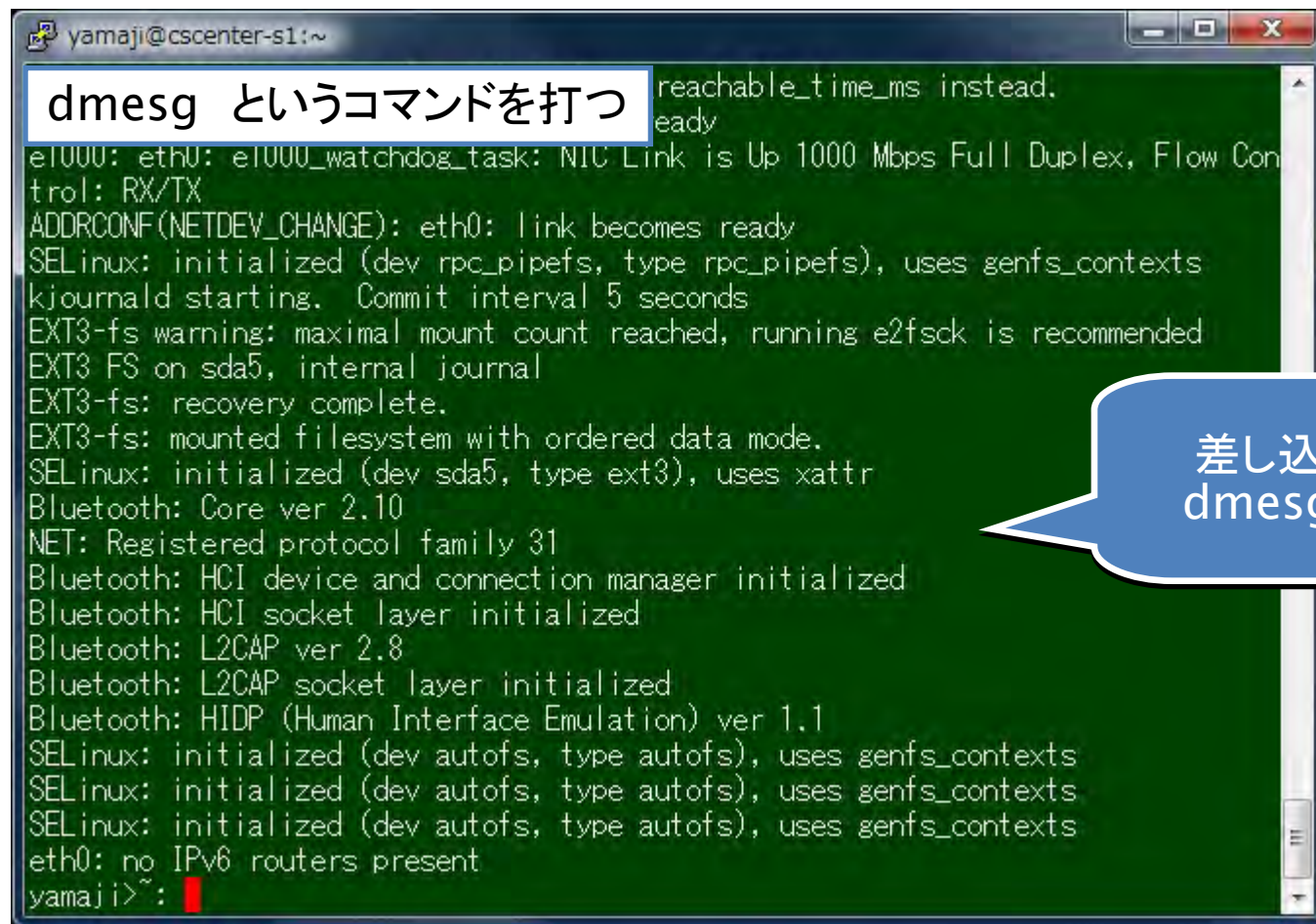
システム構築 バックアップ

- ▶ デバイスの確認
- ▶ デバイスのマウント
- ▶ バックアップスクリプト
- ▶ リストア



システム構築

USBメモリーを差し込むと



A terminal window titled 'yamaji@cscenter-s1:~' with a green background. The prompt 'yamaji>' is at the bottom. A white box highlights the text 'dmesg' and the Japanese text 'というコマンドを打つ' (type this command). The terminal displays various system messages including network link status, SELinux initialization, and filesystem recovery.

```
yamaji@cscenter-s1:~  
dmesg  というコマンドを打つ  
reachable_time_ms instead.  
eady  
e1000: eth0: e1000_watchdog_task: NIC Link is Up 1000 Mbps Full Duplex, Flow Con  
trol: RX/TX  
ADDRCONF(NETDEV_CHANGE): eth0: link becomes ready  
SELinux: initialized (dev rpc_pipefs, type rpc_pipefs), uses genfs_contexts  
kjournald starting. Commit interval 5 seconds  
EXT3-fs warning: maximal mount count reached, running e2fsck is recommended  
EXT3 FS on sda5, internal journal  
EXT3-fs: recovery complete.  
EXT3-fs: mounted filesystem with ordered data mode.  
SELinux: initialized (dev sda5, type ext3), uses xattr  
Bluetooth: Core ver 2.10  
NET: Registered protocol family 31  
Bluetooth: HCI device and connection manager initialized  
Bluetooth: HCI socket layer initialized  
Bluetooth: L2CAP ver 2.8  
Bluetooth: L2CAP socket layer initialized  
Bluetooth: HIDP (Human Interface Emulation) ver 1.1  
SELinux: initialized (dev autofs, type autofs), uses genfs_contexts  
SELinux: initialized (dev autofs, type autofs), uses genfs_contexts  
SELinux: initialized (dev autofs, type autofs), uses genfs_contexts  
eth0: no IPv6 routers present  
yamaji>:
```

