

2010年2月17日(火)

金沢大学創基150年記念「講演・シンポジウム」シリーズ(特別回)／DRF地域ワークショップ(Kan-Nihonkai)
大学図書館のミッションと機関リポジトリ:環日本海域からの情報発信

リポジトリログ分析による 学術情報流通の諸側面

筑波大学大学院図書館情報メディア研究科

佐藤翔 (min2fly@slis.tsukuba.ac.jp)

本日の内容

1. はじめに: ZSプロジェクトの紹介
2. アクセスログ分析とは？
3. 8つの機関リポジトリのログ分析結果
4. 学術情報流通における機関リポジトリの役割と問題
5. おわりに

ZSプロジェクトとは？

- “Zoological Science meets Institutional Repositories” Project
- 北大、京大、日本動物学会、筑波大等が参加

ZSプロジェクトとは？

- “Zoological Science”論文を機関リポジトリに収録
 - その後の被引用数の変化を分析
- プロジェクトの目的
 - 機関リポジトリによるオープンアクセス(OA)の効果を検証する
 - 「リポジトリに登録することが新たな引用を引き起こすか？」

OA・機関リポジトリと被引用数

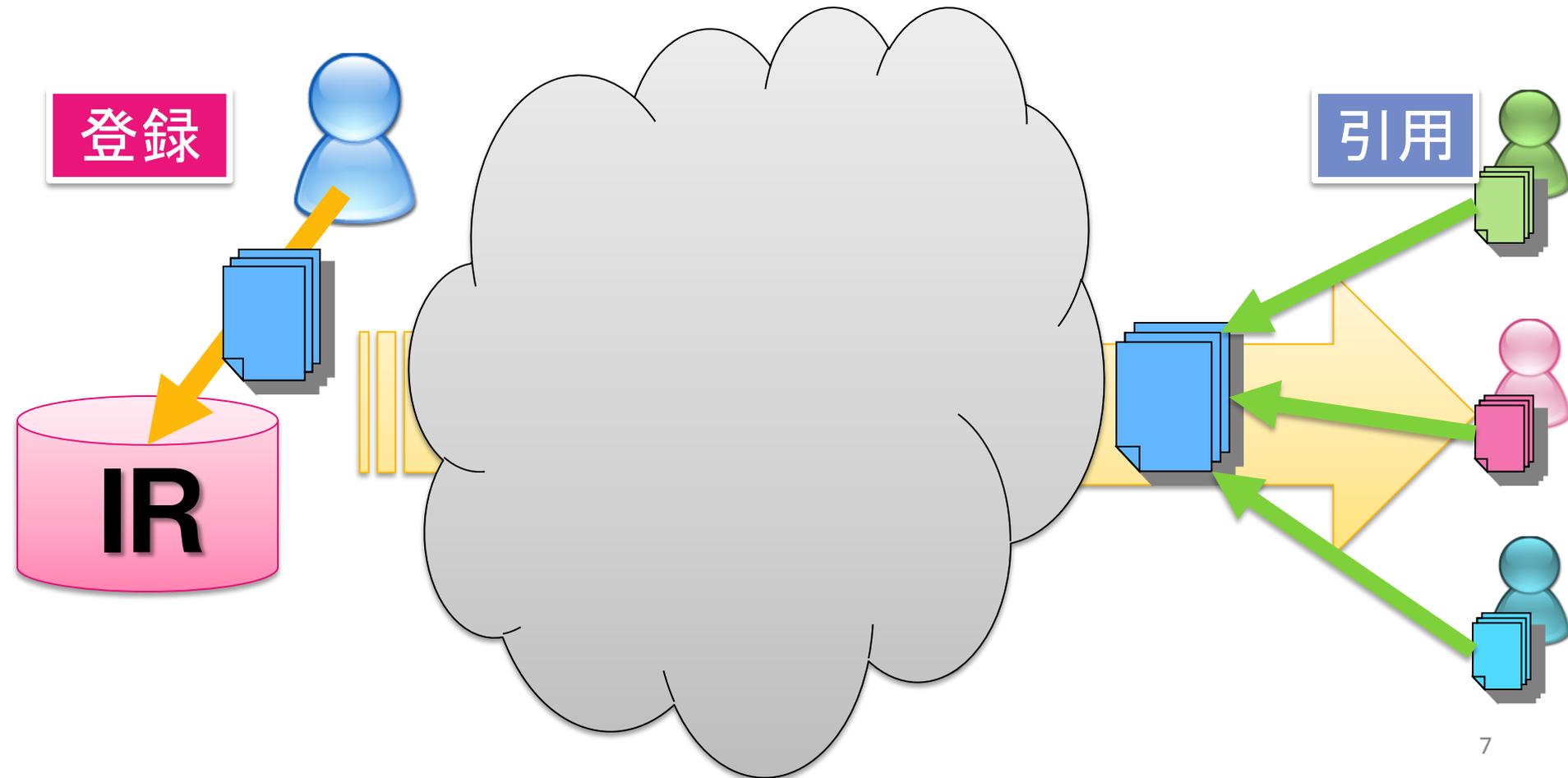
- 機関リポジトリはOAの手段の一つ
(著者自身によるアーカイブ)
- OAにすれば誰でも読める
⇒ 被引用数が増えるとする主張
- 機関リポジトリに登録する=OAにすると(しないより)よく引用される??

