OUKA
The University of Osaka
Institutional Knowledge Archive

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	

The University of Osaka Institutional Knowledge Archive : OUKA

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

The University of Osaka

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

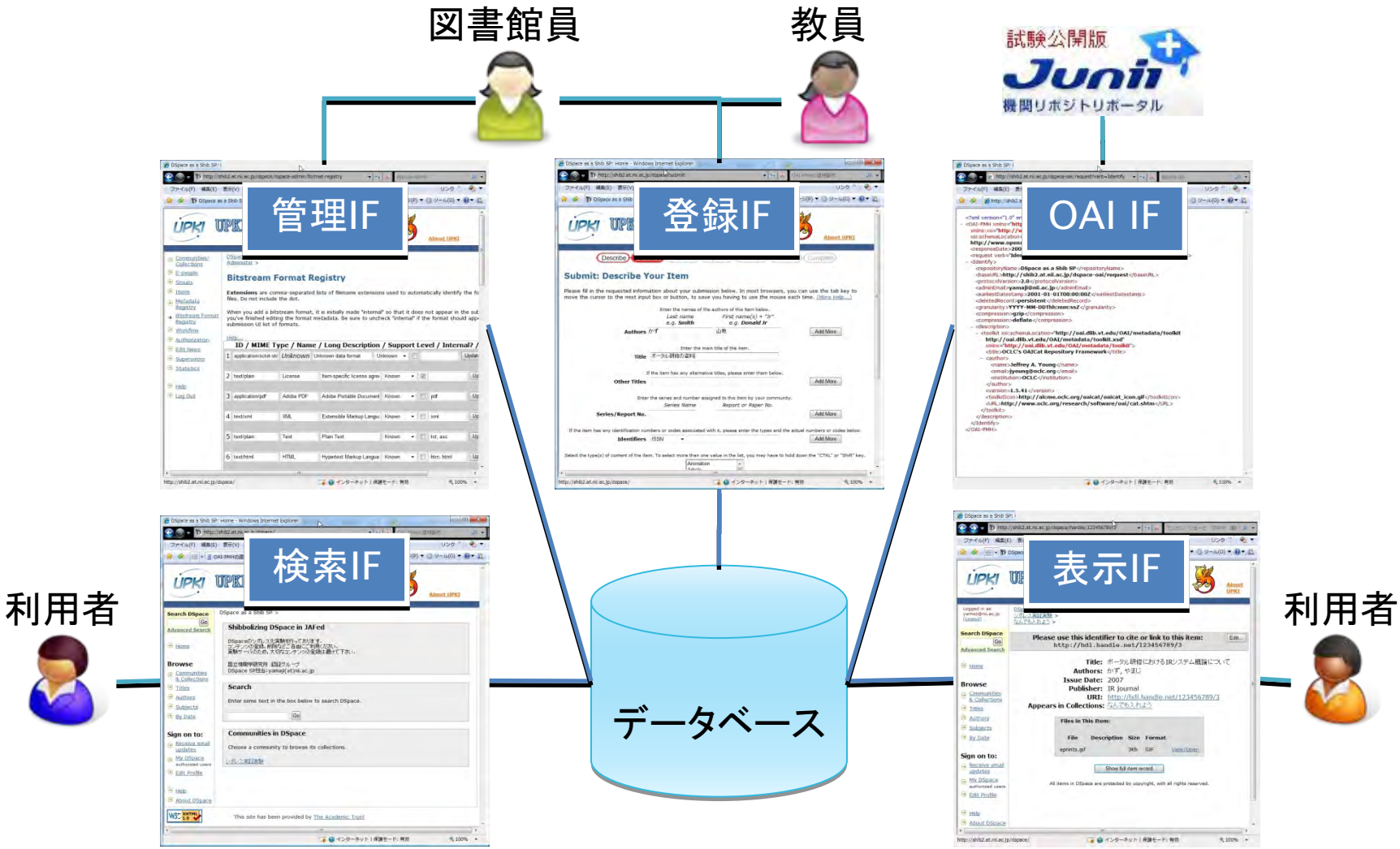
目標

- 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

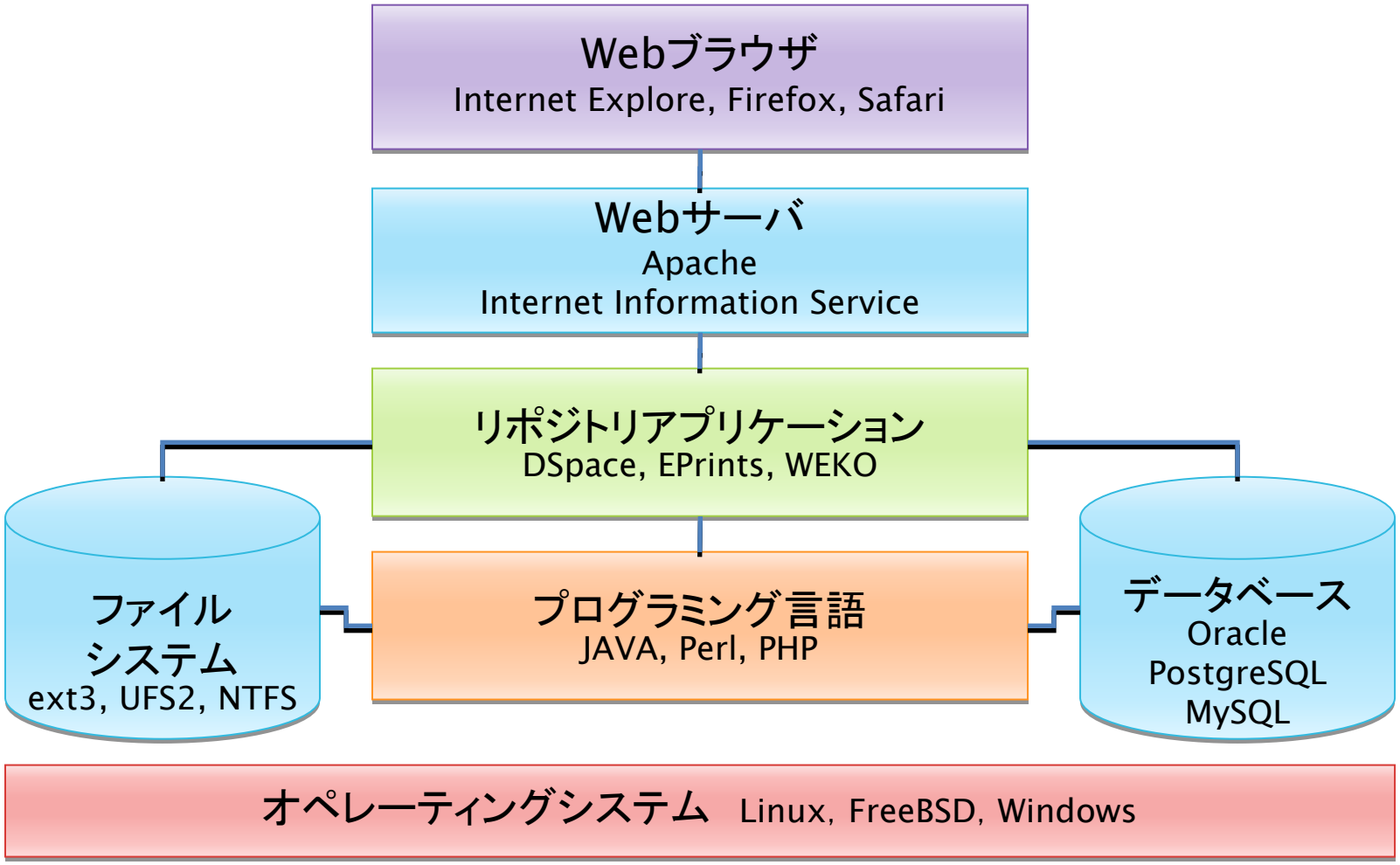
システム概略
必要な機能

- 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

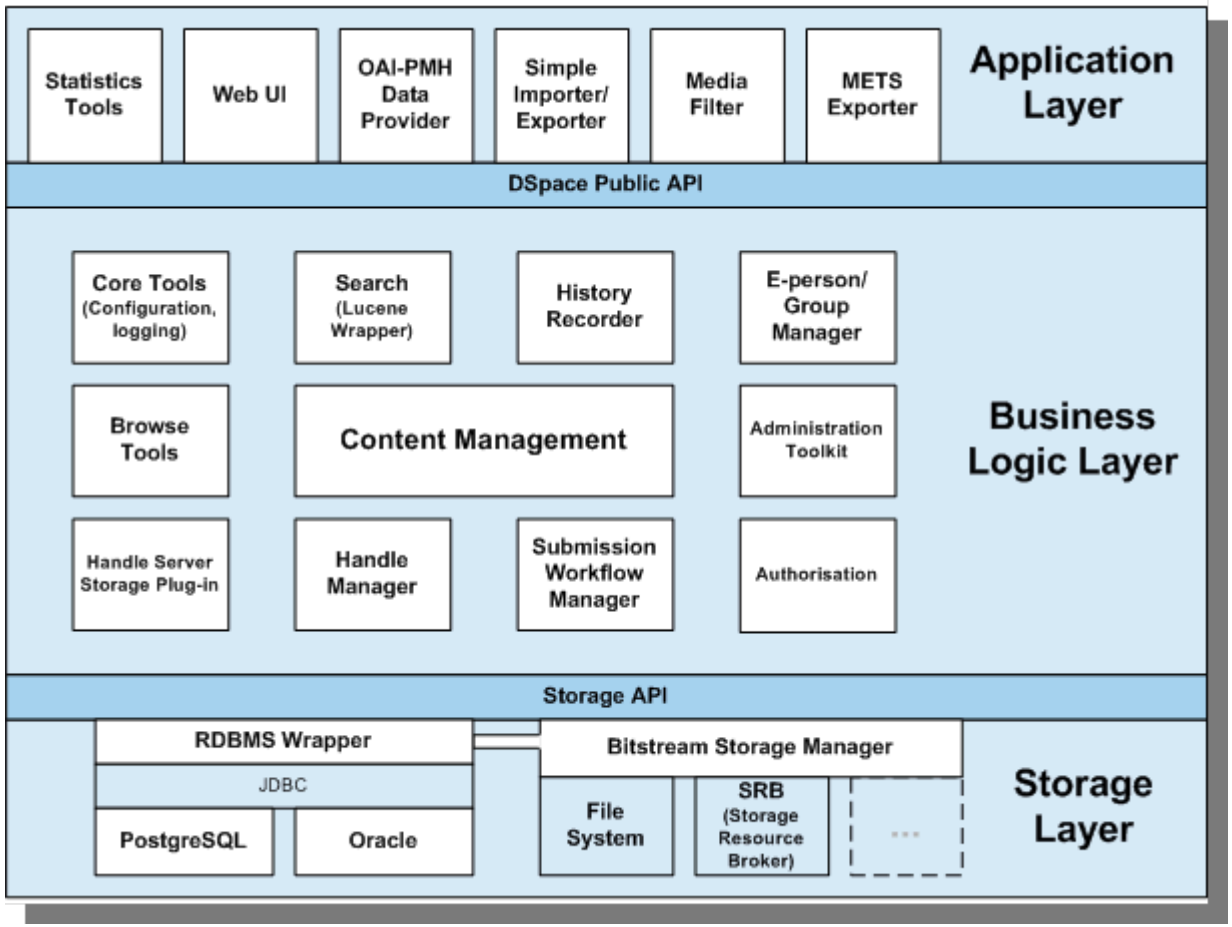
システム概略
Webインターフェース



システム概略
システム構成



システム概略
DSpace Architecture



- http://www.dspace.org/index.php?option=com_content&task=view&id=145

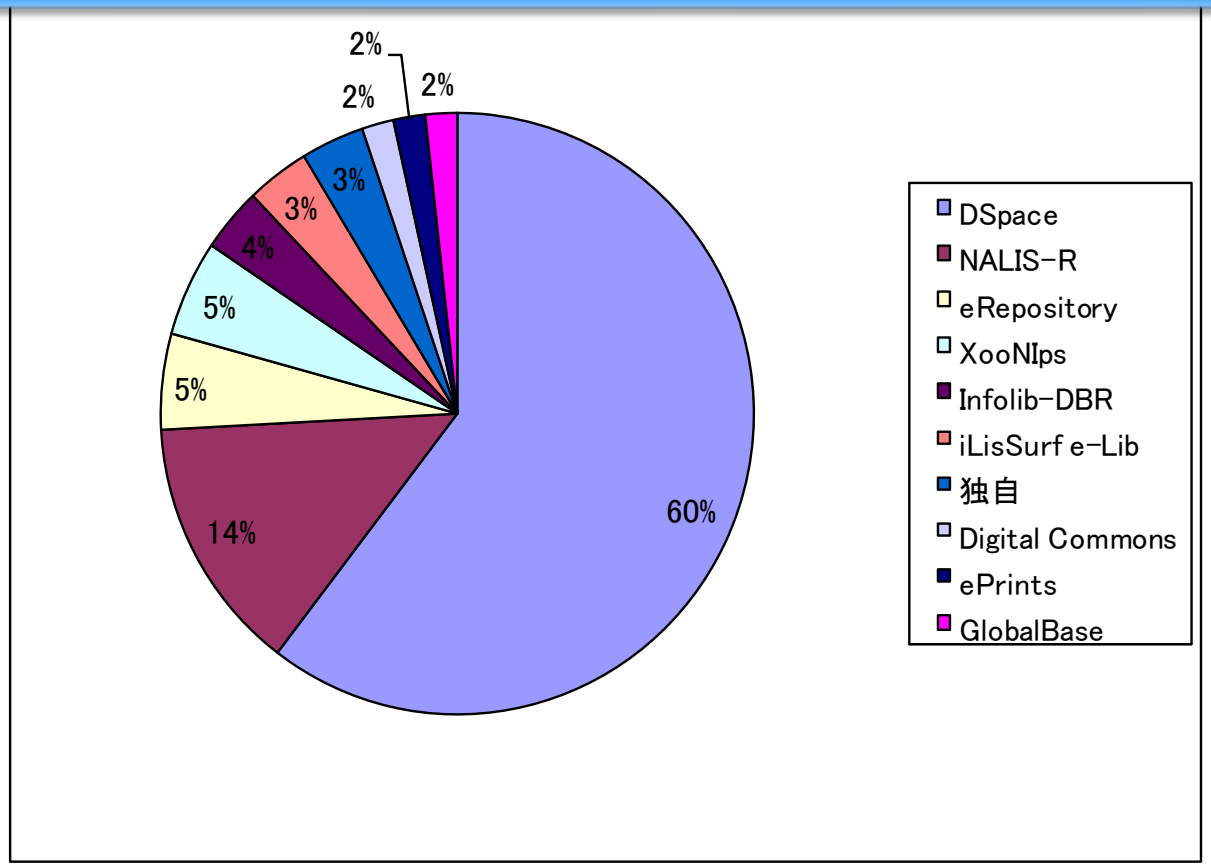
リポジトリシステム
一般的な選択肢

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）



http://www.nii.ac.jp/irp/rfp/2006/pdf/CSIH18report.pdf

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.cmsc.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

