# NIMS eSciDoc: eScience向け基盤ソフトウェア開発と実践の現場から

2009年10月30日 高久雅生 物質•材料研究機構 科学情報室 TAKAKU.Masao@nims.go.jp

### 本日のお品書き

- (背景)
- ソフトウェア共同開発
  - 何に取り組んでいるのか?
  - どう取り組んでいるのか?
- 実践
  - どんな実践を行っているのか?
  - 何が課題となっているか?

### 自己紹介

#### • 高久雅生

- 所属: 独立行政法人 物質•材料研究機構

- 部署: 科学情報室

- 職位: 主任エンジニア

#### • 経歴:

- 1994年 図書館情報大学 図書館情報学部 入学
- 2004年 筑波大学図書館情報メディア研究科博士後期 課程 修了

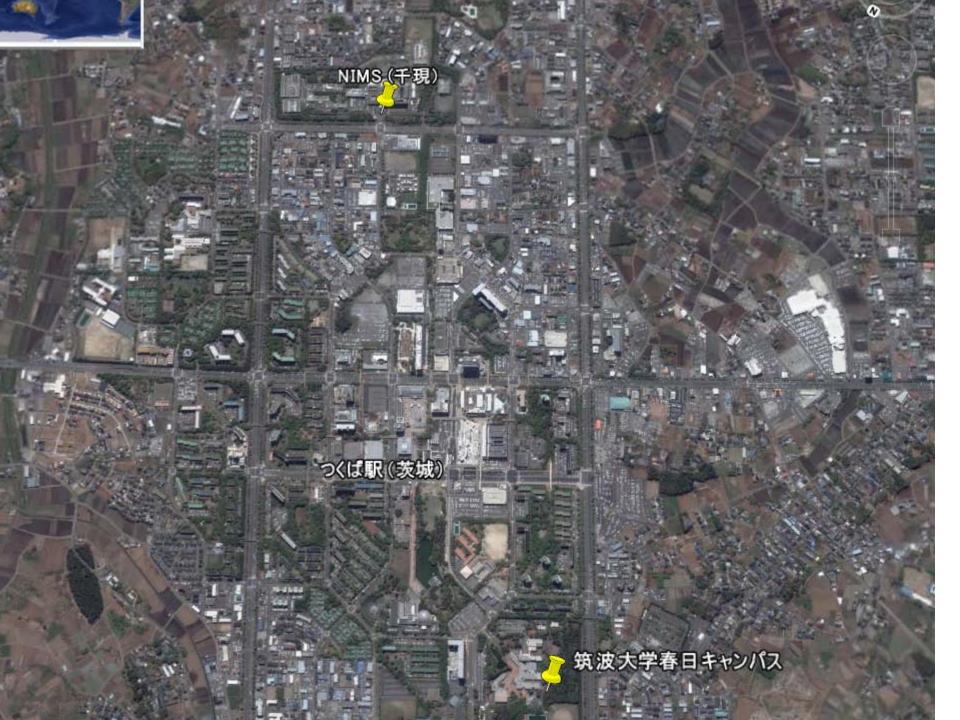
10年近く、このキャンパスに生息していました

### 自己紹介 (2)

- 属性?
  - Unix/FLOSS, TeXnician, Rubyist, Wikipedian
- オープンソースソフトウェア開発者
  - 貢献したことのあるプロジェクト:
  - Namazu, Kakasi, Ruby, QwikWeb, MoBo, etc.

### いま、何をやっているのか?

- 主任エンジニア → Library Engineer
- 図書室担当
  - おもに3拠点(千現、並木、桜)
    - 総蔵書数16,000冊程度
  - 利用者対応
- 電子リソース担当
  - 機関リポジトリ
  - 研究者総覧
  - データベース/オンラインジャーナル周り
  - リンクリゾルバー(SFX)
  - OPACシステム
  - 図書室端末PC(千現2台,並木1台)



### NIMS図書室



2009/10/30

### で、本題?

- 機関リポジトリ: NIMS eSciDoc
- eSciDoc = "eScience + Documentation"
  - ドイツ政府の公的助成に始まるプロジェクト
  - 開発チーム
    - Max-Planck Digital Library (MPDL)
    - FIZカールスルーエ (FIZ Karlsruhe)
  - eScience用基盤ソフトウェア
  - オープンソースソフトウェア
- NIMS MPDL 間で共同研究開発MOU