OA・機関リポジトリと被引用数

- OAの被引用数増効果に関する研究は多い
 - 分野別リポジトリ(arXiv等)
 - 電子ジャーナルの部分公開
 - OAであれば対象を限定しない
- 機関リポジトリに限定した研究はない
 - 分野別リポジトリや電子ジャーナルでの結果を援用できるようにも考えられない
 - 「リポジトリに登録することの効果」を検証する必要

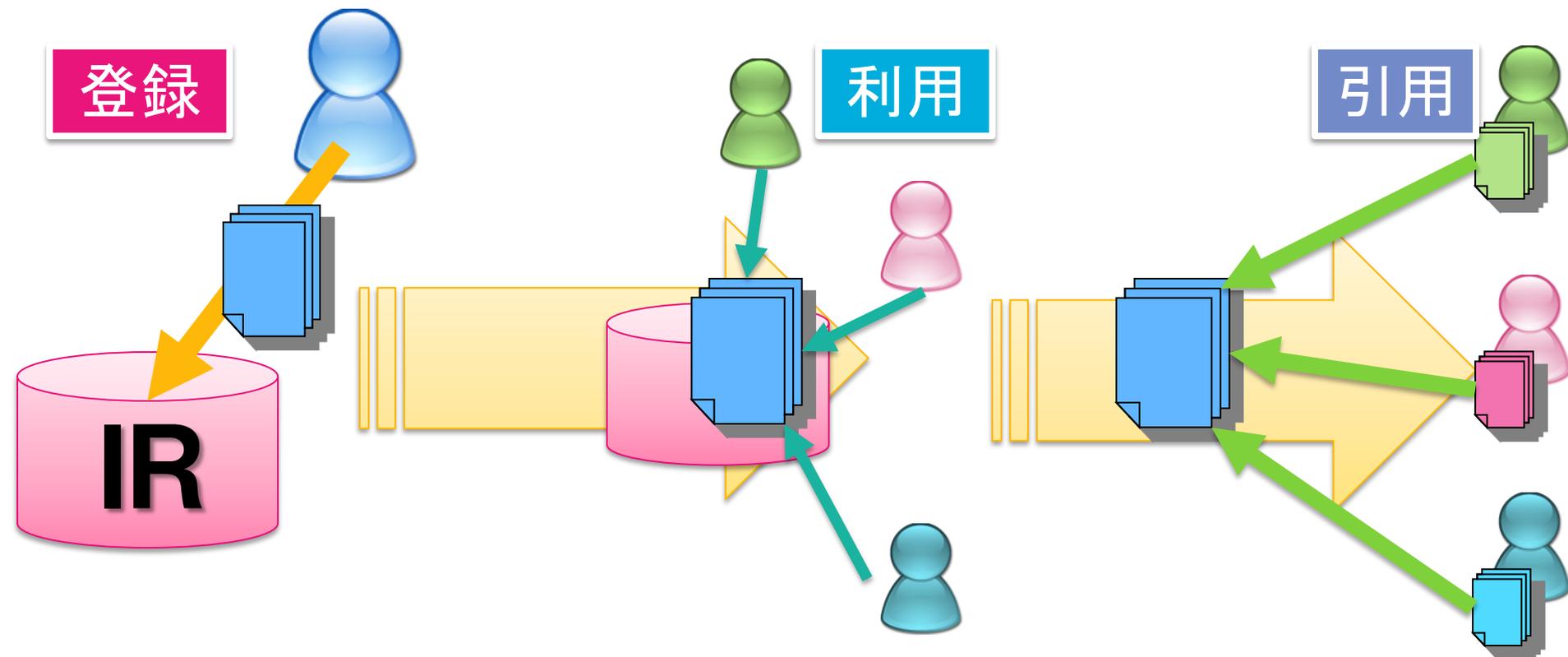
ZSプロジェクトとは？

- 「リポジトリに登録」⇒「被引用数が増える」？

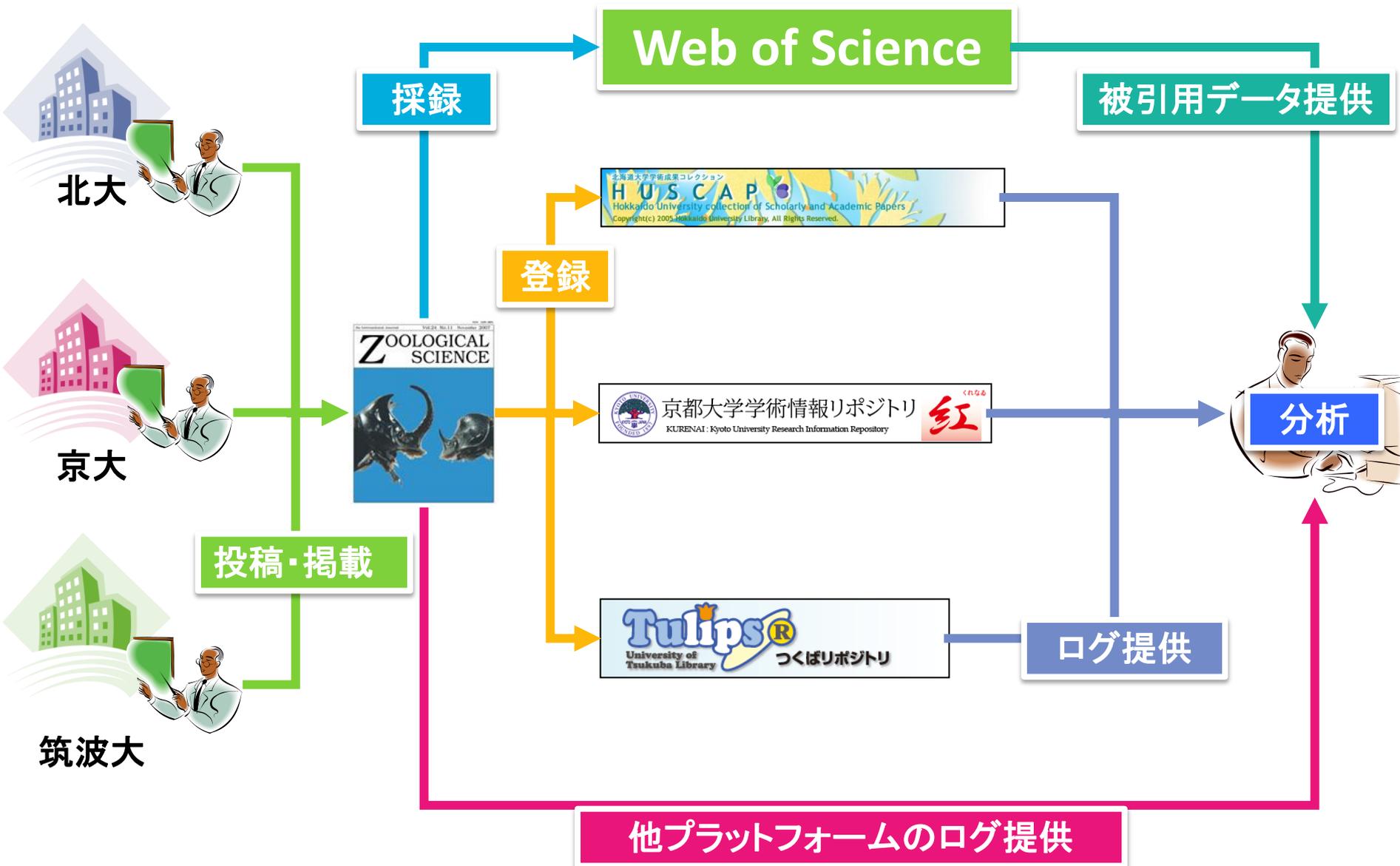


ZSプロジェクトとは？

- 利用記録を加味！



プロジェクトの流れ



ZSプロジェクトの状況と リポジトリ利用分析

- プロジェクトは鋭意進行中
 - 北大、京大、筑波大で論文収録開始
 - 被引用数データも修得
- 被引用数分析には時間がかかる
 - 閲読⇒研究⇒論文執筆⇒掲載のタイムスパン
- **まずは現在のリポジトリの利用状況を見よう**
 - 佐藤の修士論文のネタにも！
 - 2010.1.14提出、無事修了も確定しました！！！！

本日の内容

1. はじめに: ZSプロジェクトの紹介
2. アクセスログ分析とは？
3. 8つの機関リポジトリのログ分析結果
4. 学術情報流通における機関リポジトリの役割と問題
5. おわりに

ログ分析とは何か？

```
133.51.6.132 - - [27/Jul/2008:07:57:01 +0900] GET  
/dspace/handle/2433/49548/1/ HTTP/1.1 200 424264 - Mozilla/4.0  
(compatible; MSIE 6.0; Windows NT 5.1; SV1; .NET CLR 1.1.4322;  
InfoPath.2; .NET CLR 2.0.50727)
```

```
66.249.70.136 - - [01/Jul/2008:04:02:10 +0900] "GET /dspace/simple-  