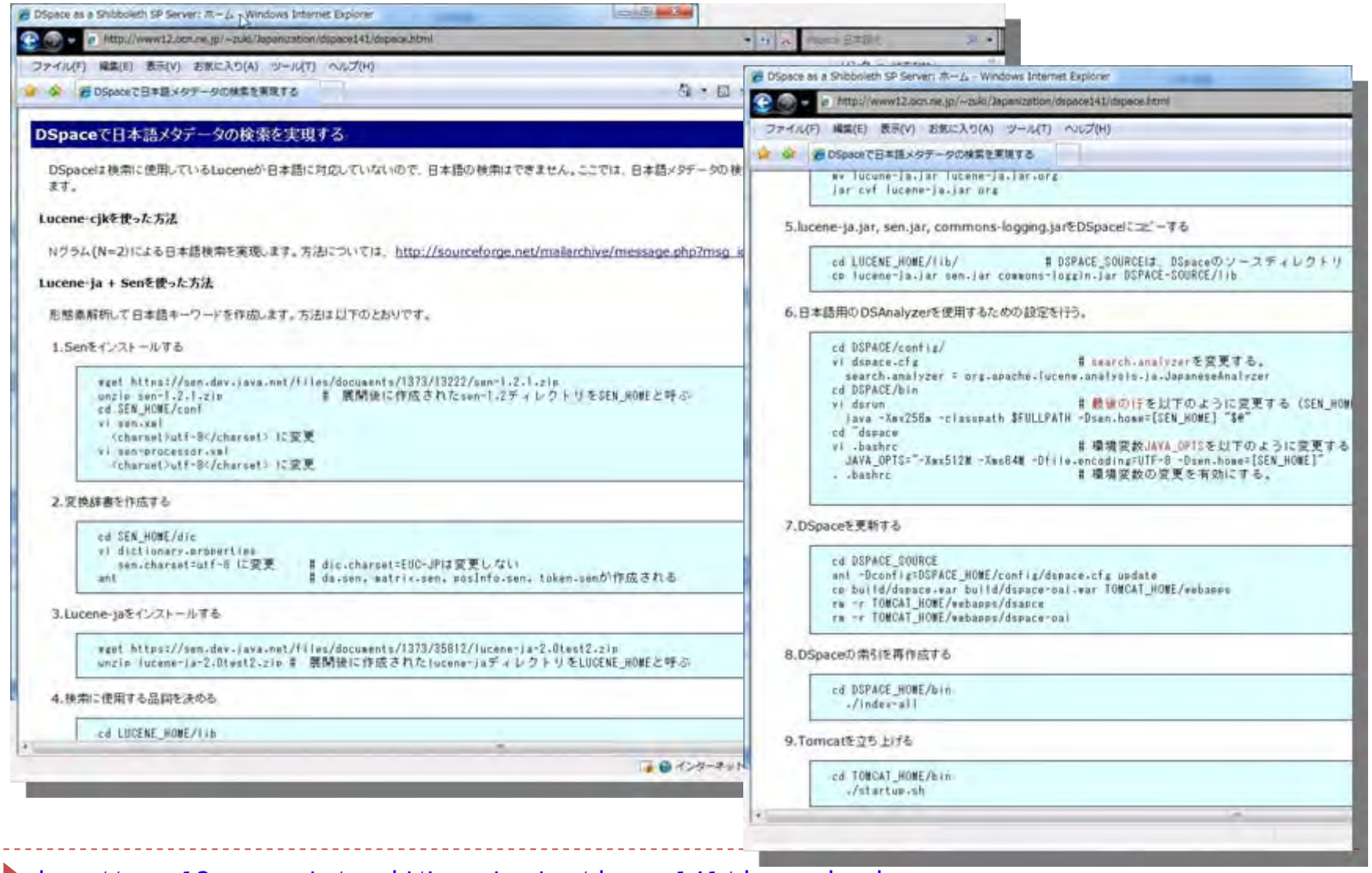
- http://shib2.at.nii.ac.jp/dspace/

リポジトリシステム
DSpaceメニューの日本語化

- http://www12.ocn.ne.jp/~zuki/Japanization/current/dspace-others.html

リポジトリシステム
DSpaceメタデータ検索の日本語化

リポジトリシステム
DSpaceメタデータ検索の日本語化

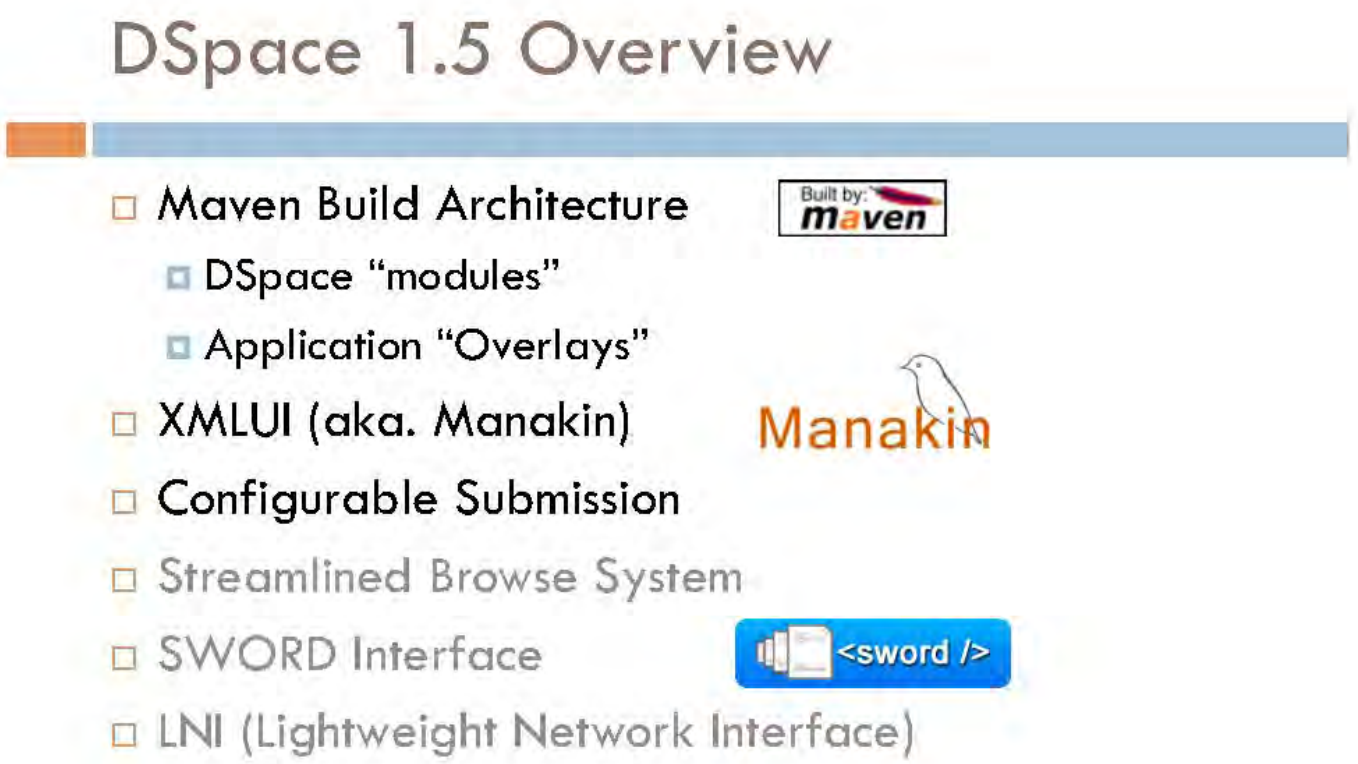


- http://www12.ocn.ne.jp/~zuki/Japanization/dspace141/dspace.html

リポジトリシステム
DSpace お茶の水女子大学TeaPotサイト

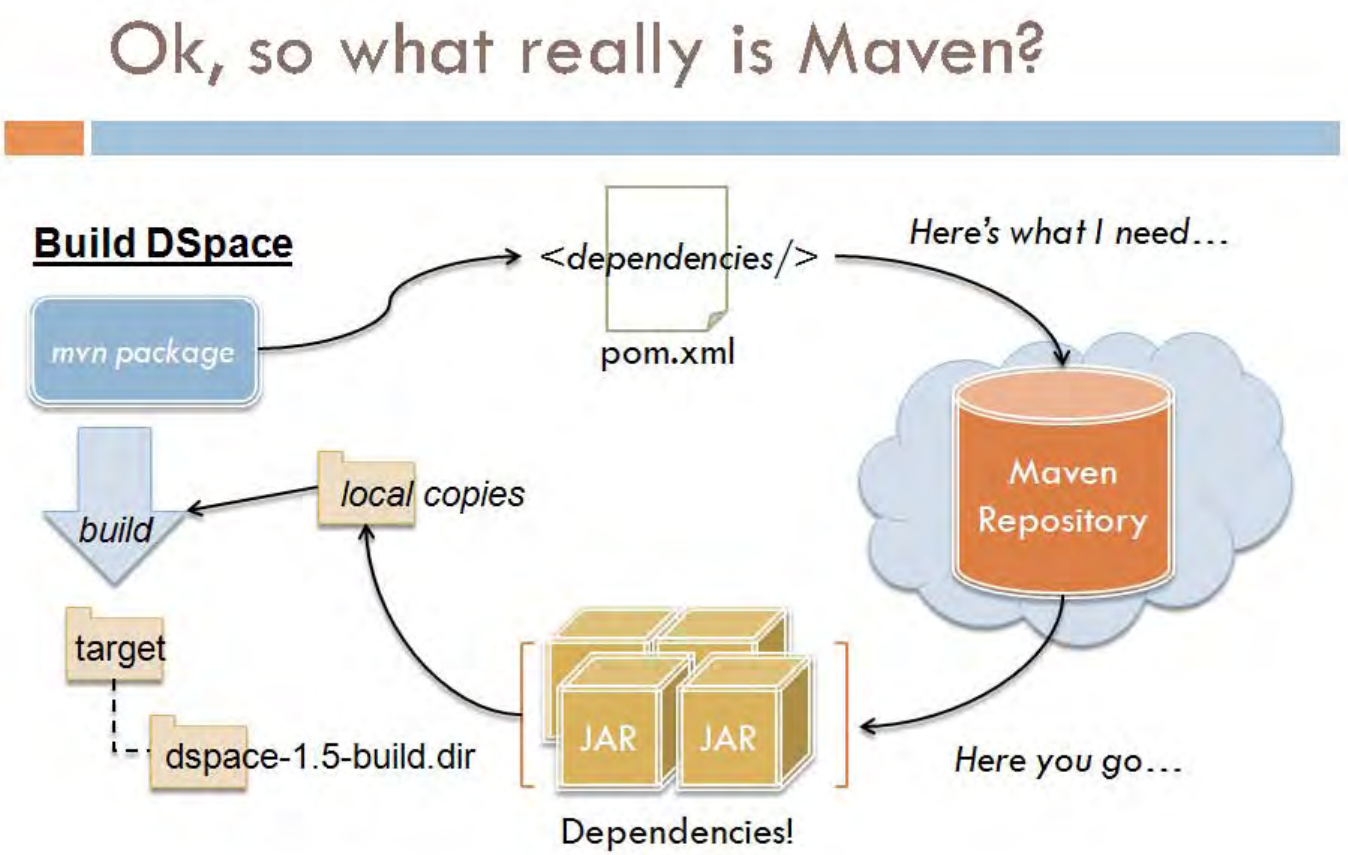
- http://teapot.lib.ocha.ac.jp/ocha/

リポジトリシステム
DSpace 1.5