差し込む前の
dmesgの内容

システム構築

デバイス名の確認

```
yamaji@cscenter-s1:~  
chable_time; Use net.ipv6.neigh.lo.base_reachable_time ms instead.  
ADDRCONF(NETDEV_UP):  
e1000: eth0: e1000_w  
trol: RX/TX  
ADDRCONF(NETDEV_CHAN  
SELinux: initialized (dev autofs, type autofs), uses genfs_contexts  
SELinux: initialized (dev autofs, type autofs), uses genfs_contexts  
SELinux: initialized (dev autofs, type autofs), uses genfs_contexts  
eth0: no IPv6 routers present  
kjournald starting.  
EXT3-fs warning: max  
EXT3 FS on sda5, int  
EXT3-fs: recovery co  
EXT3-fs: mounted fil  
SELinux: initialized  
Bluetooth: Core ver:  
NET: Registered prot  
Bluetooth: HCI devic  
Bluetooth: HCI socke  
Bluetooth: L2CAP ver  
Bluetooth: L2CAP soc  
Bluetooth: HIDP (Hum  
SELinux: initialized  
SELinux: initialized  
SELinux: initialized  
eth0: no IPv6 router  
yamaji>~  
SELinux: initialized (dev autofs, type autofs), uses genfs_contexts  
SELinux: initialized (dev autofs, type autofs), uses genfs_contexts  
SELinux: initialized (dev autofs, type autofs), uses genfs_contexts  
eth0: no IPv6 routers present  
usb 1-2: new high speed USB device using ehci_hcd and address 3  
usb 1-2: configuration #1 chosen from 1 choice  
scsi1 : SCSI emulation for USB Mass Storage devices  
usb-storage: device found at 3  
usb-storage: waiting for device to settle before scanning  
Vendor: USB 2.0 Model: Flash Disk Rev: 1100  
Type: Direct-Access ANSI SCSI revision: 00  
SCSI device sdb: 1007616 512-byte hdwr sectors (516 MB)  
sdb: Write Protect is off  
sdb: Mode Sense: 43 00 00 00  
sdb: assuming drive cache: write through  
SCSI device sdb: 1007616 512-byte hdwr sectors (516 MB)  
sdb: Write Protect is off  
sdb: Mode Sense: 43 00 00 00  
sdb: assuming drive cache: write through  
sdb: sdb1  
sd 1:0:0:0: Attached scsi removable disk sdb  
sd 1:0:0:0: Attached scsi generic sg2 type 0  
usb-storage: device scan complete
```

差し込んだ後の
dmesgの内容

USBメモリがsdb1というデバイスに割り当てられたことがわかる

システム構築

デバイスマウント

▶ 手動でマウント

```
> sudo mkdir /mnt/usb  
> sudo mount -t vfat /dev/sdb1 /mnt/usb
```

▶ 自動マウント

```
> less /etc/fstab  
/dev/VolGroup00/LogVol00 /          ext3      defaults  1 1  
LABEL=/boot              /boot    ext3      defaults  1 2  
devpts                    /dev/pts devpts    gid=5,mode=620 0 0  
tmpfs                    /dev/shm tmpfs     defaults  0 0  
proc                     /proc    proc      defaults  0 0  
sysfs                    /sys     sysfs     defaults  0 0  
/dev/VolGroup00/LogVol01 swap      swap      defaults  0 0  
/dev/sdb1                /mnt/usb vfat      defaults  0 0
```

```
> sudo mount -a
```

一行追加

システム構築

バックアップスクリプト例

▶ ローカルHDDにバックアップ(毎日)

```
> less /etc/cron.daily/backup-html
#!/bin/csh
set dt = `date "+%d"`
set backup_dir = "/var/www/backup/daily"
echo "Start backup datetime=$dt ";

# weko site backup
/usr/bin/mysqldump -u nc2 -h localhost -p*** nc2 > $backup_dir/weko-site$dt.sql
cd /var/www/html/nc2
/bin/tar zcvf $backup_dir/weko-site$dt.tgz * > $backup_dir/weko-site$dt.log

echo "Successfully completed"
```

▶ ローカルHDDにバックアップ(毎月)

- ▶ /etc/cron.monthlyに上記のようなスクリプトを配置



システム構築

バックアップスクリプト例

▶ USB-HDDにバックアップ

```
> /etc/cron.daily/backup-rsync
#!/bin/csh
set backup_dir = "/mnt/usb/backup"
echo "Start backup"

rsync -a --delete /var/www/* $backup_dir/ > $backup_dir/rsync.log

echo "Successfully completed"
```

▶ 確認

- ▶ Logwatchの実行でcronの実行結果がメールで通知
 - ▶ SendmailでSMART_HOSTを設定
 - ▶ Aliasでrootへのメールが自分に届くように設定



システム構築 バックアップのリストア

googleってやり方を調べてみよう



OAI-PMH リポジトリ構築後のステップ



<http://drf.lib.hokudai.ac.jp/drf/index.php?startup>

OAI-PMH

リポジトリ構築後のステップ: ついでに

The screenshot shows a web browser window titled "DSpace as a Shibboleth" with the URL <http://drf.lib.hokudai.ac.jp/drf/index.php?startup>. The page content includes a sidebar with a list of links such as "IR設立・運営に関する参考文献", "実務のための関連資料", "トピック", "イベント開催情報", "機関リポジトリに関する国内のメディア報道一覧", and "システム・ツール関連情報". Below these links is the "SHERPA RoMEO SCPI" logo. The main content area features a "recent(8)" section with a list of updates from 2008-07-10 to 2008-05-21. To the right of this list, there are two sections: "利用の拡大のために" (To expand usage) and "業界への広報のために" (To publicize to the industry). The "利用の拡大のために" section lists links to Google, Google Scholar, OAIster, Scirus, and JuNii?+(NII). The "業界への広報のために" section lists a link to the Budapest Open Access Initiative (BOAI).