search?query=purification&start=90 HTTP/1.1" 200 32835 "-"  
"Mozilla/5.0 (compatible; Googlebot/2.1;  
+http://www.google.com/bot.html)
```

-
-
-

133.51.6.132 - - [27/Jul/2008:07:57:01 +0900] GET
/dSPACE/handle/2433/49548/1/ HTTP/1.1 200 424264 -
Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1;
SV1; .NET CLR 2.0.50727) ath.2; .NET CLR

IPアドレス

行動の結果

アクセス先
ファイル名

アクセス
日時

アクセスログ分析登場の背景

- 電子ジャーナルの普及・・・
 - 利用頻度(ダウンロード回数)が把握できる
 - 誰(正確にはどのPC)がどんな時間に、どんな経路をたどって(どんな検索語で)アクセスしてきているかも把握できる
- ユーザのニーズをより正確に把握した資料構築・サービス構築が可能に
- 情報利用の実態を知る指標として注目される

ログ分析のメリット・デメリット

- メリット
 - 自動的に収集できる
 - 人間の意志が介在しない
 - Cf) インタビュー調査／質問紙調査
 - サンプリングの必要性がない
 - 利用の全数が記録に残っている

(出典 : Galyani Moghaddam, Golnessa; Moballeghi, Mostafa . How Do We Measure Use of Scientific Journals?: A Note on Research Methodologies. Scientometrics. 2008, vol.76, no.1, p. 125-133.)^[1]

ログ分析のメリット・デメリット

- デメリット

- ユーザを特定できない

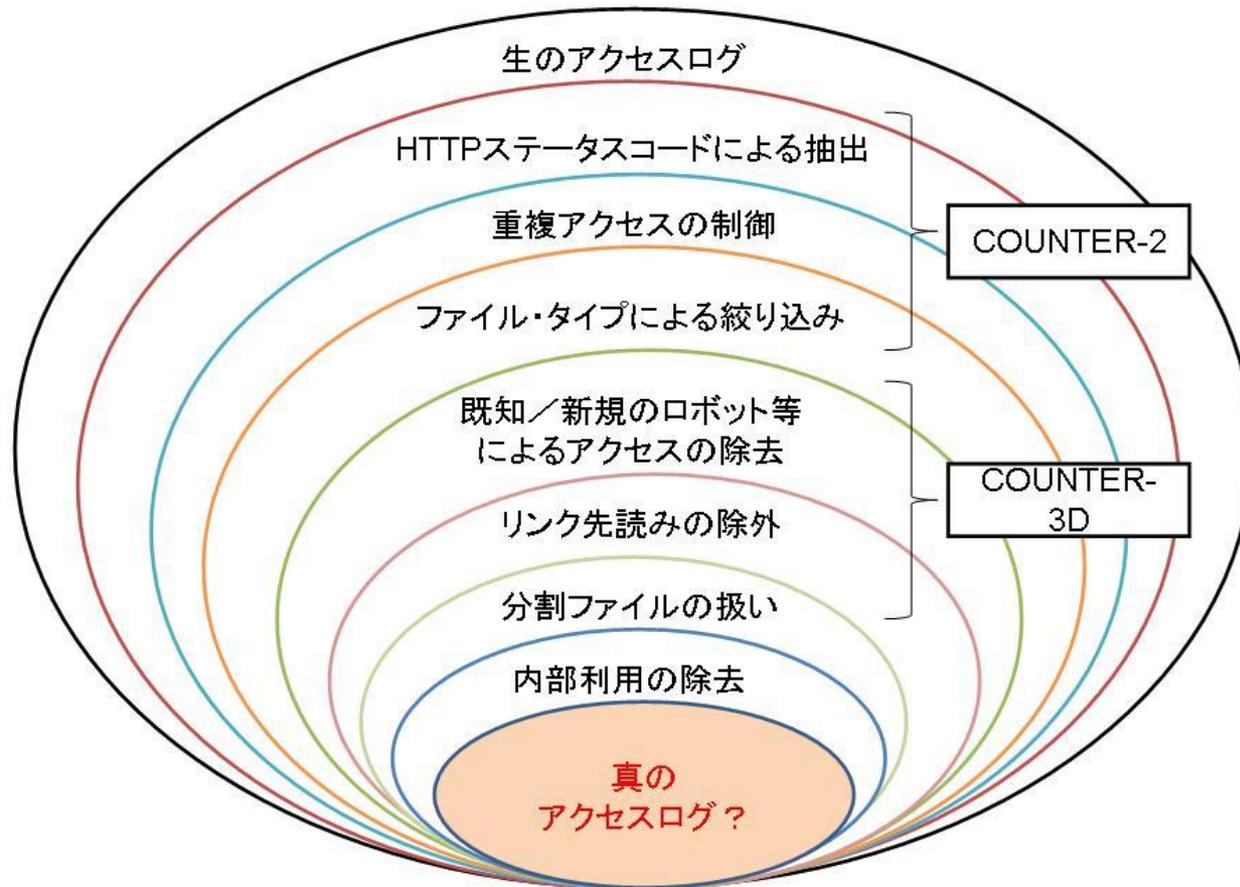
- 同じパソコンを違う人が使う
 - 違うパソコンを同じ人が使う
 - 図書館では特に起こりうること??