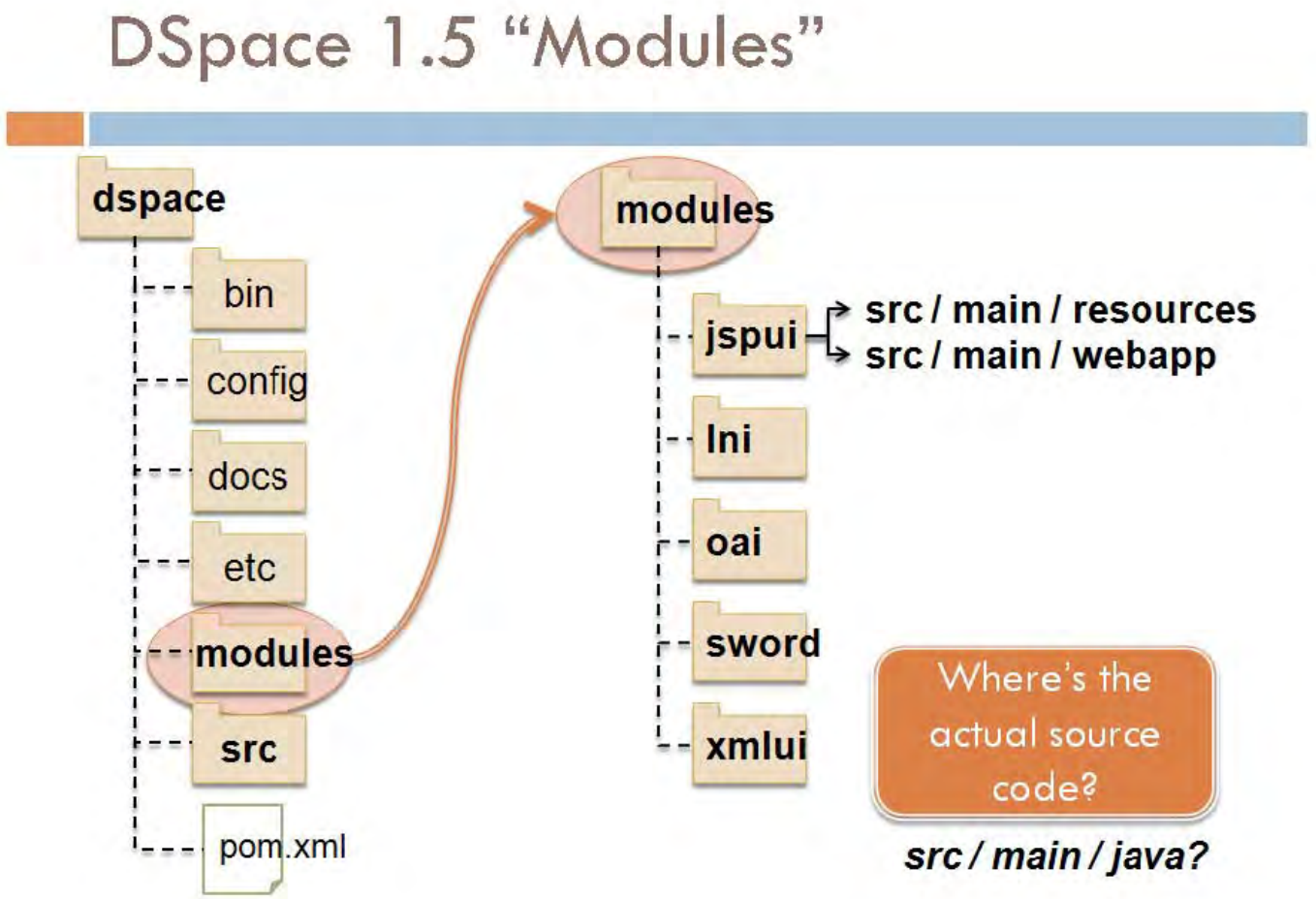
Donohue, T. and Triggs, G. (2008) *Customizing DSpace 1.5 with Basic Overlays.* In: Third International Conference on Open Repositories 2008, 1-4 April 2008, Southampton, United Kingdom.

リポジトリシステム
DSpace 1.5 Maven



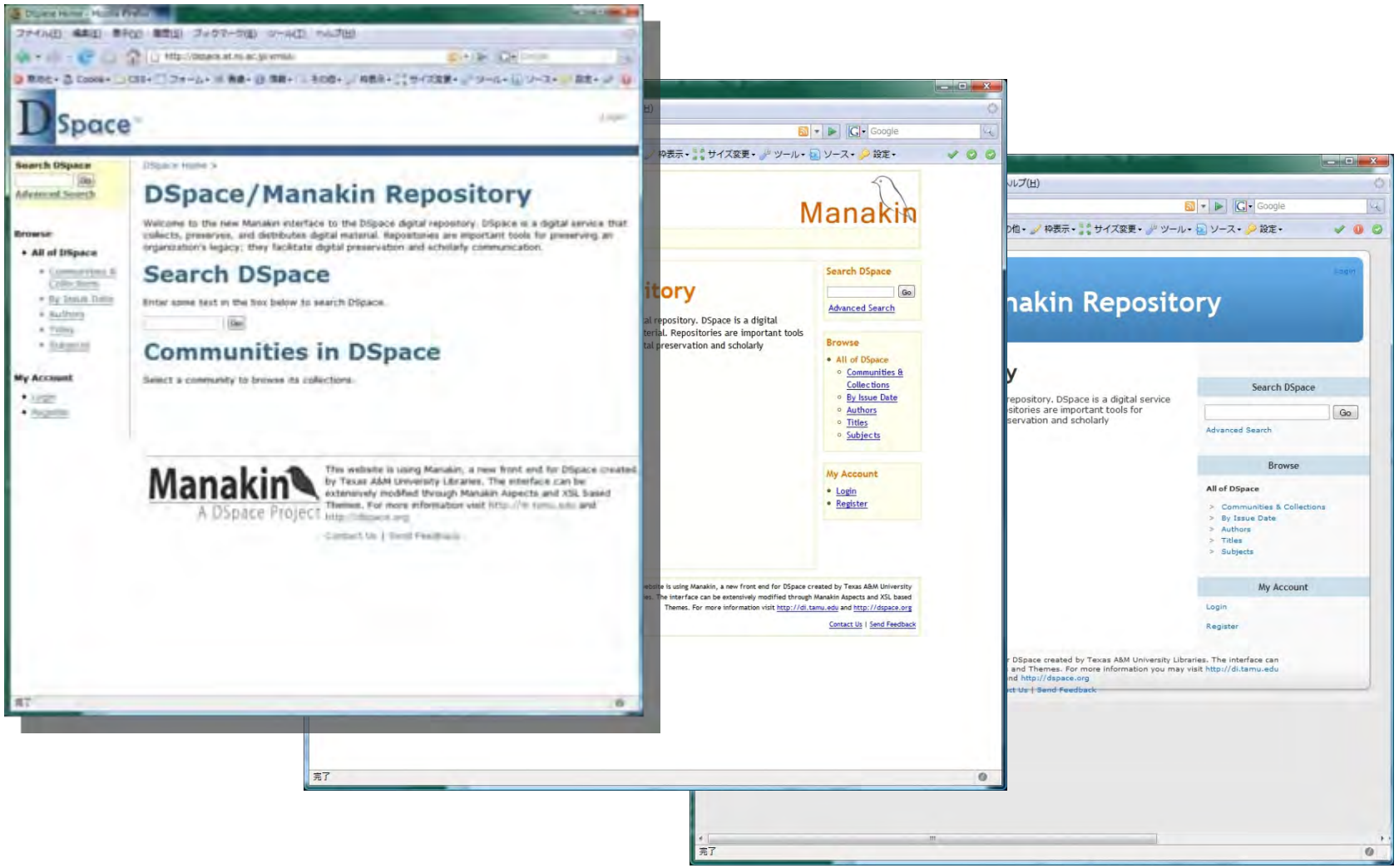
Donohue, T. and Triggs, G. (2008) *Customizing DSpace 1.5 with Basic Overlays.* In: Third International Conference on Open Repositories 2008, 1-4 April 2008, Southampton, United Kingdom.