DSpace as a Shibboleth

<http://drf.lib.hokudai.ac.jp/drf/index.php?startup>

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) リンク >>

startup - DRF wiki

0

Registering as a Data Provider <http://www.openarchives.org/data/registerasprovider.html>
の一番下に、IRの OAI-PMHのベースURLを入れて送信するだけです。
(機械的なテストののちに登録されます。テストのみを受けることもできます。OAI-PMHとかベースURLとかわからなければ、導
入業者にやってもらってもよいと思います)

利用の拡大のために ⁺

- Google
サイトの登録/削除 <http://www.google.com/addurl/?hl=ja&continue=/addurl>
からクロールを申し込みましょう。
- Google Scholar
「出版社向けサポート」<http://scholar.google.com/intl/ja/scholar/publishers.html>
を参照のこと。「図書館向けサポート」ではありません。
- OAIster
「Steps to Becoming a Data Contributor」<http://www.oaister.org/dataproviders.html>
を参照のこと。メールで申し込む方式に変わっています。
- Scirus
- JuNii?+(NII)
「JuNii?+ ハーベスタ機能実装仕様」
http://www.nii.ac.jp/irp/info/harvest/junii2_implementation_as_harvester.html

業界への広報のために ⁺

- Budapest Open Access Initiative (BOAI)

recent(8)

2008-07-10
・ Digital Repository Federation
・ 参加機関一覧
・ 参加機関一覧 (第一期)

2008-07-08
・ Journals

2008-06-26
・ DSpaceと Google Sitemap
・ Events

2008-06-04
・ MenuBar

2008-05-21
・ press

インターネット | 保護モード: 有効 100%



<http://drf.lib.hokudai.ac.jp/drf/index.php?startup>

OAI-PMH

登録データの流通

登録したデータをどのようにして広めていくか？

- ・ Googleの検索対象になるように努力
- ・ データベース間で明示的にネットワークを確立

OAI-PMH

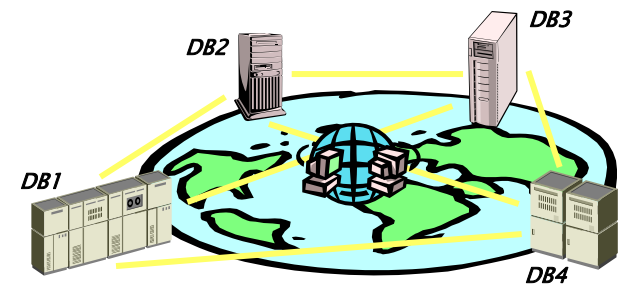
(Open Archive Initiative – Protocol for Metadata Harvesting)

サーバ間でメタデータを機械的に収集するための、
要求と結果を授受するためのプロトコル。

1999年10月：サンタフェ会議でFrameworkが合意

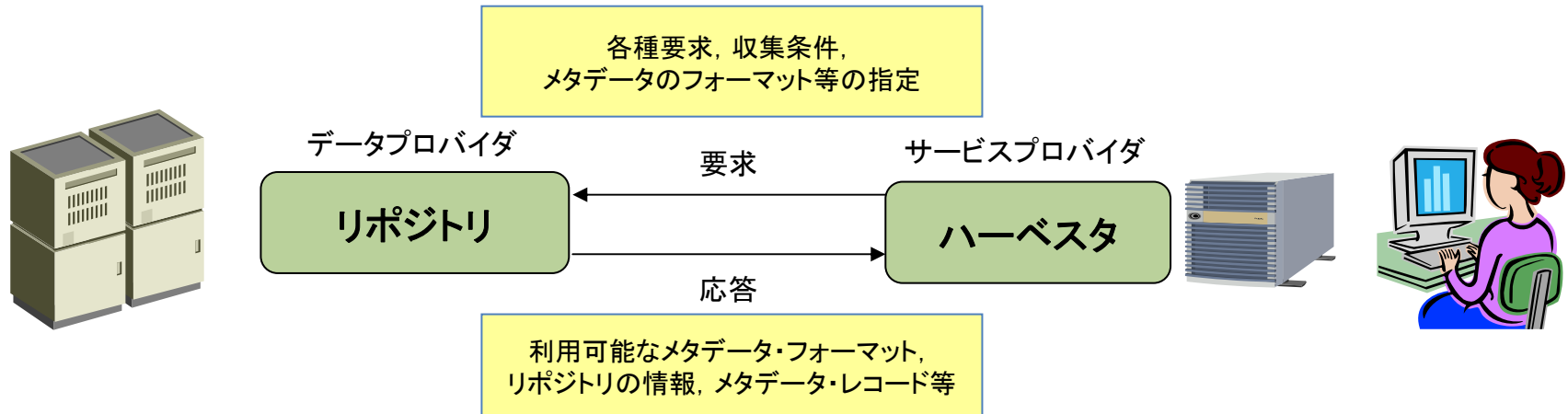
2001年 1月：OAI-PMH 第1版

2002年 6月：OAI-PMH 第2版



OAI-PMH

メタデータ通信



*Dublin Core*メタデータセットでの交換を必須

Title	Creator	Subject	Description	Publisher
Contributor	Date	Type	Format	Identifier
Source	Language	Relation	Coverage	Rights