- ユーザパフォーマンスとシステムパフォーマンスの区別ができない

- フィルタリングの必要性

(出典: Galyani Moghaddam, Golnessa; Moballeghi, Mostafa . How Do We Measure Use of Scientific Journals?: A Note on Research Methodologies. Scientometrics. 2008, vol.76, no.1, p. 125-133.)^[1]

フィルタリングの例



(出典: 佐藤義則. 動向レビュー: 機関リポジトリの利用統計のゆくえ. カレントアウェアネス. 2008, (296), p.12-16. (<http://current.ndl.go.jp/ca1666>, 2008-11-05参照))^[2]

機関リポジトリのログ分析プロジェクト

- IRSプロジェクト(サウザンプトン大学)^[3]
- JISCによるPIRUSプロジェクト^[4]
- ROATプロジェクト(千葉大学等)^[5]

⇒分析環境の整備

機関リポジトリの利用状況

- 注目されるのは・・・
 - アクセス元: **誰が?**
 - アクセス方法: **どうやって?**
 - コンテンツ属性: **何を?**
- なぜこの3点?
 - **機関リポジトリの果たしている役割**の検討
 - 各リポジトリ**運用**、機関リポジトリ**政策**の検討とフィードバック

先行研究

- アクセス元：
 - 日本のリポジトリは国内からが主 (75%)^[6]
 - 民間プロバイダ / 大学内からのアクセスが多い
- アクセス方法^{[7] [8] [9] [10] [11]}：
 - サーチエンジンが主
 - CiNiiからのアクセスも多い
- コンテンツ属性^{[12][13]}：
 - ポストプリント > プレプリント
 - そのリポジトリにしかないコンテンツは利用が多い

先行研究の限界・問題

- アクセス元とアクセス方法、コンテンツ属性等を組み合わせて見た分析が行われていない
- ファイルの形式(テキストデータの有無／セキュリティ設定)の影響を考慮していない

ファイル形式等

- テキストデータ付与の有無
 - 内部に透明テキストを持つPDFか、画像PDFか
- セキュリティ設定
 - 「内容のコピー」を許可しているか？
- ボーンデジタル／OCR
 - 最初からデジタルで作られたものか、紙からスキャンしたものか

アクセス元、アクセス方法、 コンテンツ属性の関係分析

- 目的1: 利用数とアクセス元、アクセス方法、コンテンツ属性(文献種別、記述言語等)の関係を明らかにする
- 目的2: ファイル形式等のアクセス数への影響を明らかにする

⇒ 機関リポジトリの果たす役割とは？

本日の内容

1. はじめに: ZSプロジェクトの紹介
2. アクセスログ分析とは？
3. 8つの機関リポジトリのログ分析結果
4. 学術情報流通における機関リポジトリの役割と問題
5. おわりに

4つのリポジトリのログ分析(2008)

- 分析対象とするアクセスログ・・・
 - 北海道大学: HUSCAP (コンテンツ数: 25,542件)
 - 京都大学: KURENAI (28,536件)
 - 筑波大学: Tulips-R (7,899件)
 - アジア経済研究所: ARRIDE (640件)
- 分析期間: 2008年(1年間)
- 総ダウンロード回数: 1,150,813回

環日本海リポジトリの分析

- 分析対象(分析順)
 - 金沢大学: **KURA**(分析コンテンツ数: 16,233件)
 - 福井県地域共同リポジトリ: **CRF**(2,171件)
 - 学術成果発信システムやまがた: **YOU**(2,479件)
 - 九州大学: **QIR**(12,166件)
 - 北海道大学: **HUSCAP**(29,992件)
- 分析期間
 - KURA: 2009.10.19-2009.12.31
 - CRF: 2009.4-2009.12
 - その他: 2009年1年間

フィルタリング結果について

- 生のアクセスの**大部分はロボット**
 - 悪いことではない(可視性向上)
 - ボットは随時増える(ex: 佐藤翔ほか研究者)
- **重複アクセスも多い**
 - 除外しきれないボット？
 - 悪戯？
- フィルタリング後に残るのは**2～3割**

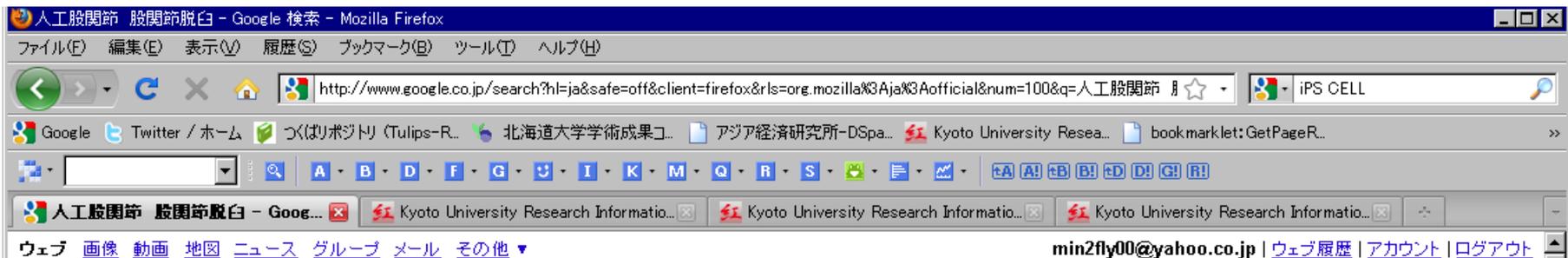
アクセス元(機関)

