九州大学学術情報リポジトリ
Kyushu University Institutional Repository

機関リポジトリの潜在需要と 利用者動向の解析

池田, 大輔
九州大学システム情報科学研究院准教授

https://hdl.handle.net/2324/19332

出版情報：附属図書館研究開発室成果報告会, 2011-03-16
バージョン：
権利関係：

機関リポジトリの潜在需要と利用者動向の解析

池田 大輔

九州大学 システム情報科学研究院 情報学部門

2011/03/16 第1回研究開発室成果報告会

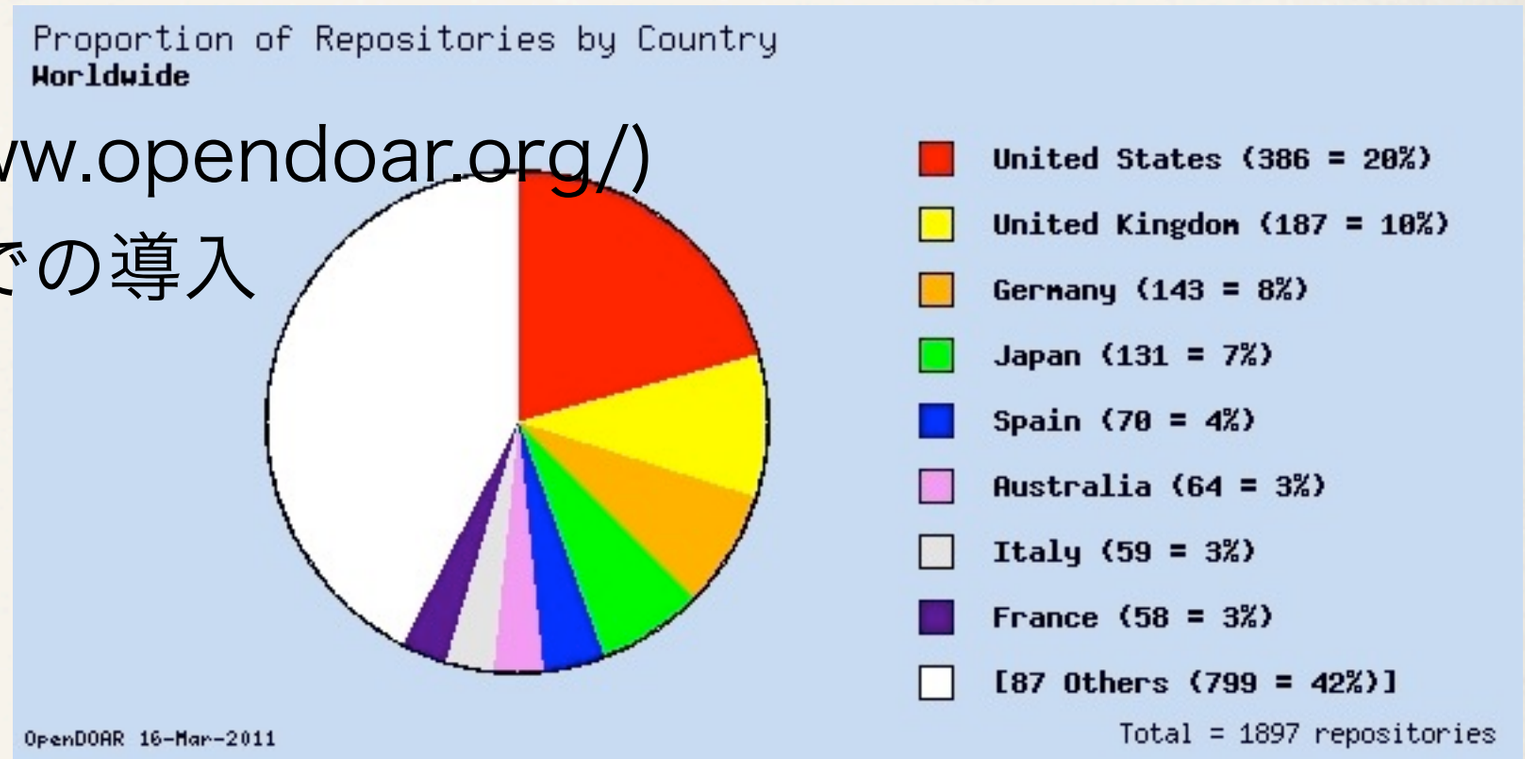
文献情報

- OR’09
  - Daisuke Ikeda and Sozo Inoue
    "Access Flows to a Repository from Other Services"
    The 4th International Conference on Open Repositories.
- デジタル図書館ワークショップ
  - 池田 大輔、星子 奈美、井上 創造
    外部連携サービスによる機関リポジトリの潜在需要の解析
    ディジタル図書館 Vol.36, pp.62-68, 2009.
- OR’11
  - Peng Wang and Daisuke Ikeda
    "Analysis of Search Keywords in Accesses from Search Engines to a Repository"
    Submitted to OR’11
  - Peng Wang (王鵬くん：シス情 情報学 M1)

背景：機関リポジトリの普及

- 雑誌危機→オープンアクセスから、機関リポジトリへ

- 世界中で1,900弱の導入
  - OpenDOAR(http://www.opendoar.org/)
  - うち、日本は131機関での導入



- 日本で142機関での導入
  - NIIの「機関リポジトリ一覧」より
    - 学術機関リポジトリ構築連携支援事業
  - 試験公開なども含まれる

今後は、内容の充実が重要

動機：機関リポジトリの影響分析が必要

- 機関リポジトリの影響
  - 研究者：メンテナンスフリーな発表の場を確保できた。
  - 機関(図書館)：ログが入手可能になった。
    - 電子ジャーナルでは不可能!!