OAI-PMH

junii2フォーマット①

メタデータフォーマット (junii2) 各データ要素の入力内容一覧

凡例: 青=必須、黄=あれば必須 表記以外=各機関リポジリでのポリシーによる

通番	大項目	項目	element	内容	必須	継返し可否	minOccurs	maxOccurs	書式(空白の場合は文字列)	xml:langにダウソクさせる ときの行先	オプションの属性
1	タイトル	タイトル	title		必須	X	1	1		title	lang
2		その他(別題等)のタイトル	alternative	別題等のタイトルやタイトルの注			0	unbounded		title	lang
3		作成者	creator		あれば必須		0	unbounded		creator	lang
4	主題	著者キーワード	subject				0	unbounded		subject	
5		国立情報学研究所 メタデータ主観語集	NDsubject	下記URL参照 http://www.nii.ac.jp/metadata/manual/ND-subject.pdf			0	unbounded		subject	version
6		日本十進分類法	NDCC				0	unbounded		subject	version
7		国立国会図書館分類表	NDLC				0	unbounded		subject	version
8		日本学名目録	BSH				0	unbounded		subject	version
9		国立国会図書館前名目録	NDLSH				0	unbounded		subject	version
10		医学抄名目録	MeSH				0	unbounded		subject	version
11		デューイ十進分類法	DDC				0	unbounded		subject	version
12		米国国会図書館分類表	LCC				0	unbounded		subject	version
13		国際十進分類法	UDC				0	unbounded		subject	version
14		米国国会図書館前名目録	LCSH				0	unbounded		subject	version
15		内容記述	description	論文の抄録 学校論文にあっては、学年年度、論文/題名の種別等 会議発表論文にあっては、会議名、開催地、日時等 その他コンテンツの内容を示すことを記述			0	unbounded		description	
16		公開者	publisher	コンテンツ本体の公開を行った個人、機関、団体			0	unbounded		publisher	lang
17		提供者	contributor	コンテンツ本体の作成に関与した個人、機関、団体			0	unbounded		contributor	lang
18		日付	date	コンテンツの作成日付			0	unbounded	YYYY-MM-DD, YYYY-MM, YYYY	date	
19		資源タイプ	type	コンテンツの種類(自由記述)			0	unbounded		type	
20		国立情報学研究所 メタデータ主観語集 (資源タイプ)	NDtype		必須	X	1	1	[Journal Article Thesis or Dissertation Departmental Bulletin Paper Conference Paper Presentation Book Technical Report Research Paper Article Preprint Learning Material Data or Dataset Software Others]	type	
21	資源 識別子	フォーマット	format	Internet Media Typeや規定されるファイル形式			0	unbounded		format	
22		その他の資源識別子	identifier	QRコード、24項以外の資源識別子			0	unbounded		identifier	
23		資源識別子URL (アイテム表示画面へのリンク)	URL	コンテンツ本体を表示する画面へのURL	必須	X	1	1	URL	identifier	
24		本文フルテキストへのリンク	fullTextURL	コンテンツ本体へのURL	あれば必須		0	unbounded	URL	identifier	
25	掲載誌 情報	ISSN	issn		あれば必須		0	unbounded	ISSN-10 or ISSN-X	採りたため識別子の 上、 identifier	
26		雑誌レコードID(国会目録ID)	NCID		あれば必須		0	unbounded			
27		雑誌名	title	本タイトルのみ	あれば必須		0	1			lang
28		巻	volume		あれば必須		0	1			
29		号	issue		あれば必須		0	1			
30		開始ページ	scope	論文の開始ページ	あれば必須		0	1	半角数字		
31		終了ページ	scope	論文の終了ページ	あれば必須		0	1	半角数字		
32		刊行年月	dateIssued		あれば必須		0	1	YYYY-MM-DD, YYYY-MM, YYYY		
33		情報源	source				0	unbounded		source	

OAI-PMH

junii2フォーマット②

番号	大項目	項目	element	内容	必須	継続し可否	minOccurs	maxOccurs	書式(空白の場合は文字列)	xml:spaceにダムダウンせざるべきの行先	オプションの属性
34		言語	language	コンテンツ本文の言語			0	unbounded	[ja, jpn, ja-JP]	language	
35		他の資源との関係	relation				0	unbounded		relation	
36	ID	URI/Media番号	uri				0	1	URI (info:uri/で記述すること)	relation	
37		DOI	doi				0	1	URI (info:doi/で記述すること)	relation	
38	関係に属する情報	異版である	isVersionOf	当該コンテンツに対して、参照先のコンテンツを主たる版とみなすことができる。			0	unbounded	URI	relation	
39		異版あり	hasVersion	当該コンテンツが参照先のリソースを異なる版として持つ			0	unbounded	URI	relation	
40		置換される	isReplacedBy	当該コンテンツが参照先のリソースによって置き換わる			0	unbounded	URI	relation	
41		置換する	replaces	当該コンテンツが参照先のコンテンツを新しく置き換える			0	unbounded	URI	relation	
42		要件とされる	isRequiredBy	参照先のコンテンツを利用するにあたり、当該コンテンツを必要とする			0	unbounded	URI	relation	
43		要件とする	requires	当該コンテンツを利用するにあたり、参照先のコンテンツを必要とする			0	unbounded	URI	relation	
44		部分である	isPartOf	当該コンテンツが参照先コンテンツの部分となる			0	unbounded	URI	relation	
45		部分を持つ	hasPart	当該コンテンツが参照先のコンテンツを部分として持つ			0	unbounded	URI	relation	
46		参照される	isReferencedBy	当該コンテンツが参照先のコンテンツから継受けられる			0	unbounded	URI	relation	
47		参照する	references	当該コンテンツが参照先のコンテンツを継受け持っている			0	unbounded	URI	relation	
48		別フォーマットである	isFormatOf	当該コンテンツが参照先のコンテンツと内容的に同じであるが、Formatは異なる			0	unbounded	URI	relation	
49		別フォーマットあり	hasFormat	当該コンテンツが参照先コンテンツに異なるフォーマットを持つ			0	unbounded	URI	relation	
50	地理的範囲	範囲	coverage				0	unbounded		coverage	
51		空間的	spatial				0	unbounded		coverage	
52		国立情報学研究所 メタデータ主題語彙集 (地域)	NSpatial				0	unbounded		coverage	
53		時間的	temporal				0	unbounded		coverage	
54	時系列範囲	国立情報学研究所 メタデータ主題語彙集 (時代)	NStemporal				0	unbounded		coverage	
55		権利	rights	コンテンツ本体に關する権利規定			0	unbounded		rights	
56		著者版フラグ	textVersion	コンテンツ本体のバージョン author:著者 publisher:出版社 none:上記以外			0	1	[author / publisher / none]		