- 全リポジトリで最も多いのは**民間・プロバイダ**
 - 詳細は不明(学生? 研究者? 民間人?)
- **大学・企業**からのアクセスも多い
 - あわせて3割前後
- リポジトリの**利用者層は多様**
 - **研究者だけのものではない!!**

アクセス元（国内／海外）

- 利用者の大半（6-8割）は国内からアクセス
 - 日本語コンテンツが多いから
- 英語は海外、日本語は国内からのアクセスが多い
 - 逆は少ない：**日本人は英語論文を読まない？**

サーチエンジンからのアクセス例



Google 人工股関節 股関節脱臼 検索 検索オプション 表示設定

ウェブ全体から検索 日本語のページを検索

ウェブ [検索ツールを表示](#)

人工股関節 股関節脱臼 に一致する日本語のページ 約 16,700 件中 1 - 100 件目 (0.74 秒)

[Kyoto University Research Information Repository: 人工股関節全置換 ...](#)

Title: 人工股関節全置換術(THR)術後の股関節脱臼について. Authors: 笠原, 勝幸 researcher_resolver 小田, 裕造 researcher_resolver 飯田, 寛和 researcher_resolver 中村, 孝志 researcher_resolver. Issue Date: 1999 ...

[repository.kulib.kyoto-u.ac.jp/dspace/handle/.../49416 - キャッシュ - 類似ページ](#)

笠原勝幸 著 - [関連記事](#)

[PDF] [人工股関節全置換術\(THR\)術後の股関節脱臼について](#)

ファイルタイプ: PDF/Adobe Acrobat - [HTMLバージョン](#)

我々は他、病院にて人工股関節全置換術を受けた後に左側、股関節を、10. 回以上習慣的に脱臼をするようにな ... め、左股関節後方進入路で約、20cm. の皮切を行。った。大腎筋を翻転して人工関節を認め、股関節を内旋すると後方へ脱臼を生じた。ソケット ...

[repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/.../19_1.pdf - 類似ページ](#)

笠原勝幸 著 - [関連記事](#)

[repository.kulib.kyoto-u.ac.jp からの検索結果](#)

[人工股関節全置換術](#) | [人工関節全置換術](#) | [人工関節の広場](#)

人工股関節は、関節の動きによって脱臼(だっきゅう)する場合があります。股関節の周辺の筋力が低下すると、関節が安定せず、人工股関節の脱臼が、前方(ひざのほう)脱臼と後方(こ

直接PDFファイル
をクリック

リポジトリ内からのアクセス例

Kyoto University Research Information Repository: 人工股関節全置換術(THR)術後の股関節脱臼について - Mozilla Firefox

ファイル(E) 編集(E) 表示(V) 履歴(S) ブックマーク(B) ツール(T) ヘルプ(H)

http://repository.kulib.kyoto-u.ac.jp/dspace/handle/2433/49416

Google Twitter / ホーム つばしリポジリ (Tulips-R... 北海道大学学術成果コ... アジア経済研究所-DSpa... Kyoto University Resea... bookmarklet:GetPageR...

Kyoto University Research Inf... Kyoto University Research Informatio... Kyoto University Research Informatio...

京都大学学術情報リポジトリ
KURENAI: Kyoto University Research Information Repository

リポジトリ検索
検索
詳細検索
ホーム

ブラウズ
コミュニティ & コレクション
タイトル
著者
主題
日付

Kyoto University Research Information Repository >
060 医学研究科・医学部・医療技術短期大学部 = Graduate School of Medicine >
京都大学医療技術短期大学部紀要 = Annual reports of the College of Medical Technology, Kyoto University >
第19号 >

このアイテムの引用には次の識別子を使用してください: <http://hdl.handle.net/2433/49416>

フルテキストリンク:

ファイル	記述	サイズ	フォーマット
19_1.pdf		664.26 kB	Adobe PDF 見る/開く

タイトル: 人工股関節全置換術(THR)術後の股関節脱臼について

著者: 笠原, 勝幸 [NII 研究者リゾルバー](#)
小田, 裕造 [NII 研究者リゾルバー](#)
飯田, 寛和 [NII 研究者リゾルバー](#)
中村, 孝志 [NII 研究者リゾルバー](#)

発行日: 1999

出版者: 京都大学医療技術短期大学部

誌名: 京都大学医療技術短期大学部紀要

このページからクリック

著者や掲載紙の詳細情報

著者: 笠原, 勝幸 [NII 研究者リゾルバー](#)
小田, 裕造 [NII 研究者リゾルバー](#)
飯田, 寛和 [NII 研究者リゾルバー](#)
中村, 孝志 [NII 研究者リゾルバー](#)

発行日: 1999

出版者: 京都大学医療技術短期大学部

誌名: 京都大学医療技術短期大学部紀要

・CiNii等のリンク先は
メタデータページが多い
⇒リポジトリ内からの
アクセス増に影響

The screenshot shows the CiNii search results page. The search query is '人工股関節 股関節脱臼'. The top result is '人工股関節全置換術(THR)術後の股関節脱臼について' by 笠原 勝幸, 小田 裕造, and 飯田 寛和. Below the title, there are several links for full-text access: 'Webcat Plus 刊行物-所蔵情報', 'JFree FullText | 機関リポジトリ本文(無料)', 'Tulips-L(SFX)', and 'OPAC Tulips-OPAC'. A pink box highlights the 'JFree FullText' link, with a pink arrow pointing to the right.