リポジトリシステム
DSpace 1.5 Maven

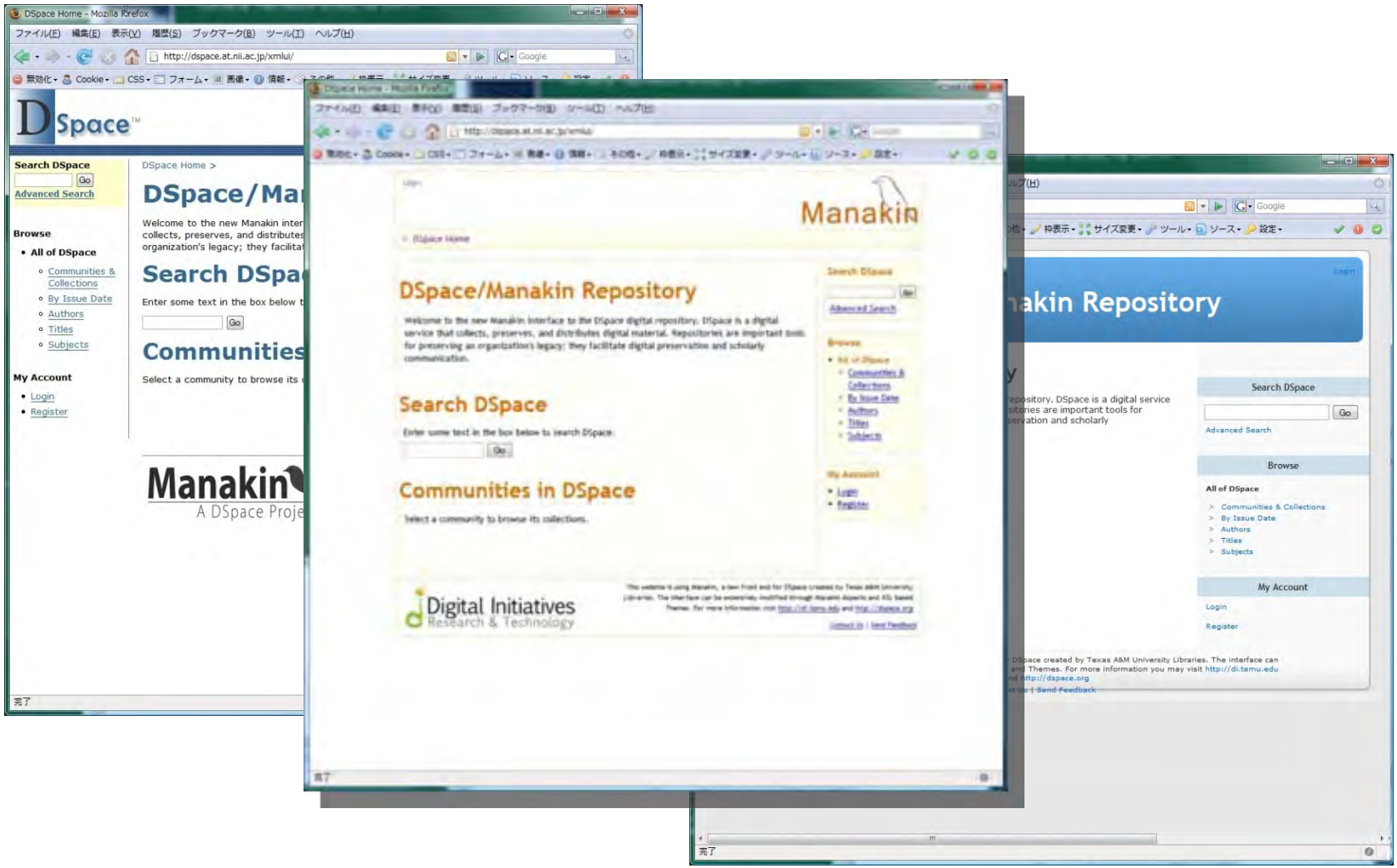


Donohue, T. and Triggs, G. (2008) *Customizing DSpace 1.5 with Basic Overlays.* In: Third International Conference on Open Repositories 2008, 1–4 April 2008, Southampton, United Kingdom.