	オプション属性	書式
lang		PPC1766
version		最新

OAI-PMH

JuNii2資源タイプ

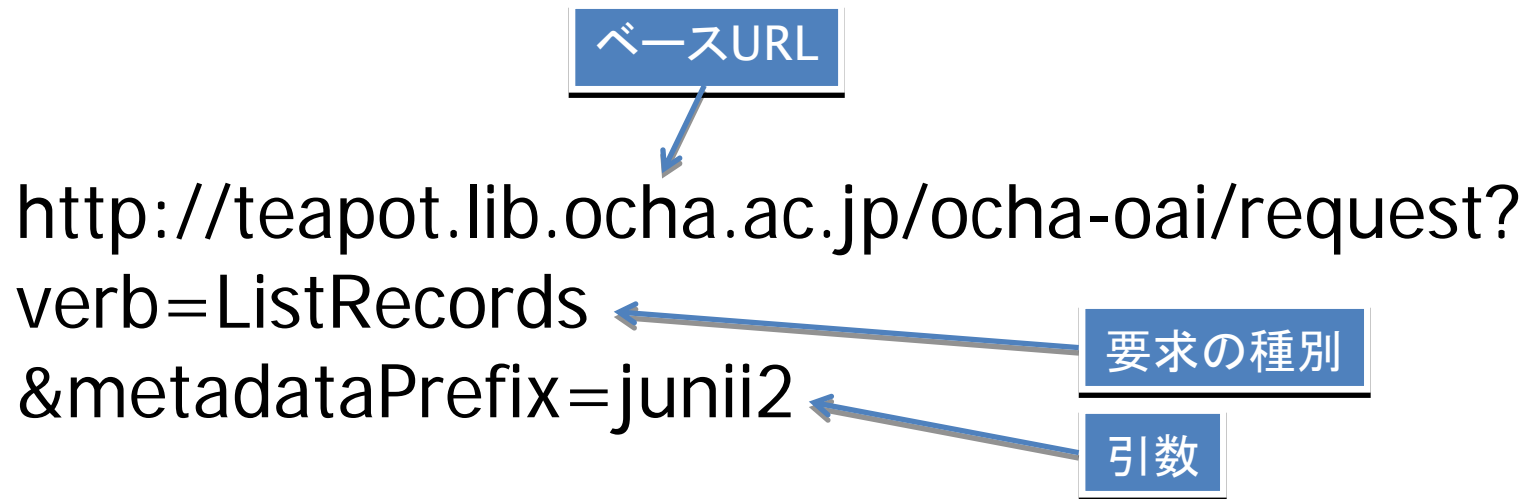
項目	内容	
Journal Article	学術雑誌論文	学術雑誌に掲載された論文
Thesis or Dissertation	学位論文	博士論文、修士論文等 授与年度、論文／課程、学位の種類等は必要に応じてDescription に記入する。
Departmental Bulletin Paper	紀要論文	紀要類に掲載された論文
Conference Paper	会議発表論文	会議の報告書等に掲載された論文 会議名、開催地、日時等は必要に応じてDescription に記入する。
Presentation	会議発表用資料	会議で発表されたプレゼンテーション資料、ポスター、口頭発表資料等 会議名、開催地、日時等は必要に応じてDescription に記入する。
Book	図書	図書の章(Chapter)も含む。
Technical Report	テクニカルレポート	テクニカルレポート、ディスカッションペーパー、ワーキングペーパー等の機関発行の報告書
Research Paper	研究報告書	科研費やCOE 等、研究助成金による研究成果の報告書 種別（助成元、研究課題の種類）等は必要に応じてDescription に記入する。
Article	一般雑誌記事	学術論文以外の記事（コラム等）
Preprint	プレプリント	プレプリント
Learning Material	教材	授業等で用いる資料類
Data or Dataset	データ・データベース	実験記録等のファクトデータ、及びそれらの集合からなるデータベース等
Software	ソフトウェア	ソフトウェア
Others	その他	上記以外のもの



http://www.nii.ac.jp/irp/archive/system/pdf/type_NII.pdf

OAI-PMH

要求例



リポジトリ (`http://teapot.lib.ocha.ac.jp/ocha-oai/request`) から
junii2メタデータフォーマットで公開されているメタデータを
全て取得せよ