The screenshot shows the KURENAI page for the article. The URL is 'http://repository.kulib.kyoto-u.ac.jp/dspace/handle/2433/49416'. The page title is '京都大学学術情報リポジトリ KURENAI: Kyoto University Research Information Repository'. The article title is '人工股関節全置換術(THR)術後の股関節脱臼について'. The authors listed are 笠原 勝幸, 小田 裕造, 飯田 寛和, and 中村 孝志. The publication year is 1999, and the publisher is '京都大学医療技術短期大学部'. The page also includes a table with file information and a list of links.

フルテキストリンク:			
ファイル	記述	サイズ	フォーマット
19_1.pdf		664.26 kB	Adobe PDF 見込/開く

タイトル: 人工股関節全置換術(THR)術後の股関節脱臼について
著者: 笠原 勝幸 [ORCID](#) [研究リポジトリ](#)
小田 裕造 [ORCID](#) [研究リポジトリ](#)
飯田 寛和 [ORCID](#) [研究リポジトリ](#)
中村 孝志 [ORCID](#) [研究リポジトリ](#)
発行日: 1999
出版者: 京都大学医療技術短期大学部
誌名: 京都大学医療技術短期大学部紀要
巻: 19
開始ページ: 1
終了ページ: 9
URI: <http://hdl.handle.net/2433/49416>
出現コレクション: [第19号](#)

アクセス方法

- **サーチエンジンからのアクセスが大半(5-8割)**
 - 直接本文へ。メタデータは見ない？
 - カバーページが重要
- **サーチエンジンからのアクセスは偏る？**
 - **テキストデータの有無が影響**
- **言語との関係**
 - **英語はサーチエンジン偏重**
 - **CiNiiのような外部サービスがないせい？**

文献タイプ

- **学位論文**はどこでも人気コンテンツ
 - **本文あり、テキストデータあり**が条件
- 他に多いのは**図書、教材**
- その他はリポジトリにより傾向の差
 - テキストデータの有無の影響も??

アクセス数上位文献

- 全体を見た場合とは異なる傾向
- 強いのは医療・保健・看護・介護文献
 - 世の中の需要を反映？
- 共同リポジトリはホスト機関以外のコンテンツへも多くのアクセス
 - 共同リポジトリの意義は大きい！

出版年との関係

- **新しい論文**を良く使う
 - 国内から／大学からのアクセス
 - リポジトリ内からのアクセス
 - **日本語**論文へのアクセス
- **出版年に関わりなく**使う
 - 海外から／企業からのアクセス
 - **サーチエンジン**からのアクセス
 - **英語**へのアクセス

ファイル形式等との関係

- テキストデータがある方がアクセス数が多い
 - サーチエンジンからのアクセスが増えるため
 - ボーンデジタル／OCRの影響は少ない
- 「内容のコピー」が不可だとアクセスは減る
 - サーチエンジンから探しにくくなるため

本日の内容

1. はじめに: ZSプロジェクトの紹介
2. アクセスログ分析とは？
3. 8つの機関リポジトリのログ分析結果
4. 学術情報流通における機関リポジトリの役割と問題
5. おわりに

機関リポジトリの役割

- 著者によるセルフアーカイビング
- 紀要の出版・流通
- 学位論文電子公開
- 機関の可視性向上・広報
- 納税者への説明責任

著者によるセルフアーカイブ

- 実現できている？
 - 大学だけでなく企業、プロバイダからも利用
 - 低所得国からも利用
- 阻害要因
 - 雑誌論文の登録数が少ない（今後に期待？）

紀要の出版・流通

- 実現できている？
 - テキストがあれば**紀要論文はアクセス数が多い**
 - 研究者の**ブラウジングと同様の利用形態**
- 実現できていない？
 - **英語／海外からのアクセスはブラウジングしない**
 - **テキストデータがないとアクセス数激減**
- 阻害要因
 - **海外向けの入口の不在**
 - **テキストデータの未整備**

学位論文電子公開

- 実現できている？
 - 学位論文はアクセス数が顕著に多い
- 実現できていない？
 - サーチエンジン以外からのアクセスは少ない
 - テキストデータがないとアクセス数激減
- 阻害要因
 - サーチエンジン以外の入口の不在
 - テキストデータの未整備

機関の可視性向上・広報

- 実現できていない？
 - 利用者は直接本文にアクセス
 - 機関名や著者名等のメタデータを読まない
- 阻害要因
 - サーチエンジン以外の入口の不在
 - 本文ファイルへのメタデータ不整備

納税者への説明責任

- 実現できている？
 - 日本国内からのアクセス数が多い
 - 大学以外(企業、民間等)のアクセス数が多い
- 実現できていない？
 - 日本人がアクセスするのは日本語論文だけ
 - 自然科学の英語論文が読まれない
- 阻害要因
 - 書かれた言語の問題

全体に共通の傾向

- 機関リポジトリは「公開」の場
 - 「探索」や「発見」の場ではない
- 外部サービスがないとアクセスされない
 - 適切なサービスがないとサーチエンジン集中
 - テキストデータがないとアクセス激減
 - テキストさえあればOCRでも問題なし
- 日本語・英語の壁
 - 言語によって利用者が異なる
 - それに見合ったサービスが必要

結論

機関リポジトリが果たしうる役割は多様で、部分的に果たせているものも多いが...

- 発見手段となる外部サービスとの連携の不在
- 言語の壁
- テキストデータの不備

がそれを邪魔している

結論

機関リポジトリが期待される役割を果たせるものになるには...