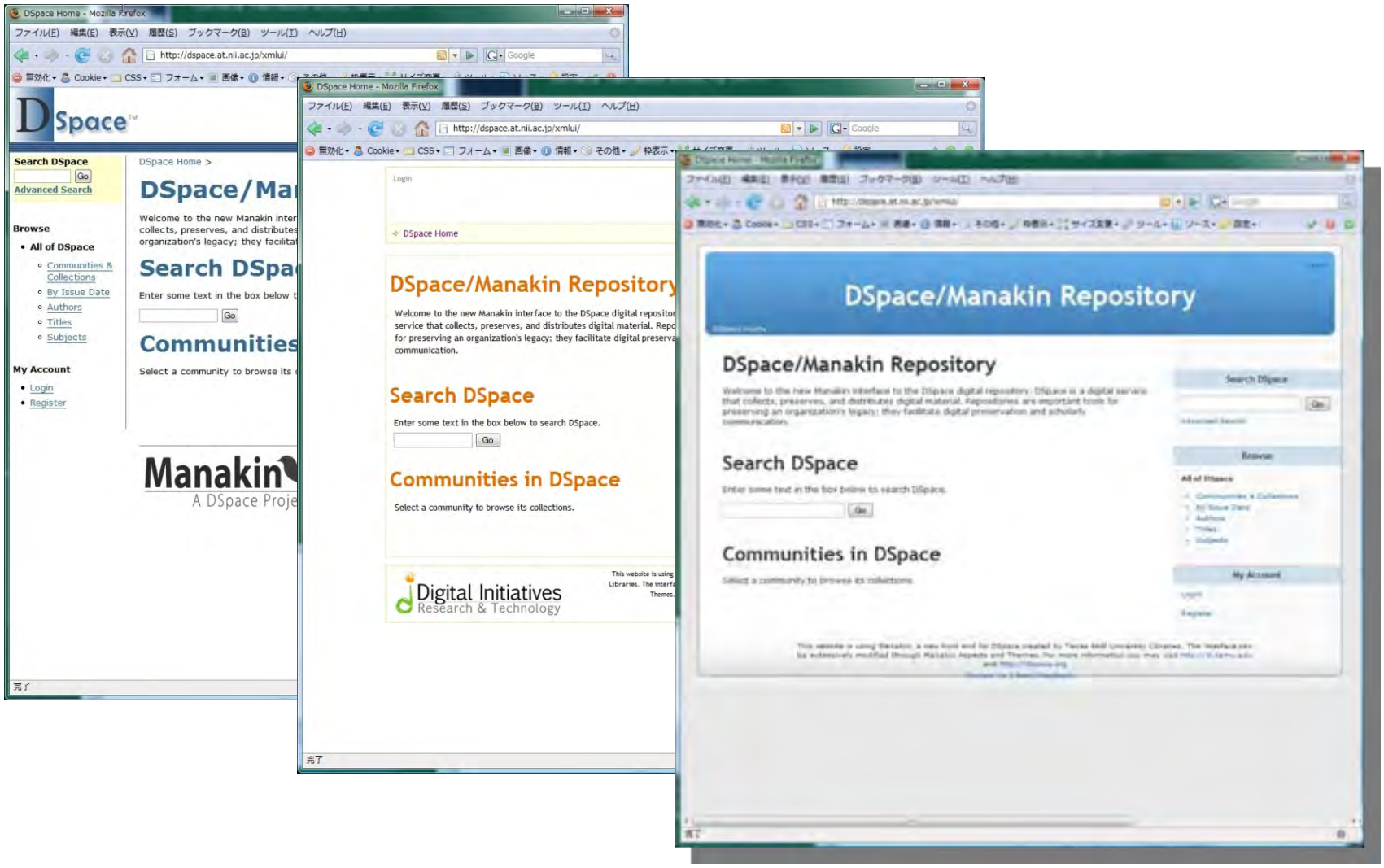
リポジトリシステム
DSpace 1.5 Manakin



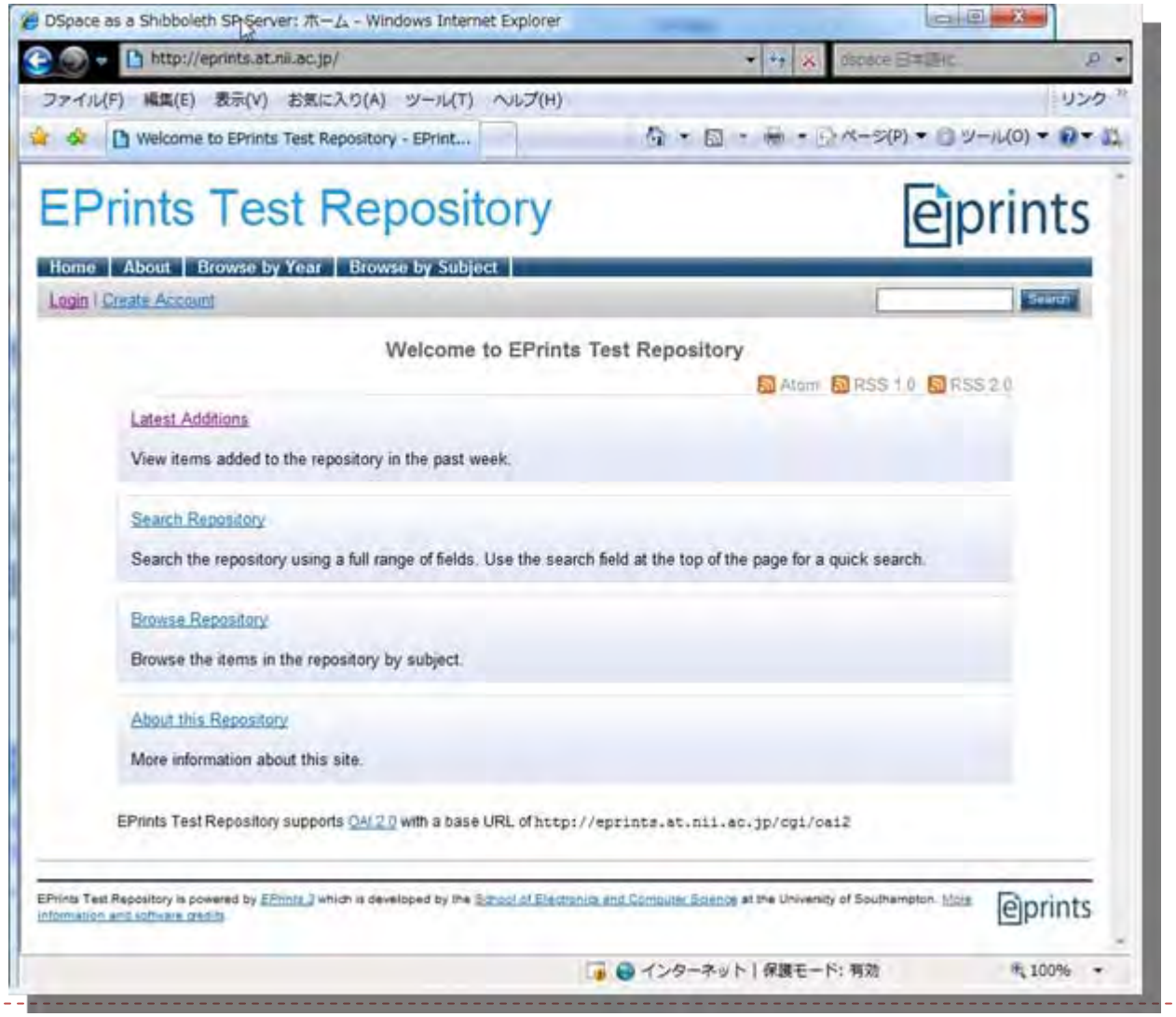
リポジトリシステム
DSpace 1.5 Manakin



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

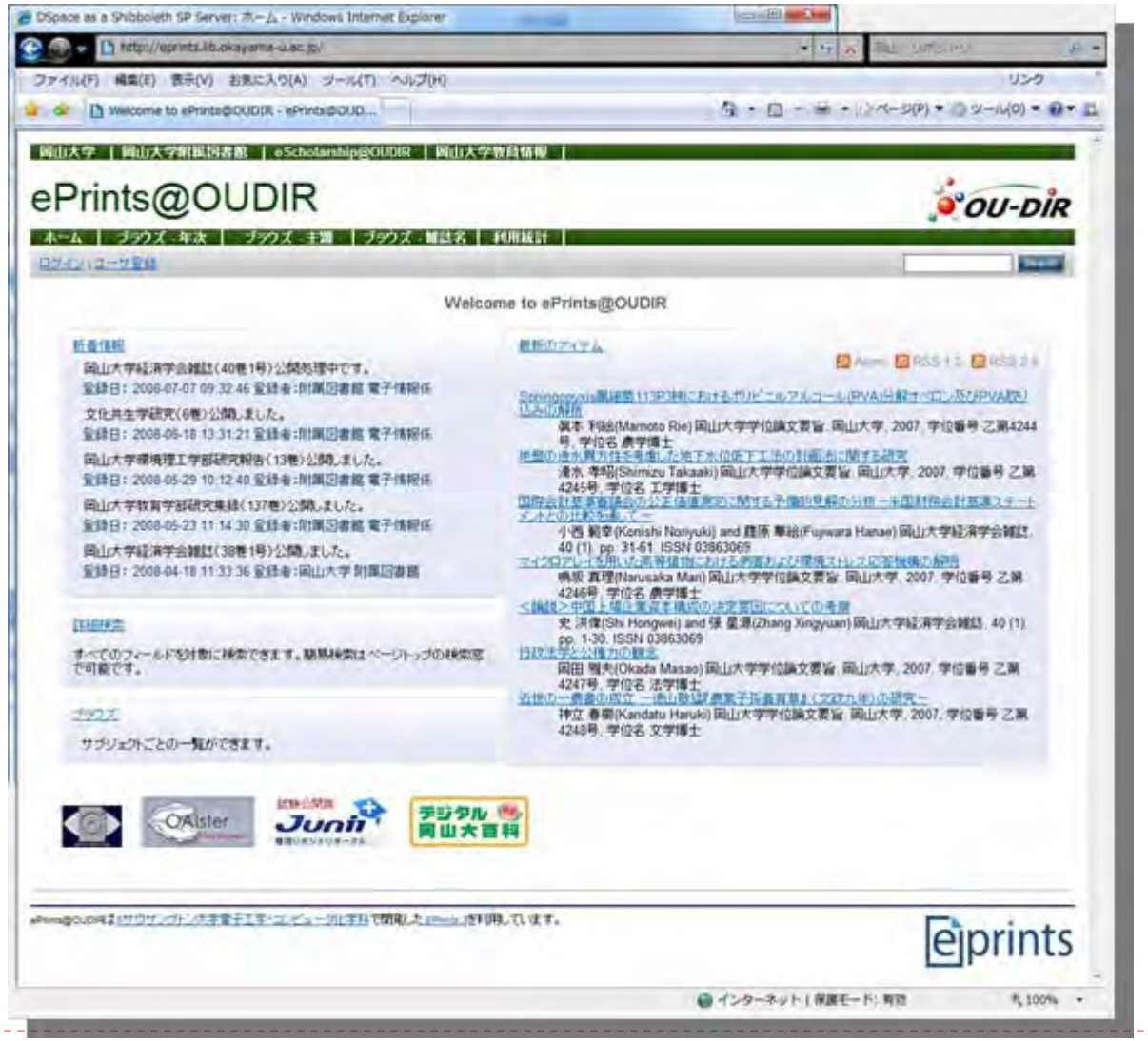


- http://eprints.at.nii.ac.jp/

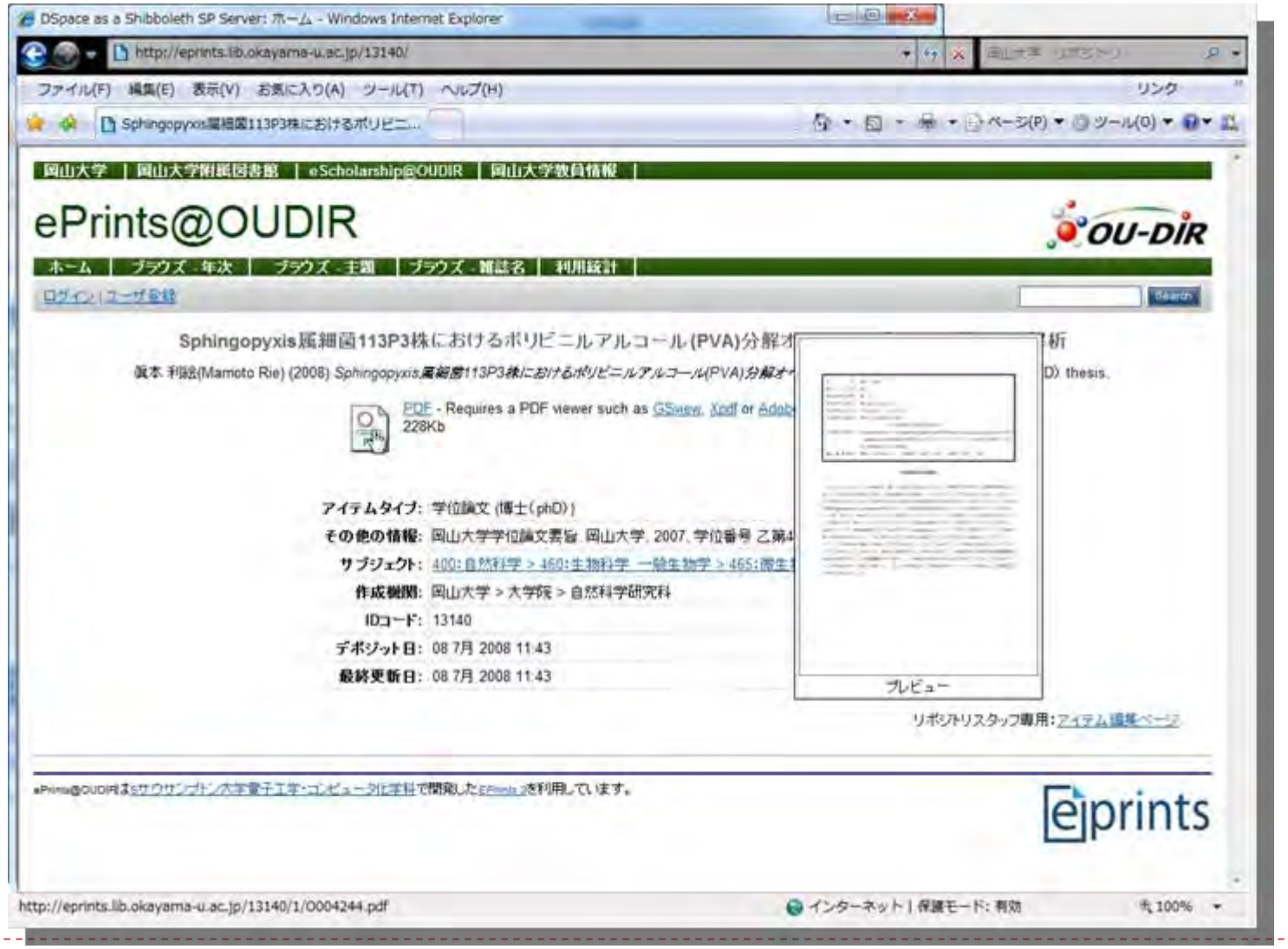
リポジトリシステム
EPrints　OR08サイト

- 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



リポジトリシステム
EPrints　岡山大学OU-DIRサイト



- http://eprints.lib.okayama-u.ac.jp/13140/

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

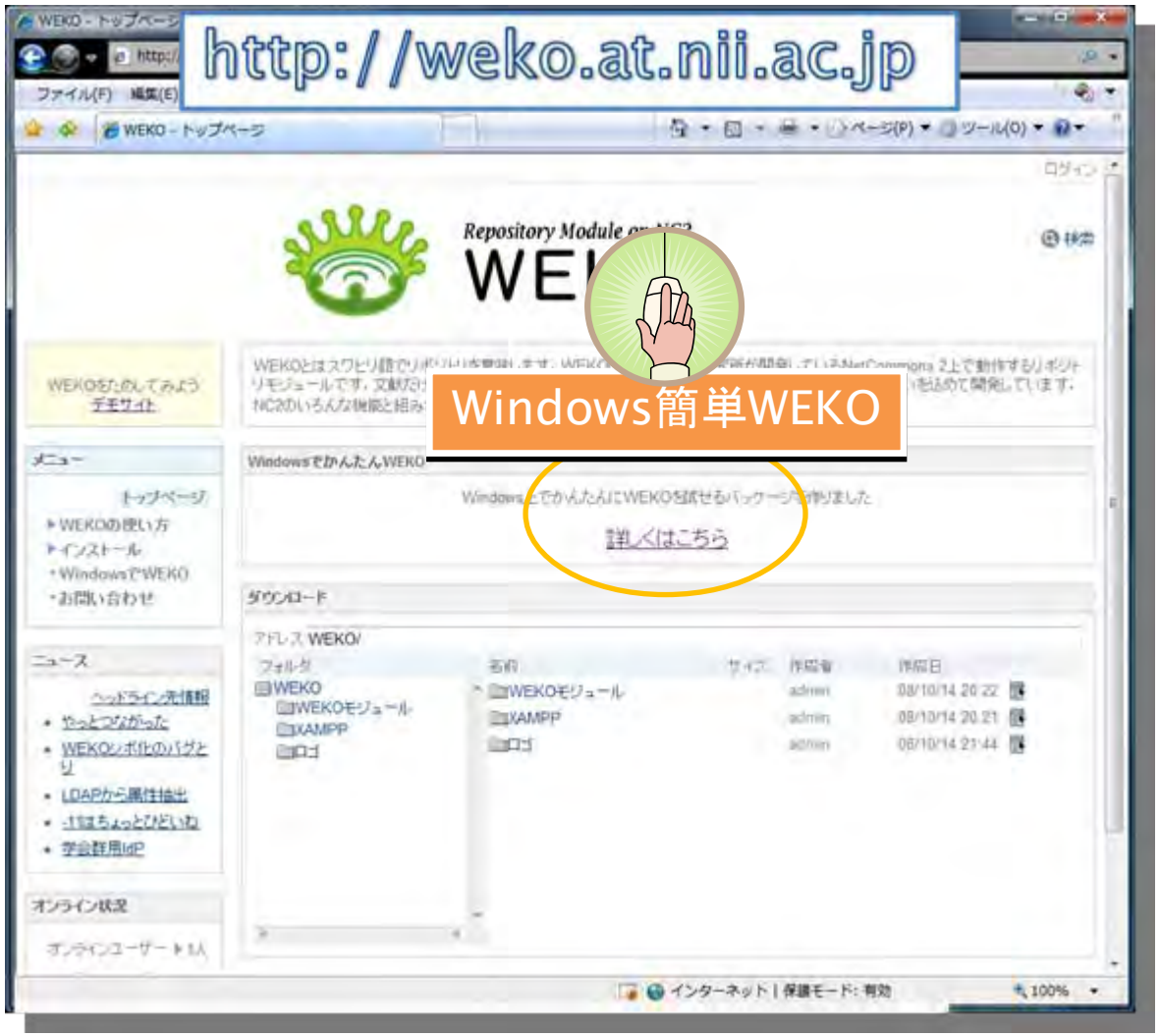
リポジトリシステム WEKO 機能

- 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

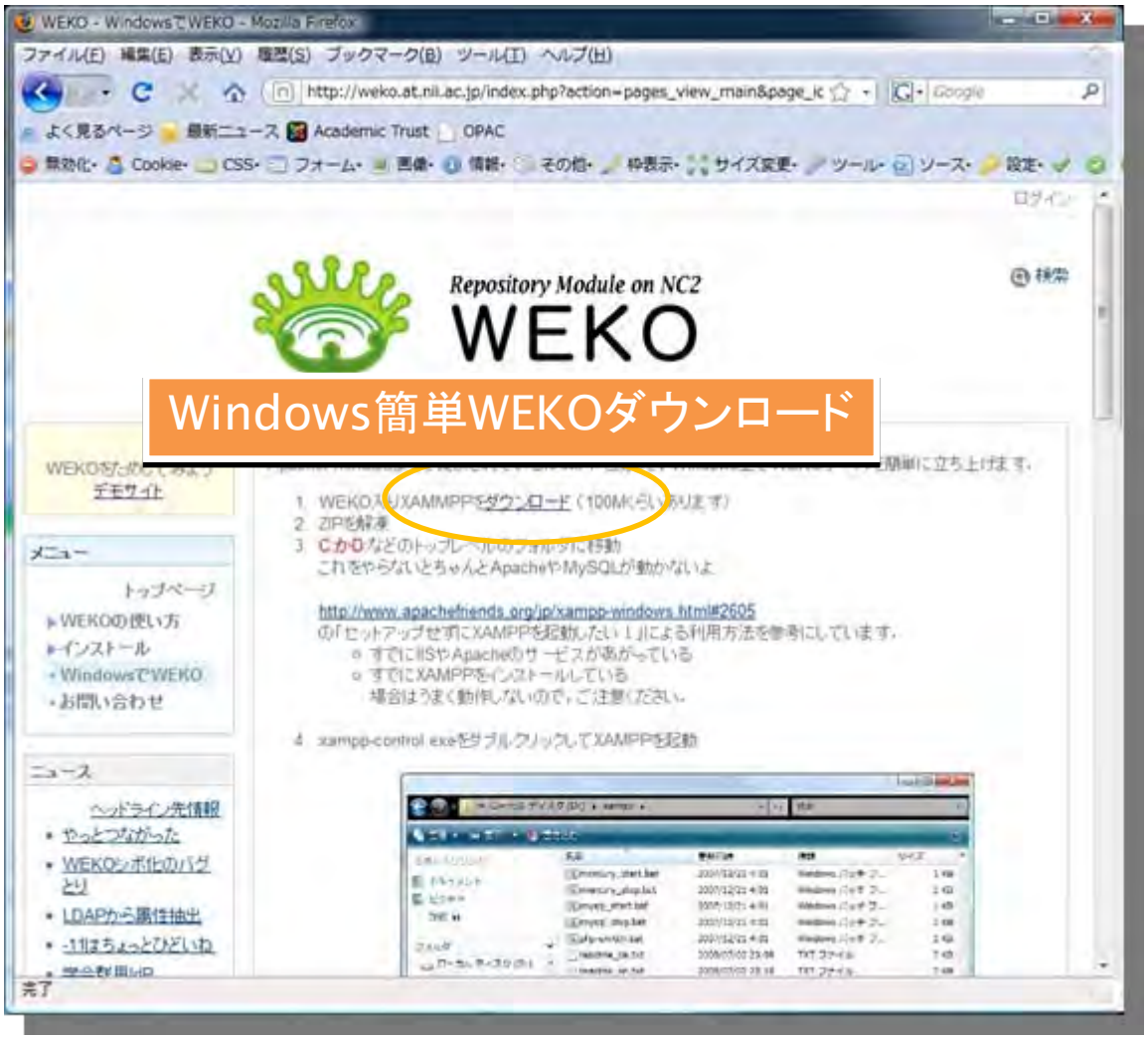
リポジトリシステム
WEKOサイト

リポジトリシステム
WEKOデモサイト

リポジトリシステム
WEKOサイト



リポジトリシステム
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

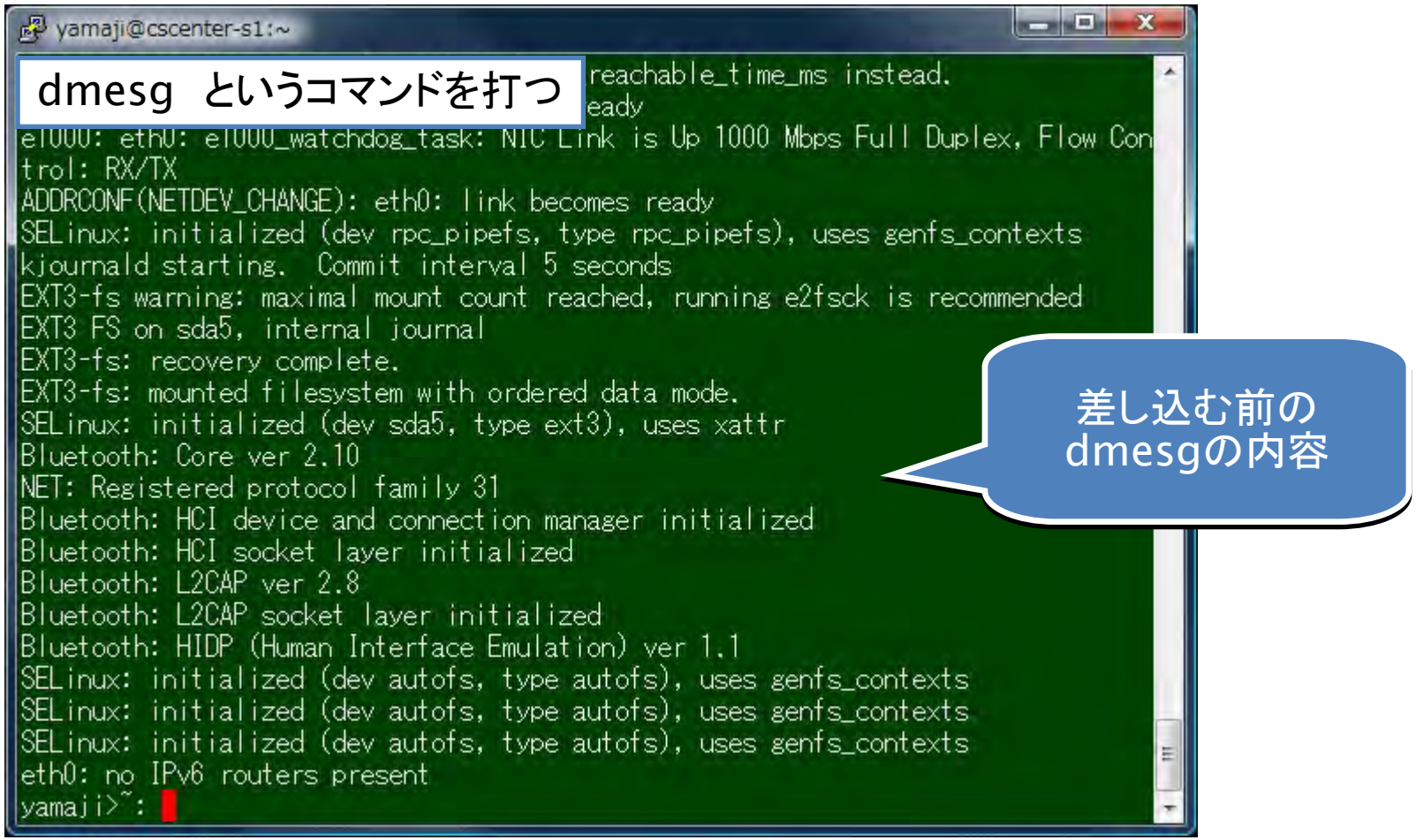
システム構築
OS，リポジトリシステムのインストール

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

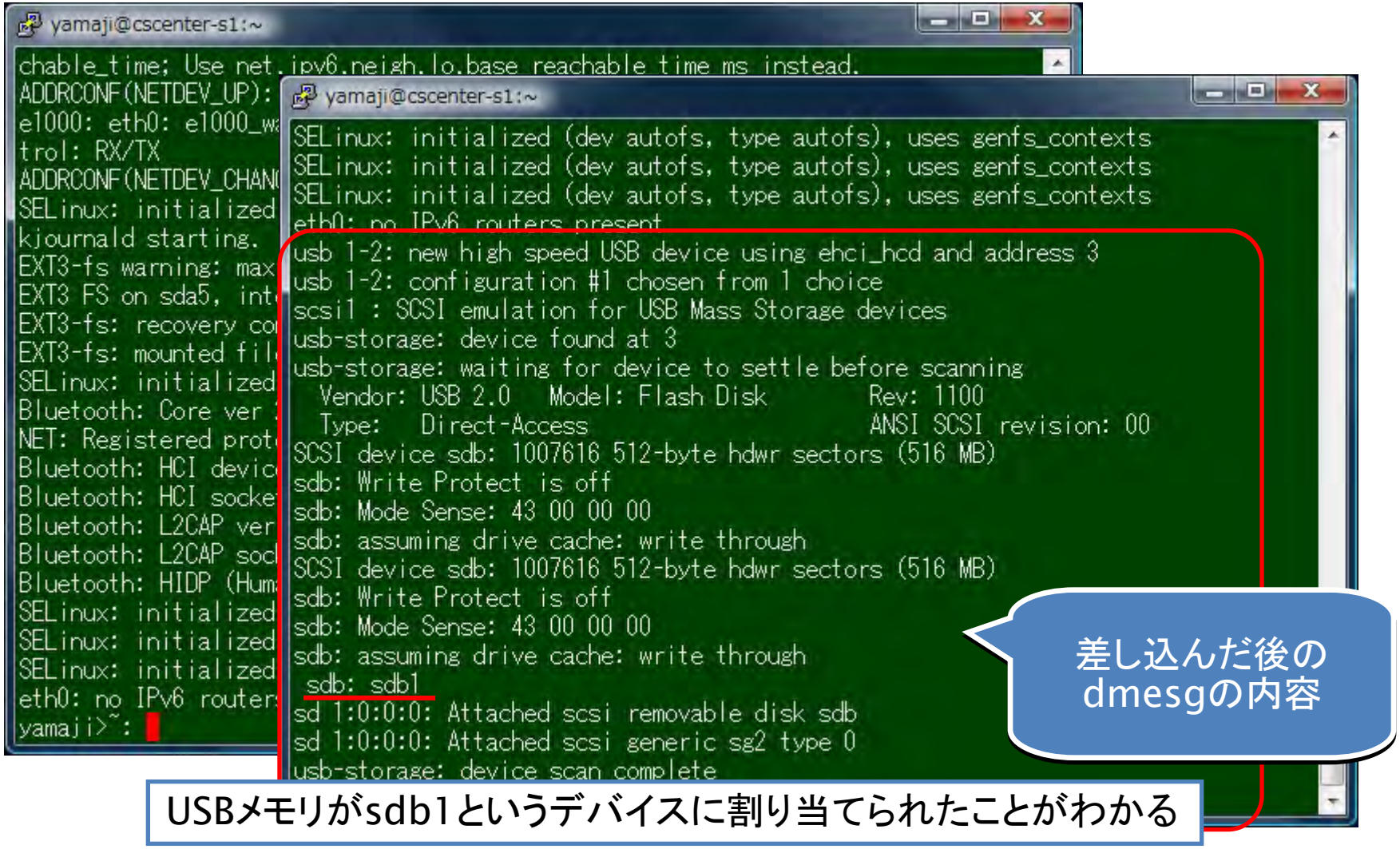
システム構築
バックアップ

- 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



システム構築
デバイス名の確認



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の実行結果がメイルで通知
    - SendmailでSMART_HOSTを設定
    - Aliasでrootへのメイ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

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

OAI-PMH
登録データの流通

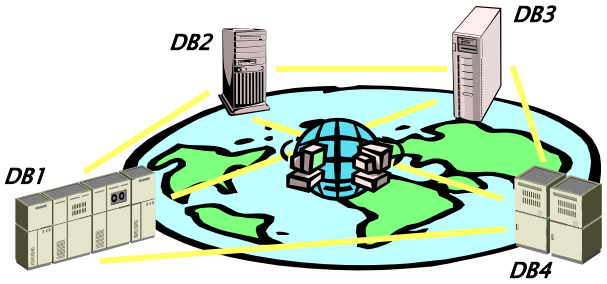
登録したデータをどのようにして広めていくか？

- 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

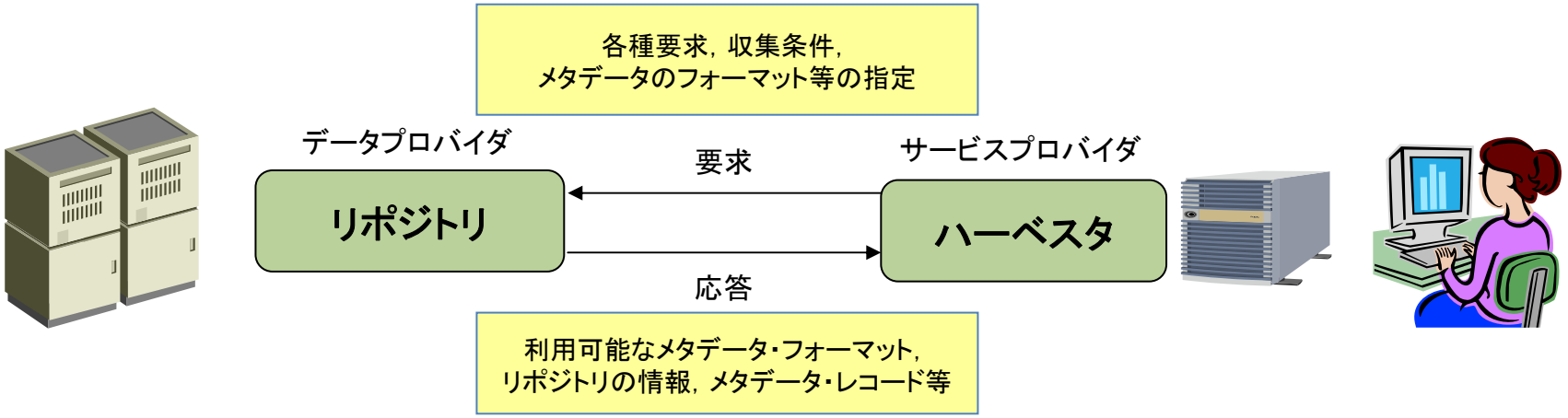
OAI-PMH
(Open Archive Initiative – Protocol for Metadata Harvesting)