### マックスプランク協会 (Max-Planck Society)

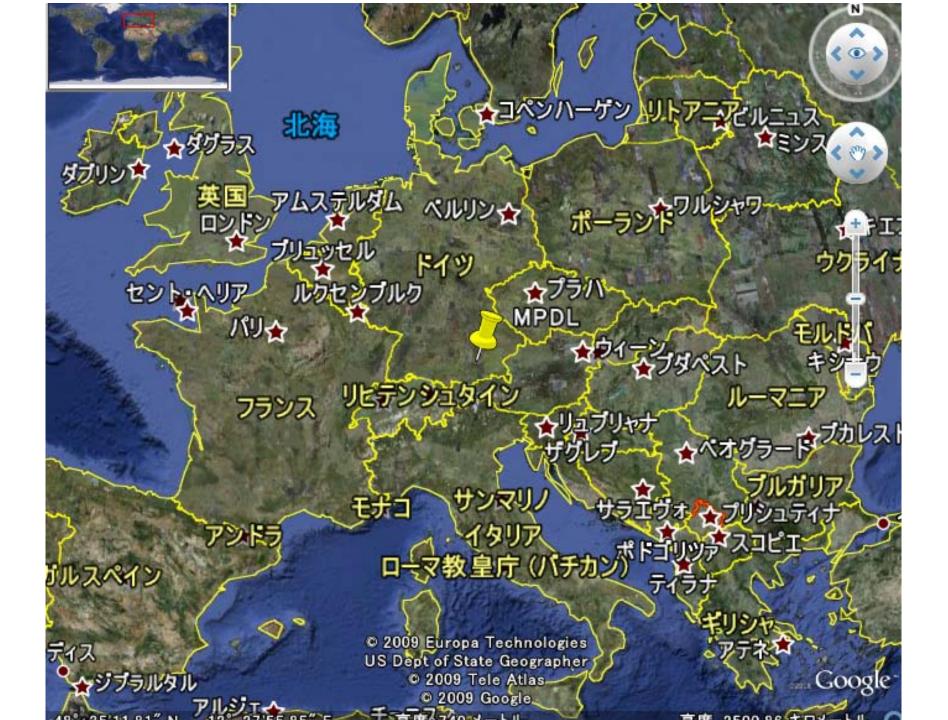
- ドイツを代表する国際的な大規模研究機関
- 80もの支部研究所により構成
  - 人文学、医学・生物学、自然科学など全研究分野を横断
  - ドイツ全土および国外に支部が散在

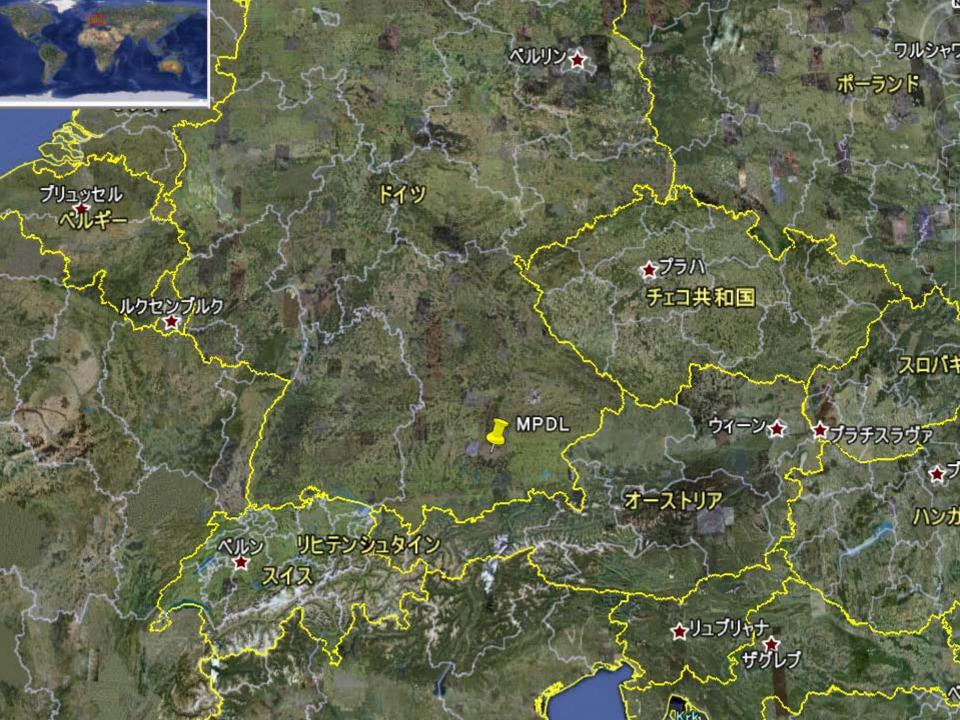
### MPDL (Max Planck Digital Library)