- 利用者とコンテンツに応じた外部サービス連携
- OCR等によるテキストデータの整備

が必要

具体的に
何をすれば
いいのか？

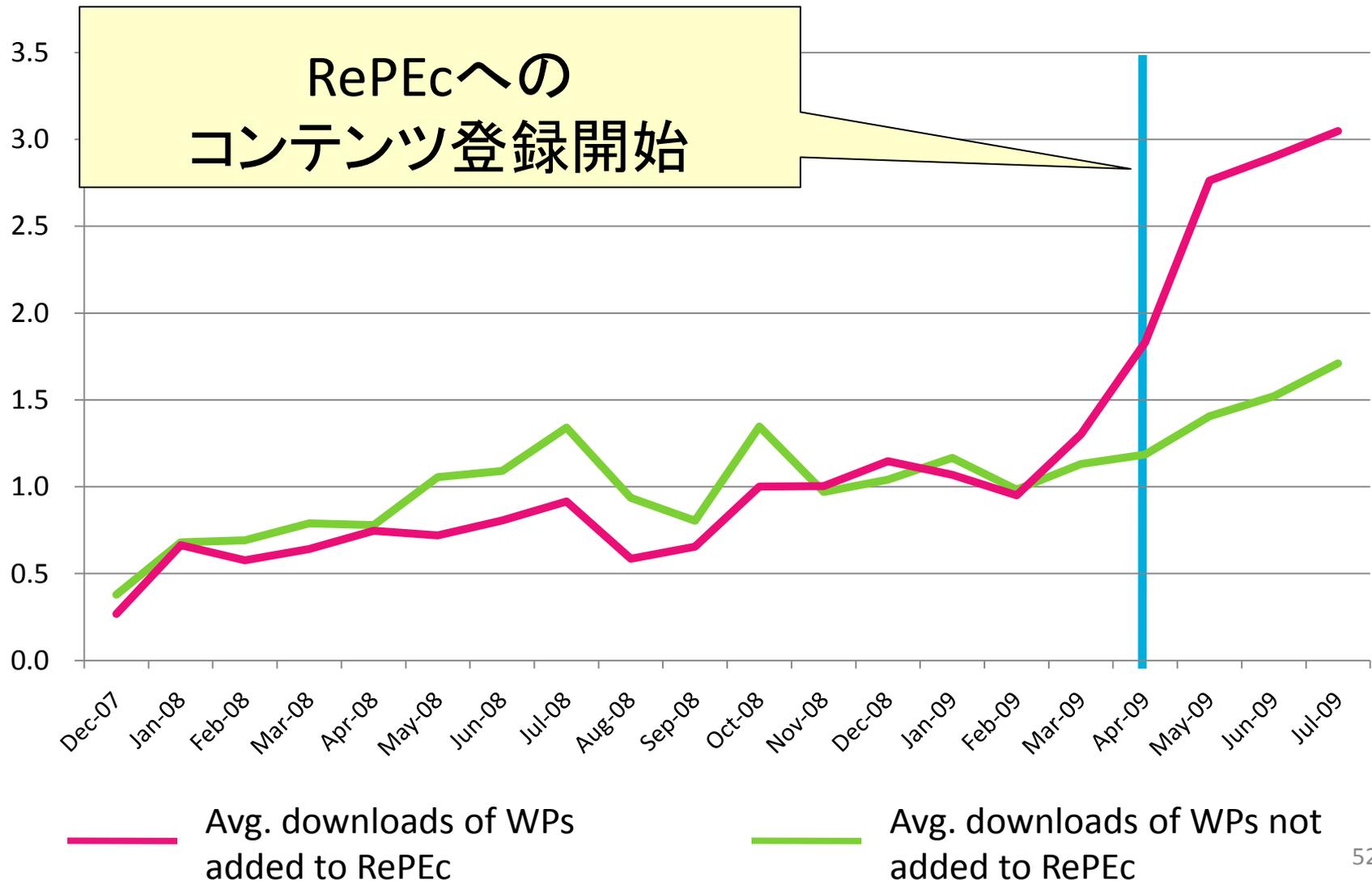
(1) OCRの徹底

(2) 外部サービス との連携

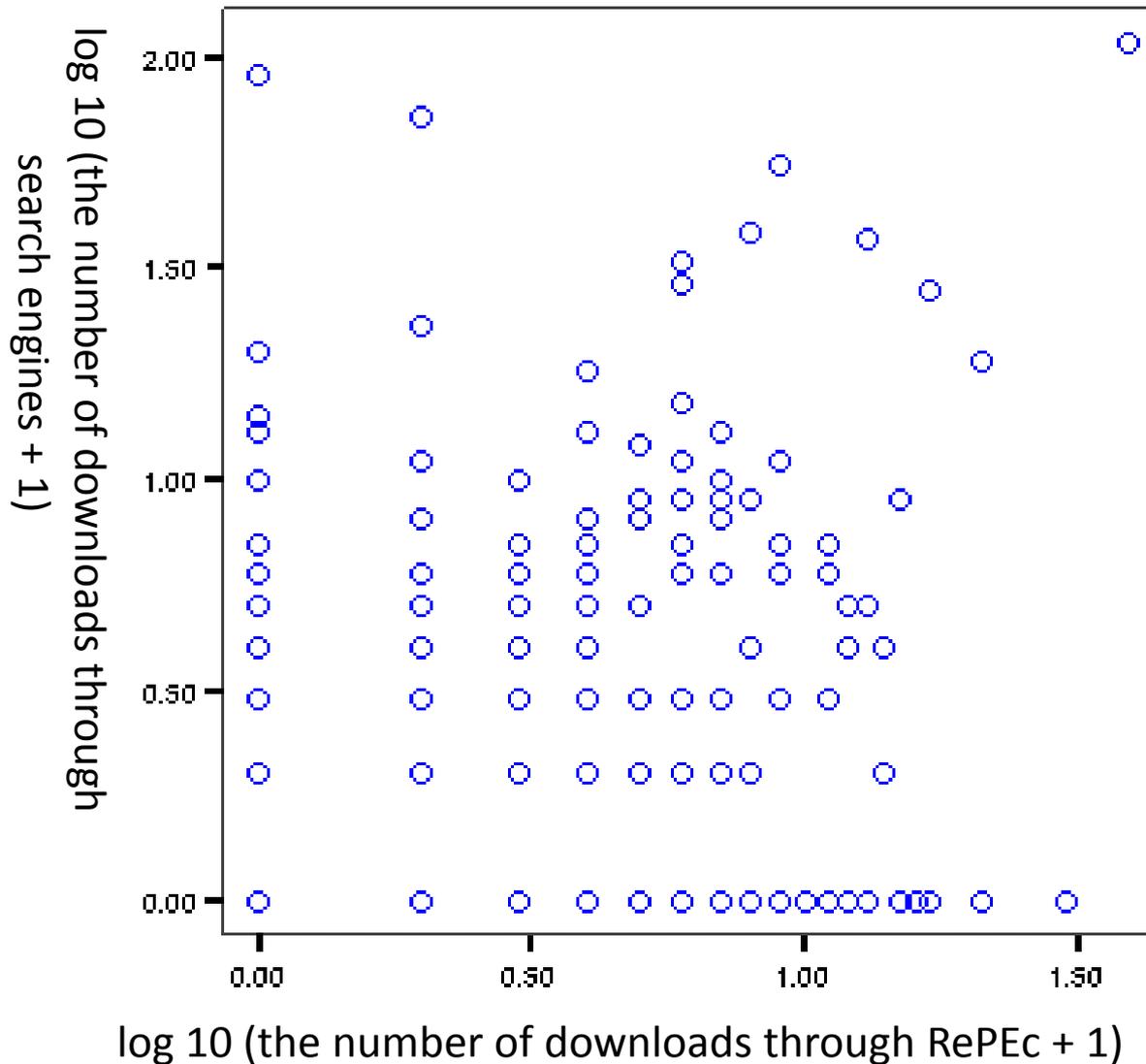
RePEcとは？

- Research Papers in Economics (<http://repec.org>)
- 経済学分野の主題リポジトリ
- 24万件以上のワーキングペーパーを収録
- アジア経済研究所、一橋大学など登録機関多数
- 本文は機関リポジトリ等に、RePEcはメタデータを収集してリンクを貼る

一橋大学 (HERMES-IR) での RePEc登録の例

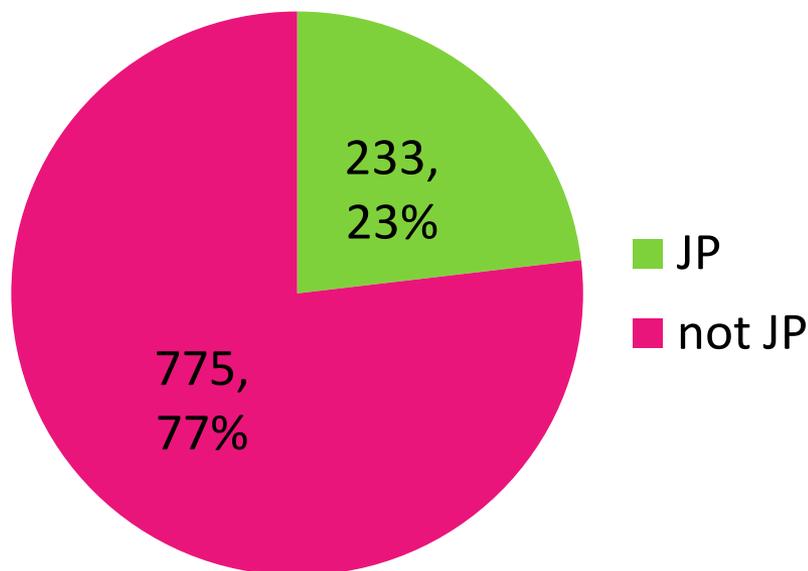


RePEcからのアクセスと サーチエンジンからのアクセスの関係



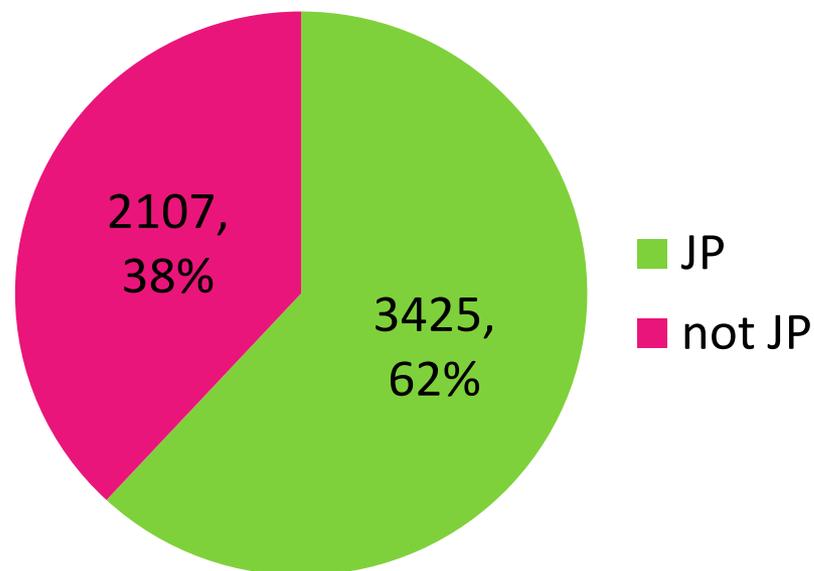
国内/海外からのアクセス数 (HERMES-IR)

アクセス元の国内／海外
(RePEcからのアクセス)



Total: 1,008 ダウンロード

アクセス元の国内／海外
(RePEc以外からのアクセス)