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

http://mitizane.ll.chiba-u.jp/curator/about/chiba_oai_NII.ppt

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、黄色＝あれば必須　左記以外＝各機関リポジトリのポリシーによる

通番	大項目	項目	element	内容	必須	繰返し可否	minOccurs	maxOccurs	書式(空白の場合は文字列)	oai_dcにダムダウンさせるときの行き先	オプションの属性
1	タイトル	タイトル	title		必須	×	1	1		title	lang
2		その他(別言語等)のタイトル	alternative	別言語のタイトルやタイトルの日主			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 メタデータ主題辞書集	NIIsubject	下記URL参照 http://www.nii.ac.jp/metadata/manual/NII-category.pdf			0	unbounded		subject	version
6		日本十進分類法	NDC				0	unbounded		subject	version
7		国立国会図書館分類表	NDLC				0	unbounded		subject	version
8		日本件名標目	BSH				0	unbounded		subject	version
9		国立国会図書館件名標目表	NDLSH				0	unbounded		subject	version
10		医学件名標目表	MeSH				0	unbounded		subject	version
11		デューイ十進分類法	DDC				0	unbounded		subject	version
12		米国議会図書館分類表	LCC				0	unbounded		subject	version
13		国際十進分類法	UDC				0	unbounded		subject	version
14		米国議会図書館件名標目表	LCSH				0	unbounded		subject	version
15		内容記述	description	論文の抄録 学位論文にあっては、授与年度、論文/課程の種別等 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、組織、団体			0	unbounded		publisher	lang
17		寄与者	contributor	コンテンツ本体の作成に関わりを持つ、個人、組織、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 メタデータ主題辞書集(資源タイプ)	NIItype		必須	×	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		フォーマット	format	Internet Media Typeで規定されるファイル形式			0	unbounded		format	
22	資源識別子	その他の資源識別子	identifier	23項、24項以外の資源識別子			0	unbounded		identifier	
23		資源識別子URI (アイテム表示画面へのリンク)	URI	コンテンツ本体を表示する画面へのURL	必須	×	1	1	URI	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	URI	identifier	