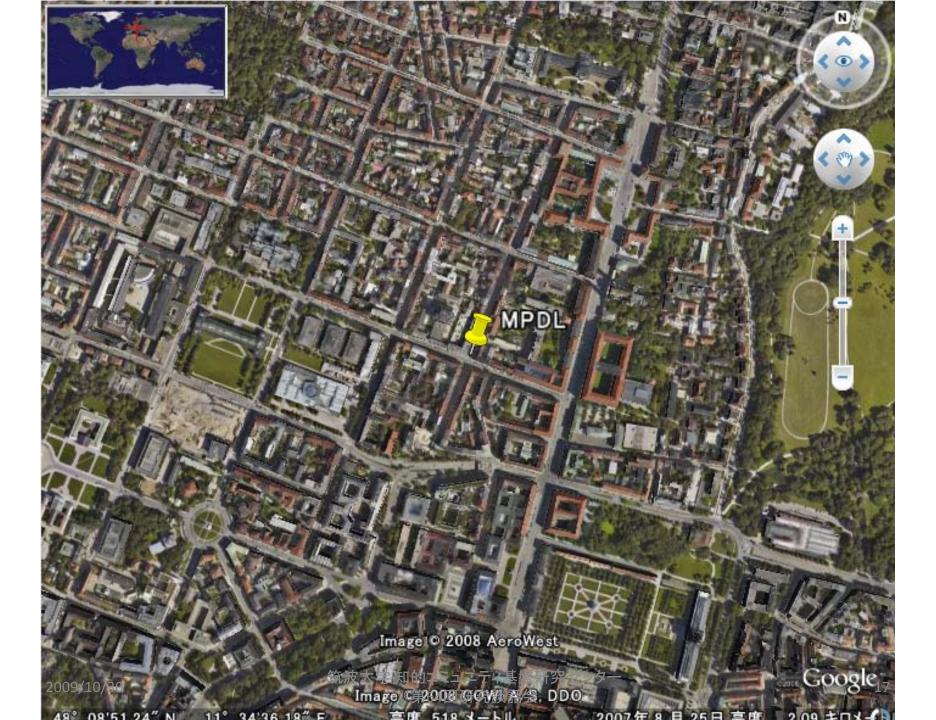
- URL: http://mpdl.mpg.de/
- マックスプランク協会の80研究所の電子図書館サービスを統括
  - 電子リソース管理
  - ポータルサイト運営
  - eScienceソフトウェアの研究開発
- eSciDocの開発、運用
  - 総勢40人前後(うち、開発チームが20人程度)
  - FIZ(ドイツ版JST)との共同開発
    - ・遠隔地: Karlsruhe-Munichは300km(東京~名古屋くらい)
    - 週に一回、ビデオカンファレンスで打ち合わせ

### 共同開発

- 相互訪問
  - MPDL 訪日: 3回
    - 2008年7月, 2008年11月, 2009年3月, 2009年9月
  - NIMS訪独: 3回
    - 2008年9月(2週間)
    - 2009年1月(1週間)
    - 2009年6月(1週間)







### MPDLエントランス



### MPDLミーティングルーム



### MPDL風景



### MPDL風景





### 共同開発の実態

- eSciDocプロジェクト
  - FIZ: コアサービス
    - 認証、データストア、検索まわりなど
  - MPDL: アプリケーション
    - 仕様策定、実装、GUI設計、テスト
- NIMS: 開発&運用
  - 日本で必要な機能の追加実装 = 日本語化(ローカライズ)
  - NIMSで必要な新機能提案 = 仕様設計

### Personas for Publication Management

Denise Unfried, Rupert Kiefl, Markus Meichau

#### Introducing Librarian "Anne"



Persona Profile Average Age Gender Education Profession Employment Working Experience Working space External Help Noise and Distraction General Applications Work spent with applications 90-100 % Using web applications

Female Librarian Max Planck Society

Average Browser

Office, Aleph. Library catalogues 80-90 % Workload Average

#### Persona Social Background

Arme is a 34 year old librarian. She was born in Germany and works in a small library at a Max Planck (national in

She is a daughter of a German father and Dotch because her mother is Dutch. Sine has a university degree in library

Her day in the library is sometimes guite busy. She does almost all library work with her computer using libero, alaphand other library catalogues. Bornetimes she is interrupted by visitors or groups of visitors like students asking for guidance During the rest 8 years working in the library, her work with web applications grew and today she works more than had at her time with web applications. In case she has a problem with an application she can ear consigues for help

#### Introducing Scientist "Simona"



#### Persona Profile

Average Age Gender Education Profession Employment Working Experience Working space External Heip Noise and Distraction Srowser: General Applications

Work spent with applications Using web applications Workload

Female Higher:

Max Planck Society 14 years

80-90 %

#### Persona Social Background

Simons is a 39 year old scientist. She was born in Hungary and works at a Max Planck institute in Berlin.

She studied in Budaped and did her post dictor in the US. She holds a doctor's degree in excreenings

Her day is the institute is quiet but very busy. She works on a study about DNA sequence data from organistist and nuclear DNA sequences to form germ phylogenes.

She reads manifold lecture spanning chamiley, strescel themstry, microbiology and anzymology. From time to line she clas one of the publications for the poper on gare. shyloganies.

EMACS. Endivide. Postsoriel.

Latest. Office other domain.