  - コンテンツの利用者：学術成果にオープンアクセスが可能。
    - オープン：フリー?
    - 「誰でも」アクセス可能になったことが重要ではないか?

新たな利用者がいるのではないか?

目的：リポジトリのコンテンツ利用の解析

- 機関リポジトリのコンテンツがどのように利用されているか調査する
  - 利用者のニーズに応えられているか？
  - 新しい利用者はいるのか？
  - どのような目的で学術成果を探しているのか?

- 期待される効果
  - 新しい利用者にあわせた新しいアクセスパスを用意できる。
    - if 留学生が留学先の先生を探すのにリポジトリを使っている
      then 留学生向けページとQIRの連携
    - if 企業の研究者が共同研究先の先生を探すのにリポジトリを使っている
      then 知財本部等とQIRの連携

手法：アクセスログの解析

- 九州大学学術情報リポジトリ(QIR)のログを解析する。
- 特に、
  - ハーベスタ、検索エンジンからのアクセス
    - どこからきているのか？
  - 九州大学研究者価情報データベースからのアクセス
    - 需要はQIRにより満されているのか?
  - 検索エンジンにおけるキーワードの解析
    - どのような意図で探しているのか？

に重点を置く。

準備：機関リポジトリの仕組み

機関リポジトリ上のコンテンツ

メタデータのページ

https://qir.kyushu-u.ac.jp/dspace/handle/2324/12515

QIR 九州大学学術情報リポジトリ Kyushu University Institutional Repository

アイテム投稿ヘルプ
アイテム編集ヘルプ
言語切替／Change Language

九州大学

Google QIR Search
Go

> ホーム

ブラウズ
> 組織別一覧
> タイトル
> 日付

> 新着情報サービス

> 文献の投稿
> ユーザ情報編集

> このサイトについて
> ヘルプ

文献を探す
検索
詳細検索

Kyushu University Institutional Repository (QIR) >
附属図書館 = Library >
九州大学附属図書館研究開

このアイテムのアクセス数

このア

タイトル:
タイトル(別表記):
著者:
著者名ヨミ:
著者名(別表記):
著者情報:

出版年:
刊行元:
刊行元(別表記):
掲載情報:

PDFのページ

https://qir.kyushu-u.ac.jp/dspace/bitstream/2324/12515/1/2007_p017.pdf

論文

学術情報リポジトリ活性化のための足跡機能

井上創造† 藤井達朗‡ 小材健§ 池田大輔**

<抄録>
学術情報リポジトリの未登録コンテンツも含めたアクセス履歴を研究者に提示することで，研究者のリポジトリへの登録の動機づけを目指すシステムを実装した例を紹介する．未登録コンテンツは九州大学研究者情報から得ることができる．研究者情報の文献を閲覧者がたどってリポジトリにアクセスした履歴を，SNS（Social Networking Service）の研究者用マイページに表示することで，研究者はリポジトリに未登録の自分のコンテンツにどのくらいのアクセスがあったかを知ることができるようになり，その研究者がコンテンツをリポジトリに登録する動機づけになるばかりでなく，アクセスの多い論文から優先的にリポジトリに登録することにつながるため，真に利用価値の高いコンテンツの蓄積が可能となる．

<キーワード> 機関リポジトリ，学術情報リポジトリ，SNS，動機付け

Footprint Visualization for Motivating Academic Information Repositories

INOUE Sozo FUJII Tatsuro KOZAI Takeshi IKEDA Daisuke

アクセスログ

- Webサーバ(apache)のログファイルを利用
  - 著名な検索エンジンやクラッキングのログ等を削除している
  - http statusは 200 or 304

アクセス先

- ログの例

1.2.3.4 - - [27/Dec/2008:00:40:55] "GET /dspace/…/6517 HTTP/1.0" 200 14962 "-" "Mozilla/5.0 (…)"

5.6.7.8 - - [27/Dec/2008:00:40:56] "HEAD /dspace-oai/request?…ac.jp:2324/1407 HTTP/1.1" 200 - "-" "-"

9.0.1.2 - - [27/Dec/2008:00:41:07] "GET /dspace/browse-title?starts_with=S HTTP/1.0" 200 25200 "-" "Mozilla/5.0 (compatible; Yahoo! Slurp; http://help.yahoo.com/help/us/ysearch/slurp)"

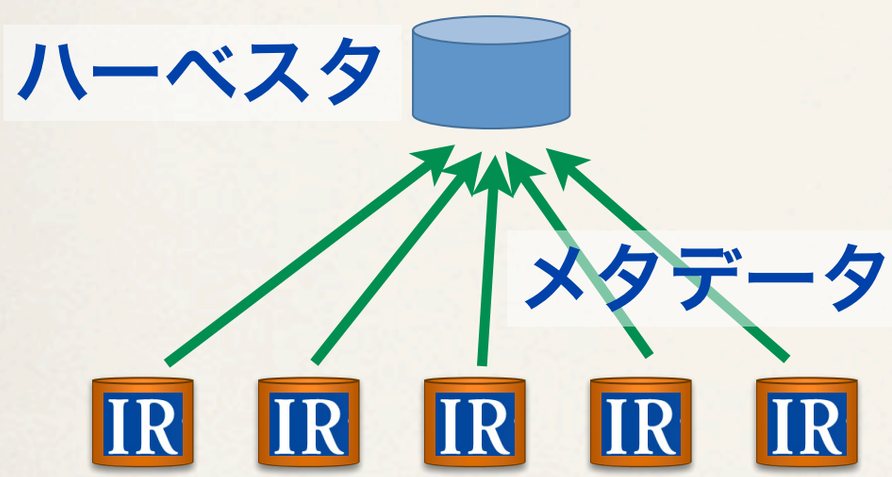
3.4.5.6 - - [27/Dec/2008:21:11:11] "GET … HTTP/1.1" 200 4369 "http://leda.lib.kyushu-u.ac.jp/… au_id=K000008…" "Mozilla…"

リファラー
(referer)

  - メタデータのページ https://qir.../dspace/handle/2324/12515
  - PDFのページ https://qir../dspace/bitstream/2324/12515/1/file.pdf

機関リポジトリ上のコンテンツの探し方

- 機関リポジトリ上の検索機能を利用
  - どこのリポジトリ上に求めるコンテンツがあるか、知っておく必要がある。
- ハーベスタを利用
  - OAI-PMHを用いて、多くのリポジトリのメタデータを収集した検索エンジン。