25	掲載誌情報	ISSN	issn		あれば必須		0	unbounded	¥d{4}-¥d{3}[¥dXx]	折りたたみ編集の上、identifier	
26		書誌レコードID(総合目録DB)	NCID		あれば必須		0	unbounded			
27		雑誌名	jtitle	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	spage	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	epage	論文の終了ページ	あれば必須		0	1	半角数字		
32		刊行年月	dateofissued		あれば必須		0	1	YYYY-MM-DD, YYYY-MM, YYYY		
33		情報源	source				0	unbounded		source	

- http://www.nii.ac.jp/irp/archive/system/pdf/junii2_elements_guide_ver2.pdf

OAI-PMH
junii2フォーマット②

通番	大項目	項目	element	内容	必須	繰返し可否	minOccurs	maxOccurs	書式（空白の場合は文字列）	oai_dcにダムダウンさせるときの行き先	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	PubMed番号	pmid				0	1	URI（info:pmid/"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	版に関する情報	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生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		事件とされる	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		事件とする	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をなす			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際のコンテンツから関連づけられる			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関連付けている			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		範囲	coverage				0	unbounded		coverage	
51	地理的範囲	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）	NIIspatial				0	unbounded		coverage	
53	時系列範囲	時間的	temporal				0	unbounded		coverage	
54		国立情報学研究所 メタデータ主題語彙集（時代）	NIItemporal				0	unbounded		coverage	
55		権利	rights	コンテンツ本体に関する権利規定			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:著者版 publisher:出版社版 none:上記以外			0	1	[author | publisher | none]		