To wait for books or publications of external stravies. Today the true access to what has been published in the field with When she started her researcher career she sometimes had she has access to what has been published in the field within seconds via web repositories.

She needs only help when a hardware problem occurs. With cofficers applications she can figure out things guille east.

#### Introducing Secretary "Barbara"



#### Persona Profile

SUMMADO ADD Gunder Female Profession king Experience Working space External have Stowen.

Education High Scrool Graduate Max Princip Stockly 35-40 Years

#### Persona Social Background

Bartiers is a 55 year old secretary. She was form in Germany

years. After High School she started as a secretary in Muand recent to the Mar Planck methods 2 Years ago She can profit from her broad experience in organizing at

ペルソナ手法に基づく 開発アプローチ

### MPDLにおける開発体制

- 仕様書、文書(一部は関係者限定):
  - CoLab (Wikiサイト): <a href="http://colab.mpdl.mpg.de/">http://colab.mpdl.mpg.de/</a>
- バグ報告 (開発者限定)
  - Apache JIRA (MPDL): <a href="http://zim01.gwdg.de:8080/">http://zim01.gwdg.de:8080/</a>
  - Bugzilla (FIZ): <a href="http://www.escidoc.de/issueManagement/">http://www.escidoc.de/issueManagement/</a>
- ソースコード管理
  - Subversionリポジトリ: https://zim02.gwdg.de/repos/pubman/
- コミュニケーション
  - 口頭,メール, ML, Yahoo! messenger
  - MPDL内の開発者ミーティング(毎週火曜日)
  - ビデオ会議(Polycom; MPDL ⇔ FIZ)(毎週)
  - ビデオ会議(Polycom; MPDL ⇔ NIMS)(不定期; 各月)

### CoLab (Wikiサイト)



Log in / create account

page

discussion

view source

history

#### about colab

- Purpose
- People
- Feedback

#### navigation

- Main Page
- Newest Pages
- Recent Changes
- All Pages
- All Categories

#### contents

- Projects
- Main Topics
- Events

#### help

- Wiki Help
- Wiki Syntax Basics

#### search



#### toolbox

- What links here
- Related changes
- Upload file
- **\$2009/10/30**
- Printable version

#### Main Page

#### Welcome to CoLab

Welcome to the MPDL CoLaboratory (CoLab), the platform for sharing information, knowledge and experiences in the context of eScience and information management. All interested users are invited to **learn more about CoLab**, to provide us with **feedback** and to participate (for participation see bottom of this page).

#### Quickstart

For a Quickstart to popular content enter CoLab via these portals:









#### CoLab categories

All content in this Wiki is organized in categories. The tag cloud displayed below links to important categories and gives you a quick overview of the topics covered by the CoLab Wiki. The bigger a category appen的themore on engine with the more possible in the important category, click on its name. 第64回研究談話会

#### CoLab News

2008-07-16: More

Improvements and Changes. CoLab has two new and nice functionalities that are worth to have a look at: The glossary extension and the RSS/Atom

feed for single

articles.

 2008-07-01: Improvements and Changes. With the update of CoLab to MediaWiki 1.12 there have been some improvements and changes. More

information is

available here

### JIRA(バグ追跡ツール)



User: Takaku Masao Filters | Profile | Log Out



HOME

**BROWSE PROJECT** 

FIND ISSUES

CREATE NEW ISSUE

**ADMINISTRATION** 

**OUICK SEARCH:** 

Manage Portal

[hide]

[hide]

#### MPDL JIRA

Category: eSciDoc Development
Common services (CSRV)
FACES Bug Tracking (FACESBUG)
PubMan Bug Tracking (PUBMAN)
PubMan Bug Tracking - Admin (ADM)
PUBMAN@NIMS (NIMS)
VIRR Bug Tracking (VIRRBUG)

Saved Filters (Create New | Manage Filters) My Issues (Unresolved) 108 NotClosedPubManBugs 13 OpenIssuesPubManBugs

Open Issues: Assigned To Me (Displaying 0 of 0) You have no assigned issues at the moment.

Category: eSciDoc Development Project: Common services (CSRV) Lead: Bulatovic, Natasa

Reports: Open Issues | Road Map | Change Log |

Popular Issues | Subversion Commits

Open Issues: (By Priority)

Filter Issues:

 Res olved recently - All Added recently Outstanding

2009/10/30

### 本日のお品書き

- (背景)
- ソフトウェア共同開発
  - 何に取り組んでいるのか?
  - どう取り組んでいるのか?
- 実践
  - どんな実践ができるようになったか?
  - 何が課題となっているか?