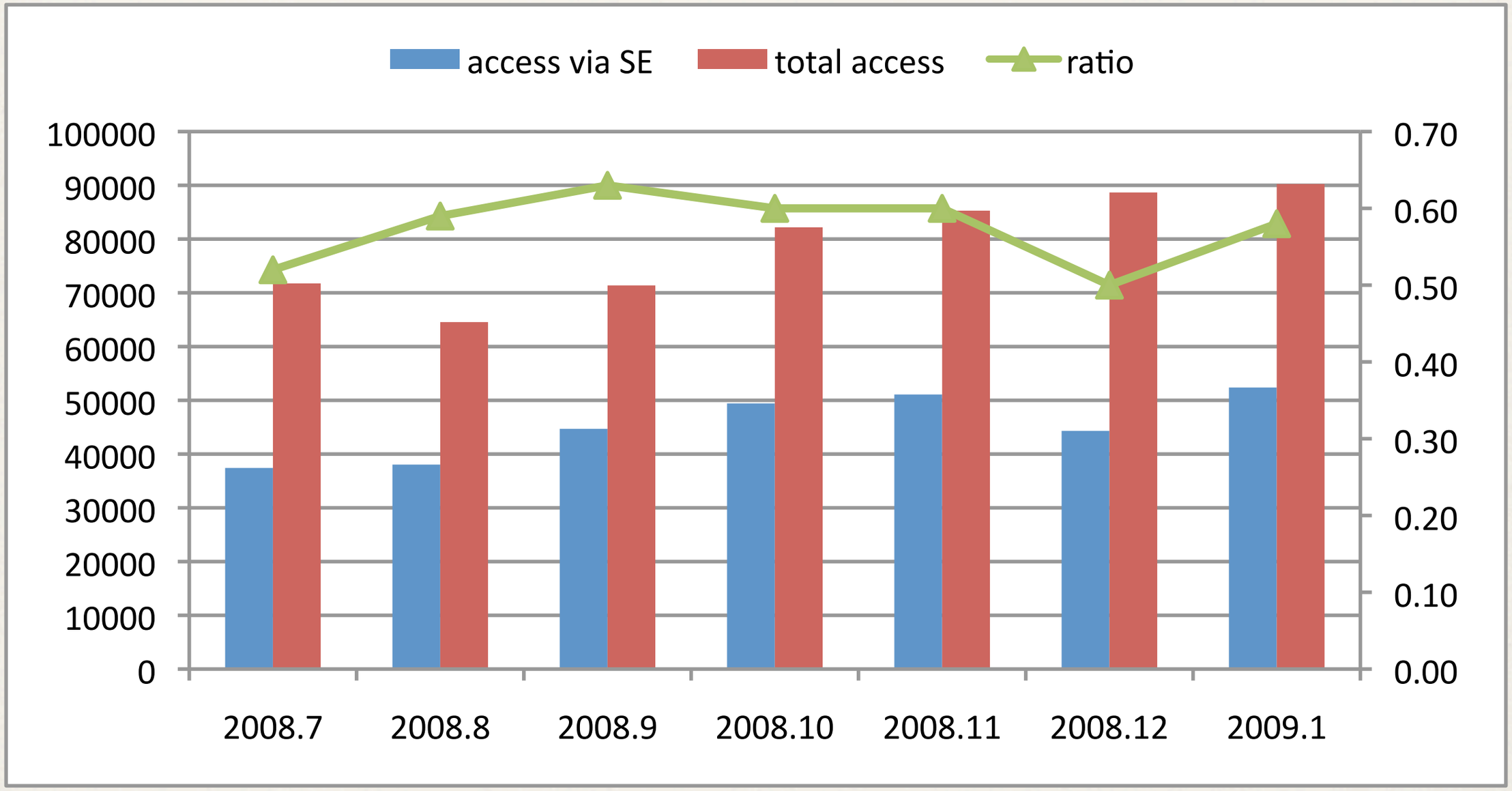


- 検索エンジンを利用
  - Google等

解析結果：どこから来たのか?