		オプション属性	書式
		lang	RFC1766
		version	実数

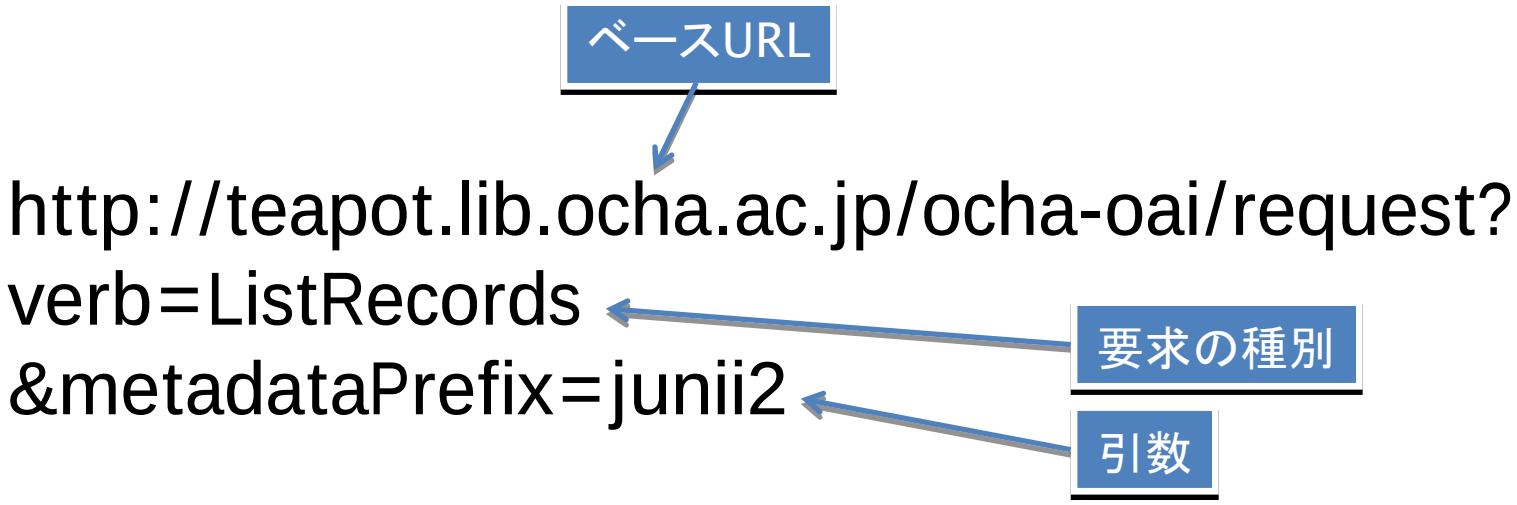
- http://www.nii.ac.jp/irp/archive/system/pdf/junii2_elements_guide_ver2.pdf

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、学位の種別等は必要に応じて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



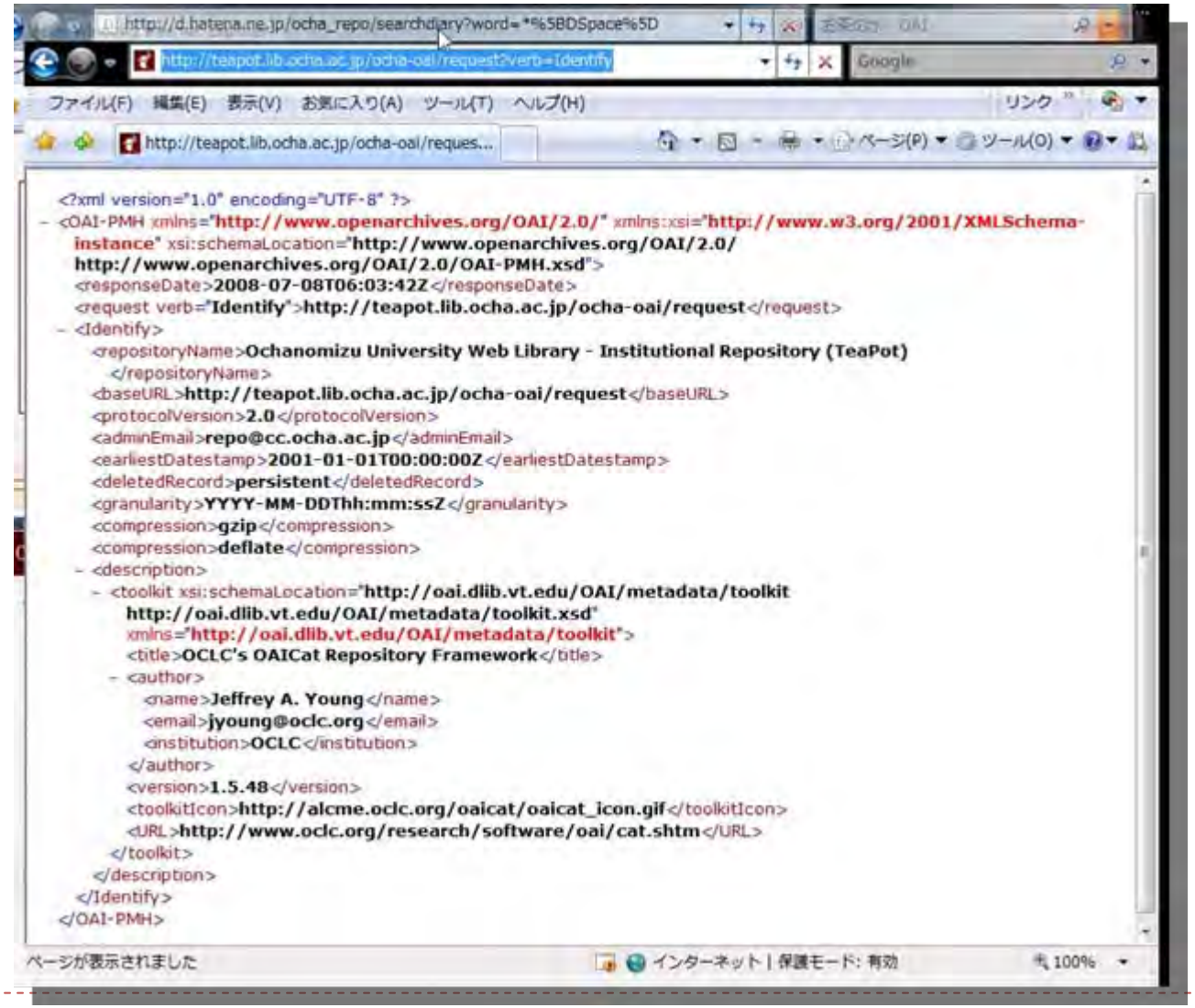
リポジトリ（http://teapot.lib.ocha.ac.jp/ocha-oai/request）から
junii2メタデータフォーマットで公開されているメタデータを
全て取得せよ

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

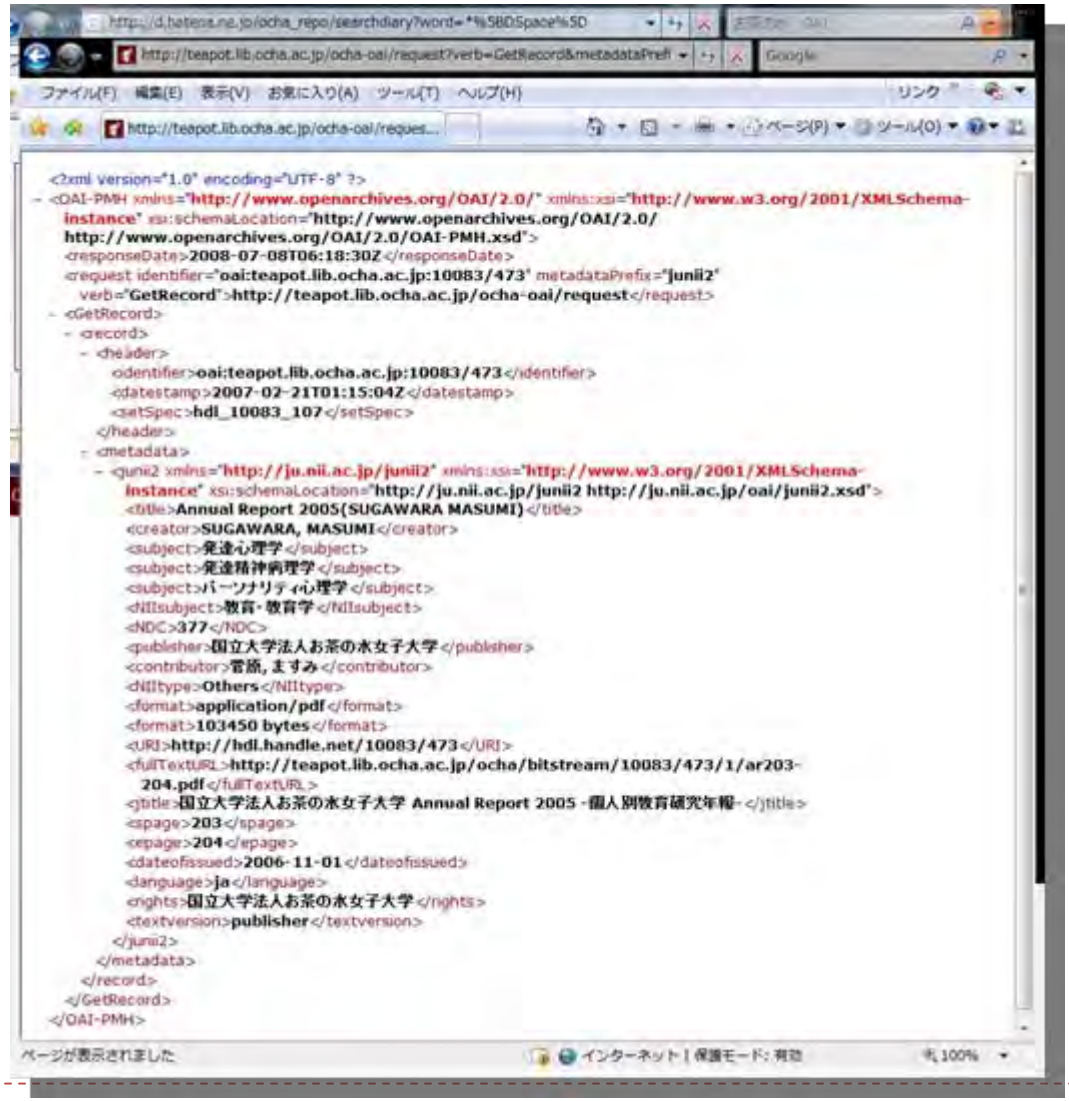


- http://teapot.lib.ocha.ac.jp/ocha-oai/request?verb=Identify

OAI-PMH ListMetadataFormats

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

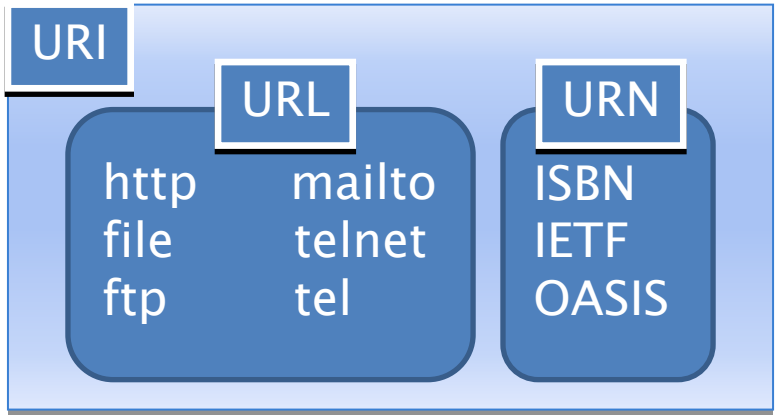
OAI-PMH GetRecord



- 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/

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：
  - システム構築にかけ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**

お気軽にメイルください

yamaji@nii.ac.jp

NII