#### eSciDocの特徴

- オープンソースソフトウェア
- eScience情報の長期保存、発信
- MPDL, FIZ Karlsruhe(ドイツ)の共同開発
- SOA (Service-oriented architecture)
  - 公開APIの活用を前提とした相互運用
- データストア: Fedora(リポジトリ基盤ソフトウェア)
  - RDFをベースとした柔軟なアイテムレコード間の関連付けが可能
- 各種アプリケーション
  - PubMan: 文献情報リポジトリ
  - FACES: 写真アーカイブ
  - ViRR: 書籍スキャン&注釈

2009/10/30

#### NIMSにおけるデジタルライブラリー構想

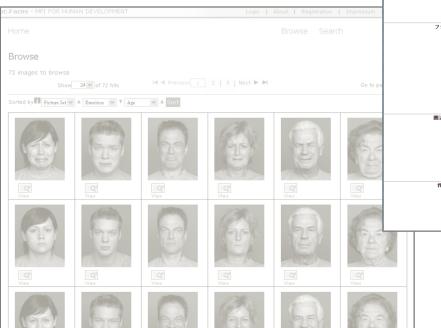
(NIMS Digital Library Project)

- eSciDoc共同研究開発 (NIMS-MPDL)
  - 2008年度から3年間の共同研究開発契約
  - MPDL (Max Planck Digital Library), Germany
  - eSciDoc(ディジタル図書館基盤ソフトウェア)の活用
  - 単一機関の枠を超えた連携可能性
- → 開発の場へコミットしながら、現場での運用を目 指す

2009/10/30

### NIMS eSciDocアプリケーション (Solution)

- 論文文献資料の保存、公開
  - PubMan
- 研究画像データの共有
  - Faces





### PubMan: 論文・出版物の保存、公開

- オープンアクセス論文・資料の保存、公開 (Open access)
- 特長 Features
  - 研究者総覧 Researcher Portfolio
  - メタデータ・本文流用入力 Fetch
  - 版管理 Versioning
  - 派生資料関連付け Relations
  - ローカルタグ Local tagging
  - アクセス統計 Visualization of usage statistics
  - 外部サービスとの多様なやり取り API

学術論文 资料種別 研究者総覧 (Researcher Portfolio) 公開 関質制限 MIME-Type / サイズ application/pdf / 2MB 出版社版 ファイル種別 到明 著者にID付与すると、 (三) 非表示: ファイル アイコンリンクが出現 関連URL 作成者 谷藤,幹子1,著者 作成者 高久,雅生1,著者 大塚, 真吾1, 著者 典, 贷市2, 著者 1 独立行政法人 物質・材料研究機構 科学情報室、〒305-0047 茨城県つくば市千現1 所属 2 独立行政法人物質・材料研究機構光材料センター、〒305-0047 茨城県つくば市千 escidoc:persistent22 著者IDリンク 非表示: 作成者 → 研究者ポートフォリオ リポジトリ、デジタルライブラリー、研究成果の 内容议明 キーワード Fedora, Scribd repository, digital library, effective dissemination DDC分類 研究成果の善穣と公開は、独立行政法人たる研究所においても重要な任務だが、それを 至旨 いう考え方はつい先ごろまでまったく浸透していなかった。そもそも図書館は、本を買っては 2009/10/30 識である。専門職員が充分にいるわけでもない。しかしさすがに情報社会の波が研究現場

mctp://da-babilan.ilibal.ilib&rae:onon/babilan/itelil/escianc:iozoo4:5

形 大棚 上東 サブノス にあり、犯 強 ヘルコルテム 病 起 小海 湯、彩 行事 にを口 ばたば ヒテキャル

51 MIUKL

### 研究者総覧

#### (Researcher Portfolio)

#### 研究者ポートフォリオ

#### PhD Todoroki, Shin-ichi

National Institute for Materials Science, Optronic Materials Center



研究者ブロフィール	現職	Senior Researcher (National Institute for Materials Science)
	現職	Faculty Member (Optronic Materials Center)
	ホームページ	Researcher Blog
	研究者ID	urn:cone:persons1008
	研究分野	Computer Sciences: Free, Libre and Open Source Softwar Libraries   Engineering: Materials Engineering; Ceramic En Engineering   Informatics   Library Science: Digital Libraries 挿入
External references	WorldCat	Search for Todoroki, Shin-ichi
	Google Scholar	Search for Todoroki, Shin-ichi
出版物		Todoroki, S. (2009). Longed-for functions for digital repository learned from my self-archiving activity. DRFtech-Kyoto: Workshop of application of Infrastructure for eScience and eResearch. Kyoto, 2009-09-07. [PubMan]
		Todoroki, S. (2009). My joy of self-archiving. 1st SPARC Japan seminar 2009: "Voluntary publication from researchers through a variety of network media in quest of dissemination to the general public". Tokyo, 2009-06-25 - 2009-06-25. [PubMan]
		Todoroki, S. (2009). Merci, XieXie, Spasibo. [PubMan]
		42