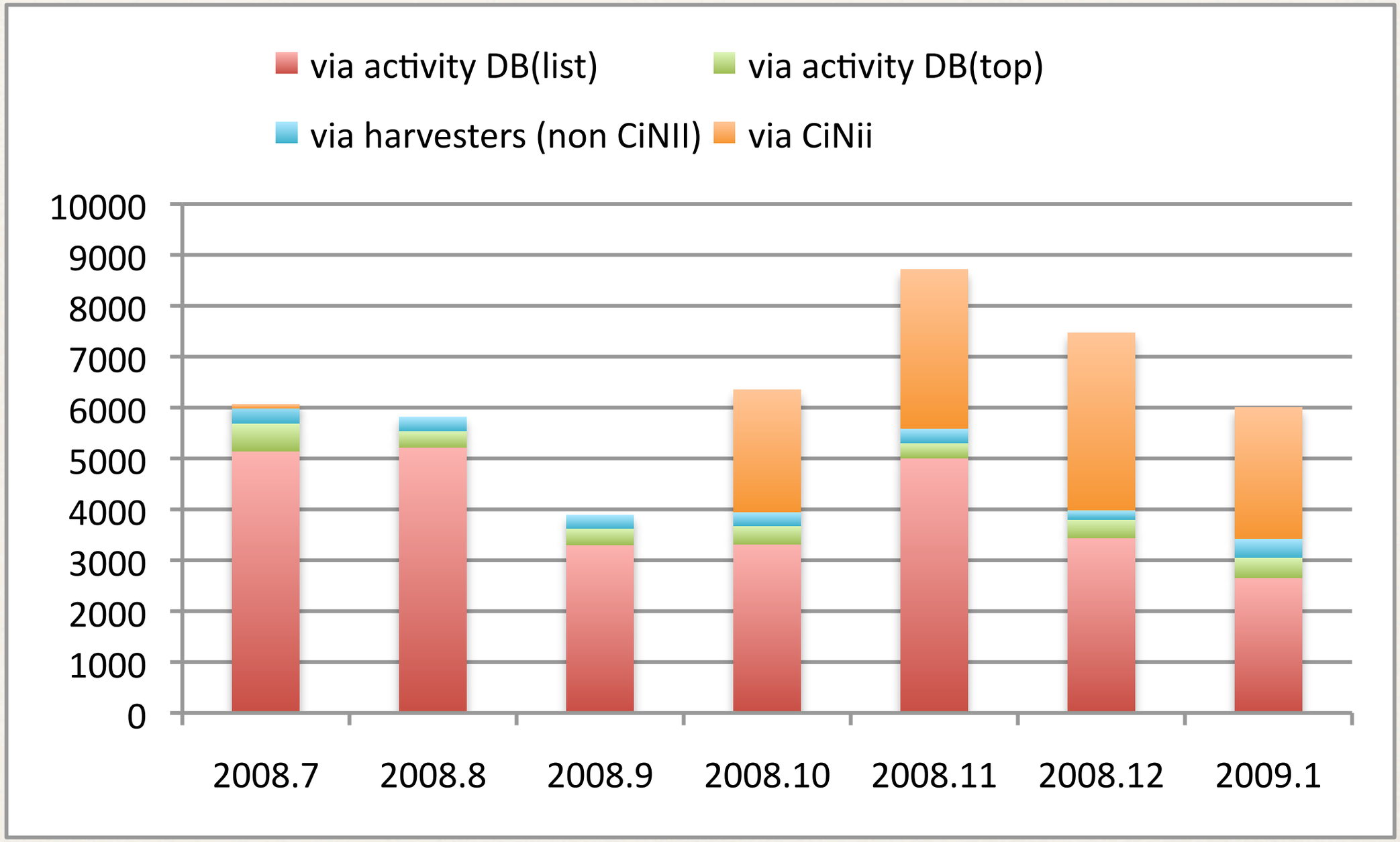
サーチエンジン経由

期間：2008.7〜2009.1



ハーベスタ経由(では来ていない!)