OAI-PMH

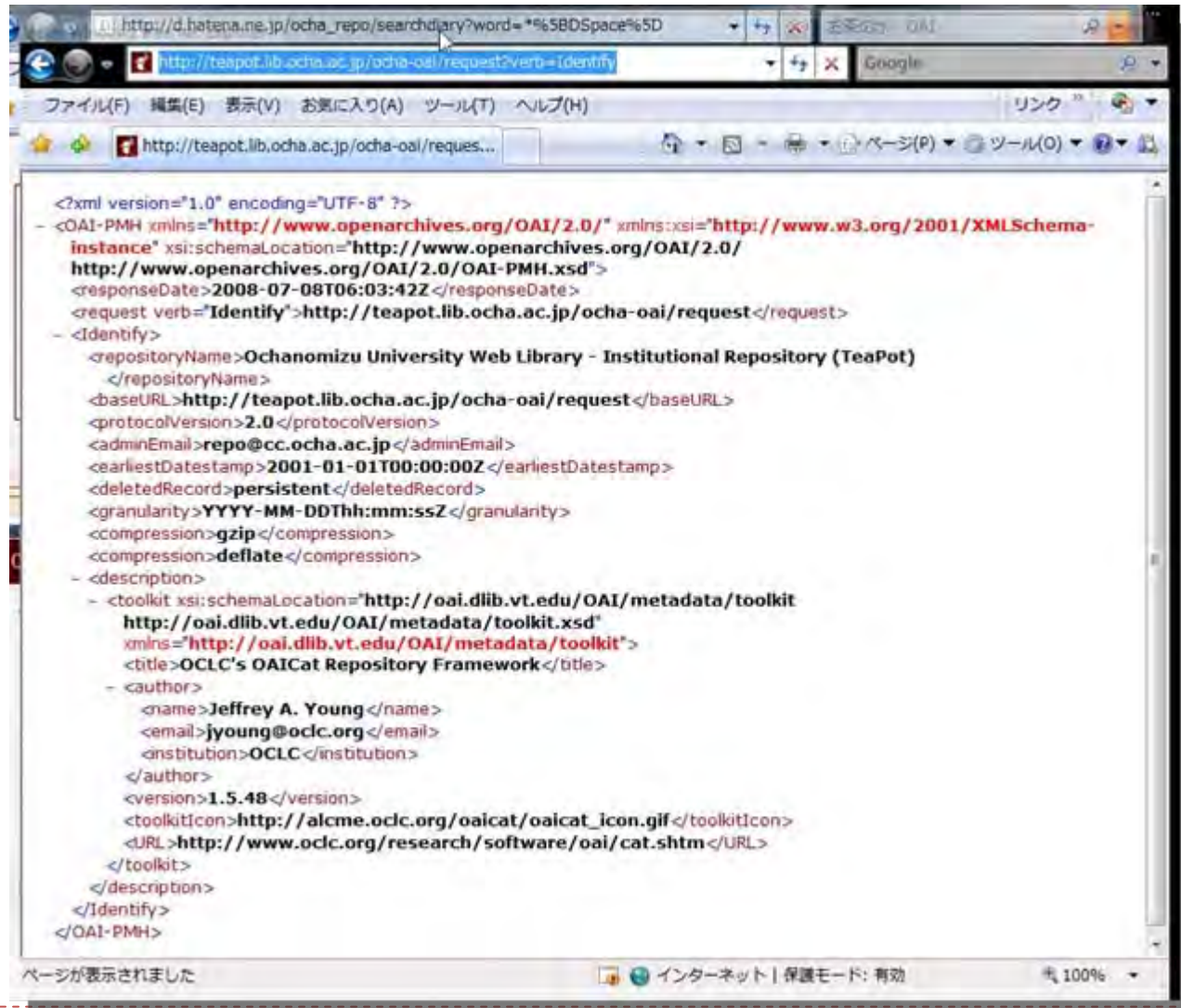
要求(verb)一覧

	要求(動詞)	機能
リポジトリに関する情報を得るための要求	Identify	リポジトリに関する情報を取得する
	ListMetadataFormats	リポジトリにおける利用可能なメタデータ・フォーマットの一覧を取得する
	ListSets	リポジトリのセット構造を取得する
メタデータ収集のための要求	ListIdentifiers	リポジトリからレコード中のヘッダー情報のみを取得する
	ListRecords	リポジトリから条件に合致するレコードを全て取得する
	GetRecord	リポジトリから個々のレコードを取得する