## PubManメタデータ/本文流用登録機能

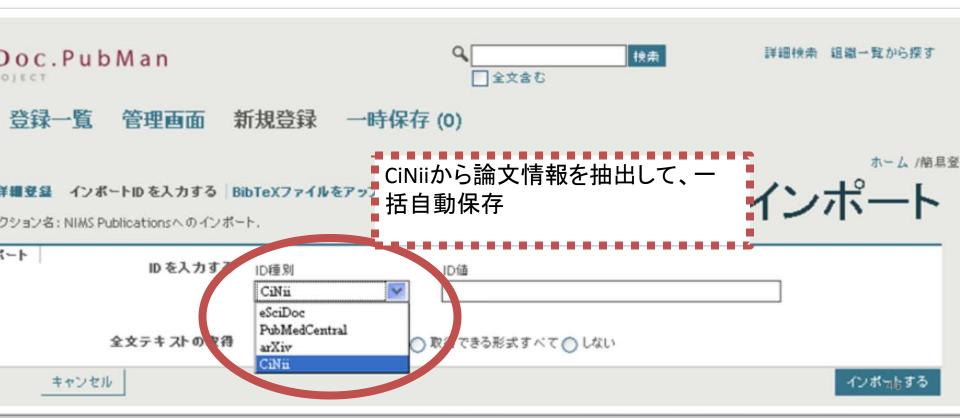
(Fetch)

- 登録を容易化し加速するための機能
  - (二度手間をなくす)
- 対応外部サービス
  - ArXiv
  - SPIRES
  - BioMedCentral
  - PubMedCentral
- 今年度開発予定:
  - 海外: CrossRef (DOI)
  - 国内: CiNii

# PubManメタデータ/本文流用登録機能

(Fetch)

- 登録を容易化し加速するための機能
  - (二度手間をなくす)
- 対応外部サービス



#### PubMan版管理機能

(Versioning)

検索!

#### メタデータ/本文の改訂履歴を自動保存

リリー ス履歴を表示 改訂版を表 ブレゼンテーション資料作成作業への Rule of Three 適用支援 ハンドアウト登録 公開日: 2009-08-29 3 バージョン 10 コメント handout fixed 公開 バージョン 8 公開日: 2009-08-29 14:16 コメント File open バージョン 6 公開日: 2009-08-28 03:52 公開日を確定 コメント handout should be open for ter 2009-08-29 ページ番号を確定 バージョン 4 公開日: 2009-08-27 14:24 コメント Page number changed 初期版を登録 バージョン 2 公開日: 2009-08-20 06:39 コッツト First release

発表前から発表後までの成果文献情報を一元管理

One single platform for pre- and post-publication info.)

### PubMan派生資料関連付け

基本情報

ファイル

曰 非表示:全項目

一 非表示:基本情報

引用URL

资料種別

ファイル名 説明

閲覧制限

MIME-Type / サイズ

その他のタイトル

講演

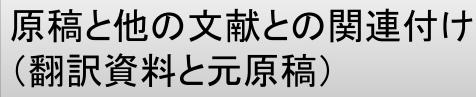
handout

公開

| 詳細 | リリニス履歴を表示 | 図訂版を表示 | 利用統計を表示

(Relations)

組織一覧から探す >検索結果



(translation and original manuscript)

詳細 | リリース履歴を表示 | 改訂版を表示 | 利用統計を ree and LaTeX http://pubman.mpdl.mpg.de/pubman/item/escidoc:94908:11 【英訳】プレゼンテーション資料作成作業への Rule of Three 適用支援 Etex pres+sem4.pdf (付録資料) [MD5]

プレゼンテーション資料作成作業への Rule of Three 適用支援



application/pdf / 101KB

Upgrade your slides with Rule of Three and LaTeX 易一覧 Todoroki, Shin-ichi 2009-08-29, 紙版出版 資料種別 1ファイル

個別文献にとどまらない文献同士 の関連付けを明示できる。

※他にも、ユーザ参加によるキー ワードタグ付けなどでの整理提示

### PubManアクセス統計機能

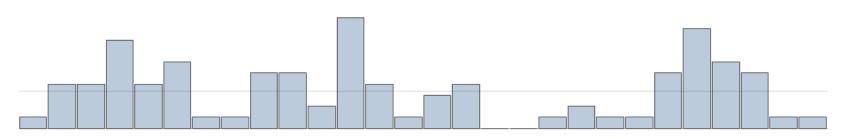
(Visualization of usage statistics)

View Results (Author = todoroki) (Combined Paper ID) のアクセス分析 2月2009
20098月20日木曜日,15:56 (5分前)に最後のアップデートをしました。 今月の95アクセス件数(62ユニーク)の合計、3.1一日の平均(2.0ユニーク)。