期間：2008.7〜2009.1



解析結果：充足率

研究者情報からQIRへのリンク

研究業績

主要原著論文

全てを見る >>

1. Daisuke Ikeda and Sozo Inoue,A Sustainable Model based on the Social Network Service to Support the Research Cycle,*The 3rd International Conference on Open Repositories*,,2008.04 http://pubs.or08.ecs.soton.ac.uk/17/. FulltextQIR
2. 案浦 スミタカ, 荒木 寛幸, 井上 創造, 池田 大輔,SNSを用いた研究支援と認証認可付き検索システム,*ディジタル図書館*,Vol. 34, pp. 21--32,2008.03. FulltextQIR
3. Daisuke Ikeda and Sozo Inoue,A New, Sustainable Model for the Institutional Repository: A CSI Project "Integration and Presentation of Diverse Information Resources",*DRF International Conference Open Access and Institutional Repository in Asia-Pacific*,p58,2008.01.
4. 池田 大輔, 宗 孝, 納富 貞嘉, 井上 創造,研究支援ツールとしての機関リポジトリにおけるバージョン管理機能,*ディジタル図書館*,Vol. 33, pp. 31--38, November, 2007.,2007.11. FulltextQIR
5. 池田 大輔,ユビキタス時代におけるハイブリッドライブラリー型電子図書館モデル, *名古屋大学附属図書館研究年報*,pp.13-19, No. 4,2006.09.

主要総説, 論評, 解説, 書評, 報告書等

全てを見る >>

1. 池田 大輔, 安浦 寛人,*全学共通ICカードの学内実証実験報告とその基盤技術について*,九州大学情報統括本部 ITマガジン,pp13-21,2007.06. FulltextQIR

研究者情報

九州大学 Academic Staff Educational and Res

研究者情報

基本情報 研究活動

池田 大輔 (いけだ だいすけ)

准教授

システム情報科学研究院 情報理学部

大学院(学府)担当

システム情報科学府 情報理学専攻

学部担当

理学部 物理学科情報理学 コアセミ

理学部 物理学科情報理学 データ科

理学部 物理学科情報理学 情報検索

その他の教育研究施設名

092-802-3786

取得学位

中間システム

http://leda.lib.kyushu-u.ac.jp/link/link.php?…&au_id=K000021&…

&au_f_name=池田&au_l_name=大輔&title=A%20Sustainable%20Model…

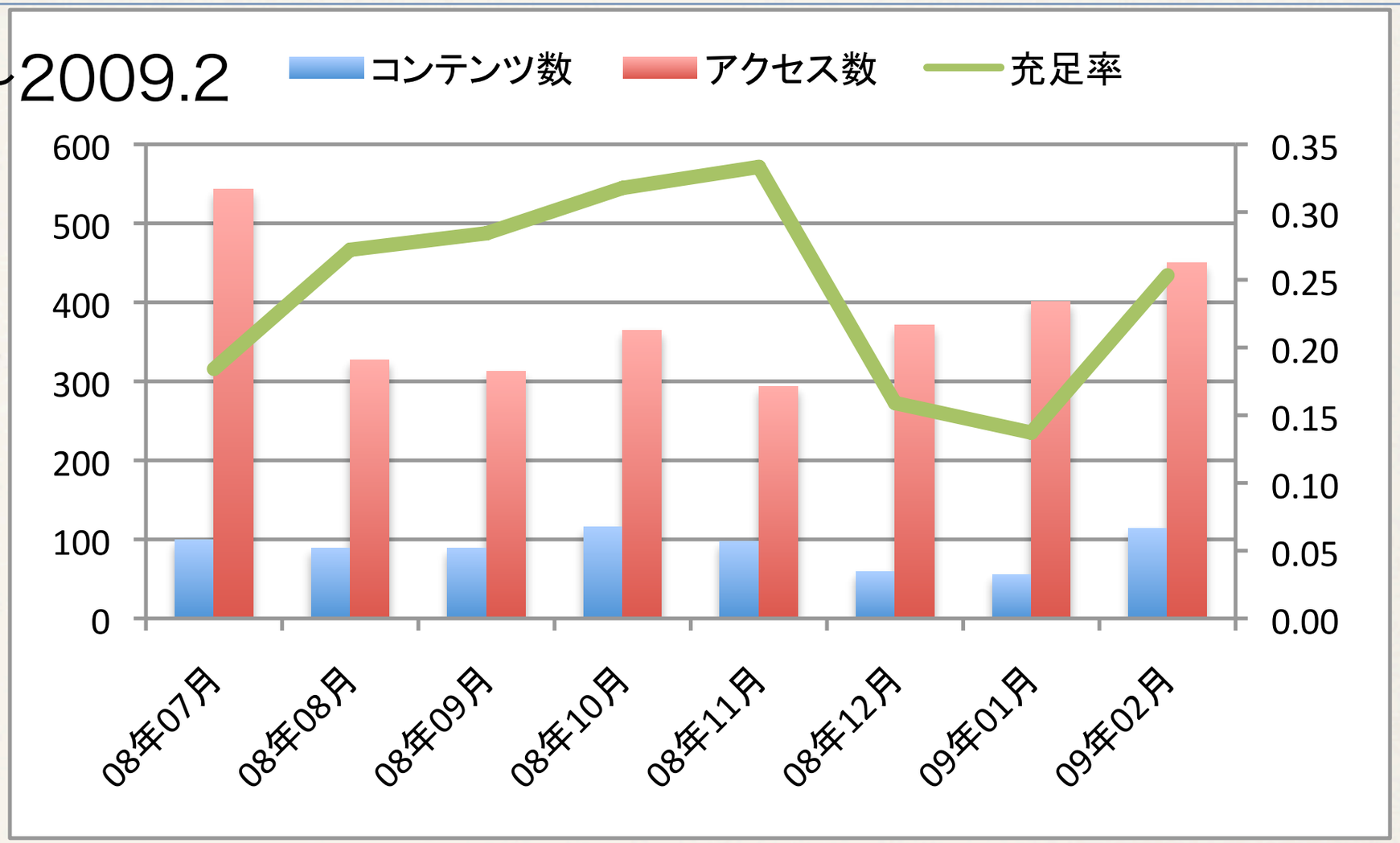
QIRにコンテンツがなくてもリンクできる!