予備: <http://weko.at.nii.ac.jp/demo/oai?verb=Identify>



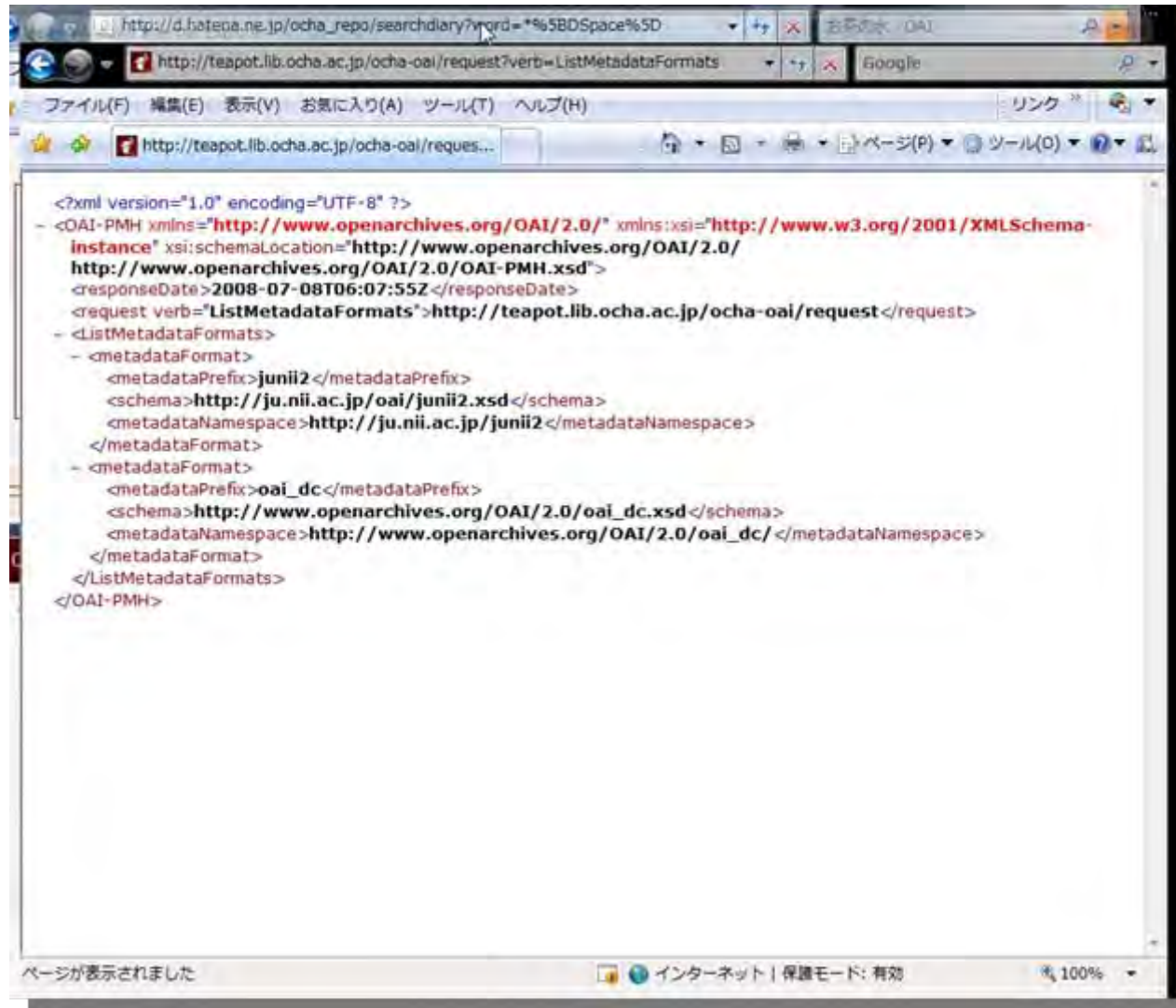
OAI-PMH Identify



▶ <http://teapot.lib.ocha.ac.jp/ocha-oai/request?verb=Identify>

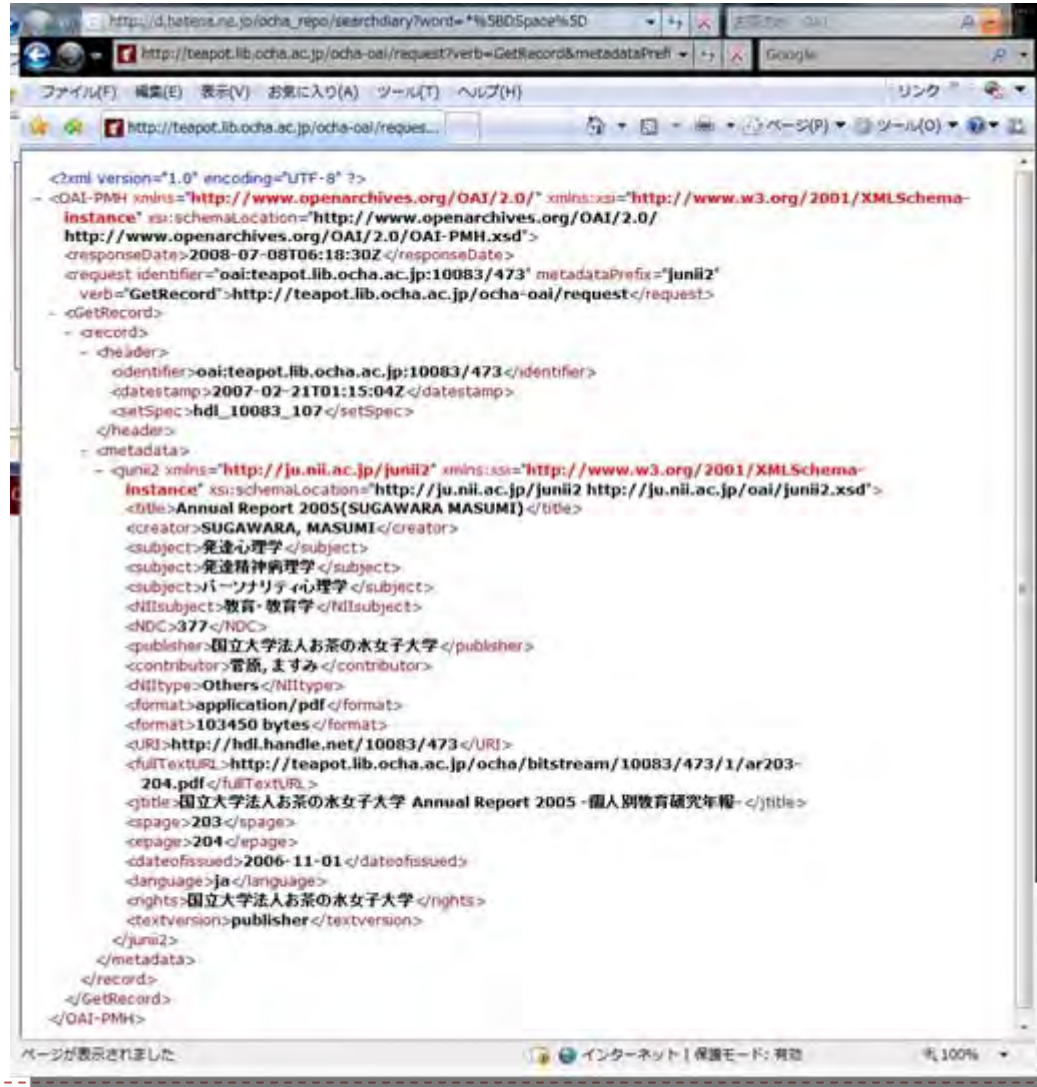
OAI-PMH

ListMetadataFormats



▶ <http://teapot.lib.ocha.ac.jp/ocha-oai/request?verb=ListMetadataFormats>

OAI-PMH GetRecord

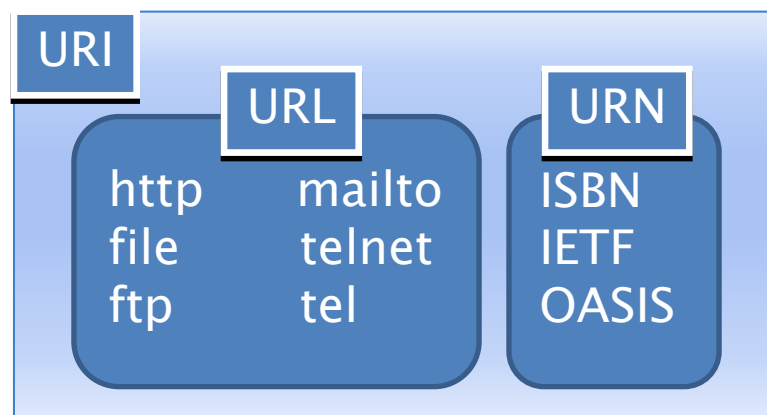


▶ <http://teapot.lib.ocha.ac.jp/ocha-oai/request?verb=GetRecord&metadataPrefix=junii2&identifier=oai:teapot.lib.ocha.ac.jp:10083/473>

OAI-PMH

参考情報: 情報の固定化

- ▶ URI Uniform Resource Identifier
 - ▶ リソースを特定する文字列
- ▶ URL Uniform Resource Locator
 - ▶ リソースをアクセスする方法で特定
 - ▶ `http://domain/path` , `mailto:mbox@domain`
 - ▶ <http://hdl.handle.net/1721.1/29466>
- ▶ URN Uniform Resource Name
 - ▶ リソースを恒久的な名前で特定
 - ▶ `urn:isbn:4-8399-1796-5`
 - ▶ `urn:handle:1721.1/29466`



- HandleやDOIは単なるURN
- 大切なのはWeb資源の固定化

OAI-ORE



► <http://www.openarchives.org/ore/>

OAI-ORE

「ORE仕様書とユーザガイド 日本語訳」

<http://www.nii.ac.jp/irp/archive/translation/oai-ore/1.0/>



The screenshot shows a web browser window displaying the OAI-ORE 1.0 Japanese translation page. The page header includes the Open Archives Initiative logo and the text "Open Archives Initiative Object Reuse and Exchange". The main title is "ORE仕様書およびユーザガイド - 目次" (ORE Specification and User Guide - Table of Contents), dated "2008年10月17日". The page content is organized into sections with bullet points:

- ORE入門** (ORE Introduction): A section for those who don't know OAI-ORE, introducing basic concepts and the specification.
- ユーザガイド** (User Guide): A section providing methods for creating various resource maps and publishing them on the web. It lists several implementation guides:
 - Atomによるリソースマップの実装 (Implementation of Resource Maps using Atom)