今月 月 時間 ブラウザ 国 ファイル オペレーティングシステム ページ リファラー 訪問話問者 検索 セッション ステータス

今月のアクセス件数

概略 | アクセス件数をカレンダーで見る | アクセスページ件数をカレンダーで見る | ヒット件数をカレンダーで見る | バンド幅使用量をカレンダーで見る



В	<b>\$</b>	日付	1	間者数	\$	ページ 💠	平均 ♦	ヒット件数 💠	平均 ♦	バンド幅 🏻 💠	平均 ◆
日曜日		12月'09			1	2	2.0	2	2.0	106.0k	106.0k
月曜日		22月'09			4	21	5.3	21	5.3	1.1M	284.0k
火曜日		3 2月 '09			4	6	1.5	6	1.5	322.1k	80.5
水曜日		42月'09			8	29	3.6	29	3.6	1.5M	193.3
木曜日		5 2月 '09			4	16	4.0	16	4.0	834.4k	208.6
金曜日		6 2月 '09			6	15	2.5	15	2.5	794.7k	132.4
土曜日		7 2月 '09			1	1	1.0	1	1.0	53.2k	53.2
日曜日		8 2月 '09			1	1	1.0	1	1.0	53.3k	53.31
月曜日		9 2月 '09			5	8	1.6	8	1.6	436.6k	87.3
火曜日		10 2月 '09			5	10	2.0	10	2.0	522.6k	104.5
水曜日		11 2月 '09			2	2	1.0	2	1.0	109.2k	54.6k

### PubManアクセス統計機能

(Visualization of usage statistics)



光ヒューズ 光ヒューズ 光ファイバヒューズ **光ファイバ ファイバーヒューズ** 

**国際出願 査定 実験ノート パソコン** 実験ノート 例

### 外部サイトでのデータ再利用

(API)

#### **Researcher** Page

Home

自動的に年月順の発表論文アーカイブ 論文付与キーワードによるブラウジング

#### **Home**

Last modified: August 25, 2009

#### Todoroki, Shin-ichi

Senior Researcher

- 🗂 Optronic Materials Center
- Mational Institute for Materials Science



#### About

Working on optical fibers being exposed to high power laser beam.

Interested in digital publishing using free software, including writing light essays.

外部サイト上にPubMan上の情報を 自動挿入、リンク付与 (Wordpress (CMS) 対応)

#### Categories

- > Blog (1)
- > Carbon (2)
- > critical cooling rate (1)
- crystallization tendency (1)
- electronic notebook(2)
- > Erbium (1)
- > Fiber fuse (12)
- > finite element method (1)
- > fluorescence lifetime(1)
- > Glass (10)
- > high-throughput experimentation (1)
  - brid device (2)
  - formatics (7)
  - ertion loss (1)
  - titutional repository

ser-induced damage

ht-induced formation (1)

Material Informatics

#### Archives

- □ 2009(7)
  - > September (1)
  - > August (2)
  - > July (3)
  - > February (1)
- ± 2008(2)
- **±** 2006(3)
- ± 2005(6)
- **±** 2004(9)
- ± 2003(2)
- \_\_\_\_\_
- ± 2002(4)

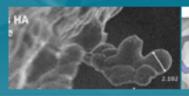
#### RSS

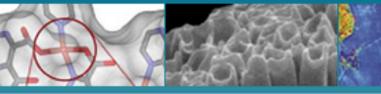


🚮 Comments

### 外部サイトでのデータ再利用 (API)

#### Researcher Page





Home

#### Archive for July, 2009

escidoc:69880

Thursday, July 23rd, 2009

Todoroki, S., **My joy of self-archiving** (2009). (Translated viewgraphs of the talk given in the **1st SPARC Japan seminar 2009**: "Voluntary publication from researchers through a variety of network media in quest of dissemination to the general public".)

[pubman]

Original in Japanese: [pubman]

#### 整形済み文献情報とPubManへのリンク

Posted in Fiber fuse, Self-archiving, Serendipity | No Comments

#### Categories

- > Blog (1)
- > Carbon (2)
- critical cooling rate (1)
- crystallization tendency (1)
- electronic notebook(2)
- > Erbium (1)
- > Fiber fuse (12)
- > finite element method
  (1)

> Glass (10)

> high-throughgut experimentation (1)

### 主題リポジトリ・多機関共同リポジトリ

複数の国・機関をまたいだ、同居型共同リポジトリ (Cooperative repository)

组筐



- 🖮 吞 Kaiser-Wilhelm-Gesellschaft [組織説明]
- 由 Max Planck Society [組織説明]

#### 今年度後半には国内での連携テストを開始予定

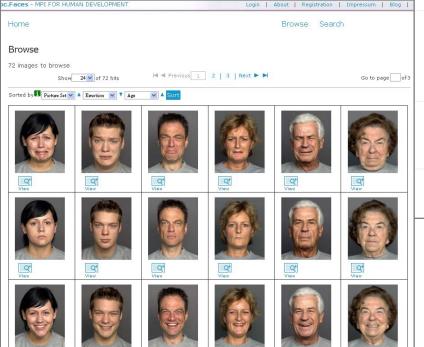
・日本大学新学術創成プロジェクトにおける4研究分野

医学 (Medical)、物性化学 (Chemistry)、アートアーカイブ (Archives)