館の個人情報管理：…

- Coloring for Pattern Detection
- Pattern Matching

研究者情報DBを経由したQIRへのアクセス数の推移

期間：2008.7～2009.2



- 充足率＝ $\dfrac{\text{中間システムからQIRへ到達できた回数}}{\text{研究者情報DBでのアイコン(RSSも含む)をクリックした回数}}$

充足率の平均値は0.24

解析結果：検索キーワード

検索キーワードの調べ方

- 検索エンジンからのアクセスの場合、リファラに検索キーワードが埋め込まれている。

12.34.56.78 - - [21/Jun/2008:15:43:00 +0900] "GET /dspace/handle/2324/10596 HTTP/1.1" 200 6626 "http://www.google.co.jp/search?hl=ja&q=%E8%81(中略)7%B4%A2&lr=" "Mozilla/4.0"

アクセス先

復号化→

http://www.google.co.jp/search?hl=ja&q=職場におけるワーク・ライフ・バランスの必要性&btnG=検索&lr=

用語の説明

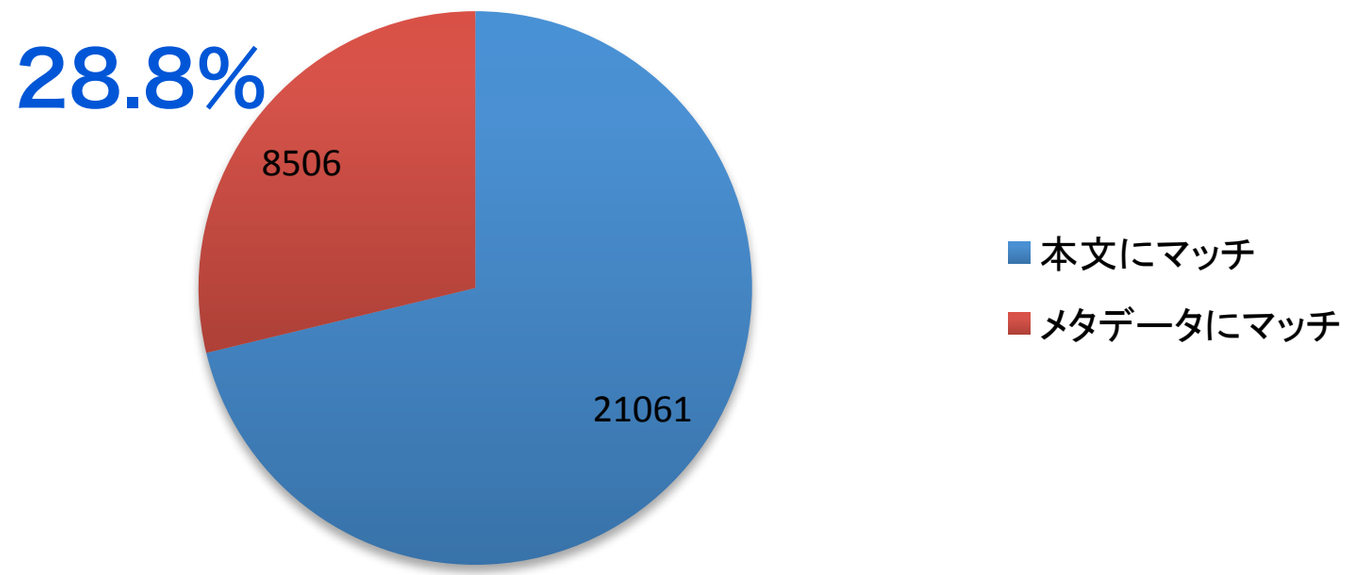
- キーワード
  ⇔QIR に、検索エンジン経由できた場合のキーワード
  - 複数でもよい