 - RDF/XMLによるリソースマップの実装 (Implementation of Resource Maps using RDF/XML)
 - RDFaによるリソースマップの実装 (Implementation of Resource Maps using RDFa)
 - HTTPによる実装 (Implementation using HTTP)
 - リソースマップの発見 (Discovery of Resource Maps)
- 仕様書** (Specification): A section defining the vocabulary for entities and properties used in the OAI-ORE data model. It lists:
 - 抽象データモデル (Abstract Data Model)
 - 語彙 (Vocabulary)
- 次のツールとリソースを使ってOAI-OREベースのアプリケーションを構築して欲しい。** (We want you to build applications based on OAI-ORE using the following tools and resources.):
 - ツールおよび追加リソース (Tools and Additional Resources)
 - GoogleグループOAI-ORE上の意見と議論 (Opinions and Discussions on the OAI-ORE Google Group)

At the bottom, there is a section for versions:

- 本バージョン: <http://www.openarchives.org/ore/1.0/toc>
- 最新バージョン: <http://www.openarchives.org/ore/toc>
- 旧バージョン: <http://www.openarchives.org/ore/0.9/toc>

The page also features a Creative Commons Attribution-Noncommercial-Share Alike 3.0 Unported License logo and text stating that the book is licensed under this license.

OAI-ORE

どんなものなのか？

インターネット上のリソースの関連付けを明確にして発信

▶ 対象

- ▶ 関連付けられたHTMLページのセット
- ▶ ソーシャルネットワークサイト内でのリソースセット
 - ▶ 本文, 画像, コメント
- ▶ リポジトリ内での複数フォーマットからなるリソースセット
 - ▶ LaTeX, PDF, HTML
- ▶ オーバーレイジャーナル

▶ 御利益

- ▶ サーチエンジンがより賢くクロールできる
- ▶ ブラウザがより賢くナビゲートできる
- ▶ コンテンツアーカイブのためのシステムがより賢く働く
- ▶ 引用関係などの情報をより賢く機械処理できる

学術情報のリユースを可能にし次世代学術コミュニケーションを形成



税金や授業料を何に使うか？

- ▶ リポジトリシステムは発展途上
 - ▶ 現在でも新しい機能がどんどん開発・追加されている
 - ▶ リポジトリをシステムレベルで理解することの重要性
 - ▶ リポジトリとは何かを考えることが重要
 - ▶ リポジトリに関する思想はシステムの機能として反映されている
- ▶ システム構築はそれほど難しくない
 - ▶ システム導入のハードルは低くなってきた
 - ▶ 経験者、情報サイト、各種サポートベンダーあり
 - ▶ 試験運用や実験では、パソコンレベルでも可
 - ▶ 本格運用でも多分問題ない
 - ▶ バックアップ装置は要必要
 - ▶ 1万円のUSB-HDDでも十分かも知れない
- ▶ システムよりコンテンツが大事：
 - ▶ システム構築にかかるお金は極力少なく
 - ▶ サクッと立ち上げてコンテンツ収集に注力する
 - ▶ **リポジトリとは何かをよ〜く考える**



お気軽にメールください

yamaji@nii.ac.jp

NII