組織説明

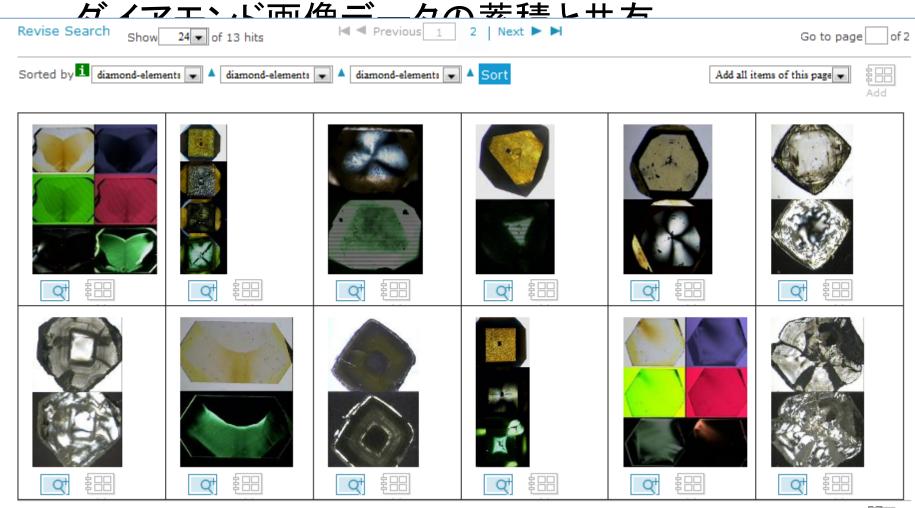
### NIMS eSciDocアプリケーション (Solution)

- 論文文献資料の保存、公開
  - PubMan
- 研究画像データの共有
  - Faces

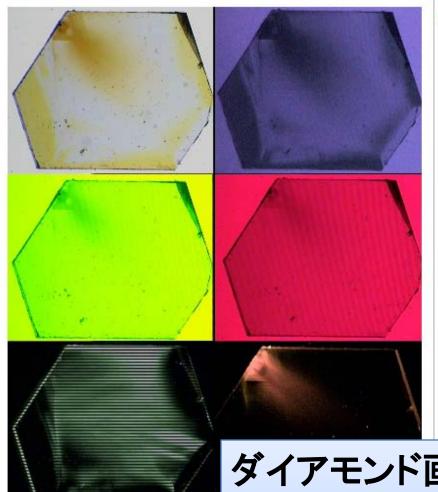




#### Diamond-Faces



#### Diamond-Faces (Detail view)



diamond-item

diamond-elements: C

shape : crystal

color:

mechanism : HPHT

condition : Optical microscope

temperature: 300 K

brief-details : synthetic diamond

full-details : Polished synthetic diamond plate viewer

transmission (white, violet, green or rea

or luminescence (337 or 515 nm exc.). Stripes in the violet and red im are interference; green and red luminescence is due to dissimilar Ni-N complexes. Image width 2 mm.

date: 26:03:2009

identifier: 2707

#### ダイアモンド画像のメタデータスキーマを

属性定義 (User-defined metadata schema)

元素、形、色、合成手法、実験条件、温度、説明、日付など (→物質・対象に応じて柔軟に定義可能)

temperature

### Diamond-Faces: eScienceのための特長

(Features for eScience)

- 属性定義 (Configurable metadata schema)
  - 柔軟に属性(色・形・温度など)を定義可能
- 参照URL
  - 固定URL (Permalink)
  - 複数の画像集合を単位としたひもづけ (Album)
  - → 研究論文から対象データとして引用・参照でき、 関連資料としてまとめた情報を共有できる。

### まとめ

(Conclusion)

- 研究者の活動、ニーズによりそったリポジトリシステムの要件とその研究開発
- 2つのアプリケーションの運用を進めている
  - PubMan: 文献資料 (+研究者総覧)
  - Faces: 画像データ
- 課題(今後の取り組み)
  - データ入力・登録の簡便化に向けた工夫
  - 所内システムおよび外部サービスとの連携を図る
    - 所内: 研究発表許可願システム、人事DB
    - 外部:トムソンResearcherID、NII研究者リゾルバ

### 今後の予定(宣伝)

- 図書館総合展: 2009年11月10日~12日
  - @横浜
  - FIZ Karlsruheの開発者責任者 Matthias Razum氏を招聘
  - フォーラム開催予定:
    - 11月11日 13:30~14:30
    - eSciDoc/Fedora: e-Researchのための機関リポジトリの新展開ーデジタルライブラリーからeScienceへ向けて
- 次回研究談話会
  - 11月13日(金)
  - Matthias Razum氏による講演
- デジタルリポジトリ連合国際会議2009
  - @東京工業大学
  - 12月2,3日
  - ポスター発表および口頭発表