- キーワードがメタデータにマッチ
  ⇔キーワードのうち、一つでもメタデータに含まれている

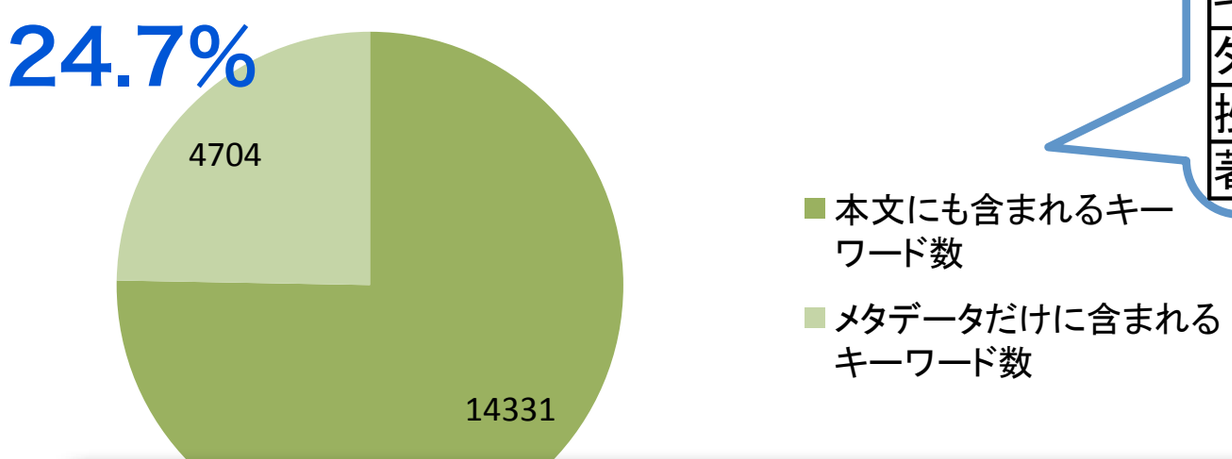
- キーワードが本文データにマッチ
  ⇔キーワードが一つもメタデータにマッチしない

解析結果

PDFページへのアクセスの割合



「メタデータにマッチ」の内訳



本文にも含む属性	キーワード数
キーワード	3429
タイトル	8282
抄録	2190
著者	430

52%=PDFファイルへのアクセス

純粋なメタデータ中のキーワードアクセス

52%×28.8%×24.7%≒3.7%

まとめ

- 機関リポジトリのアクセスログ解析により、利用のされ方の一部を明らかにした。
- 特に、
  - 検索エンジン経由が多く、ハーベスタはあまり使われていない。
  - 研究者情報DBからのアクセスのうち1/4程度しかコンテンツを収集していない。
  - 検索エンジで用いられるキーワードのほとんどは、本文に埋めこまれており、メタデータは検索時には役にたっていない。
    - メタデータが不要なわけではない!!
    - サービス連携、データ共有などにはメタデータが重要
- 今後の課題：
  - より詳細な利用者像を描きだしたい。