Total: 5,532 ダウンロード

その他の特徴的な参照元

- その他の学術系サービス
- JAIRO / CiNii
 - NII提供サービスは重要
- Wikipedia
 - Wikipediaはメタデータへのアクセスも多い
- Yahoo! 知恵袋ほか Q&Aサイト
- 大学広報ページ / ニュースサイト
 - 積極的に広報すると◎

もう一度...機関リポジトリの役割

- 著者によるセルフアーカイビング
- 紀要の出版・流通
- 学位論文電子公開
- 機関の可視性向上・広報
- 納税者への説明責任

実現できるかは
運営担当者によ
って決まる！

- 株
-
- 納税

利用者を迎えに行こう！

- “If you built it, he will come” (『フィールド・オブ・ドリームス』より) はリポジトリには通じない
 - (1) 作っただけではコンテンツは来ない
 - (2) コンテンツがあるだけでは利用者は来ない

本日の内容

1. はじめに: ZSプロジェクトの紹介
2. アクセスログ分析とは？
3. 8つの機関リポジトリのログ分析結果
4. 学術情報流通における機関リポジトリの役割と問題
5. おわりに

アクセスログ分析の効用

- リポジトリの利用状況、アクセスの阻害要因は**詳細なログ分析**によって明らかにできる
- **ぜひこれからもよろしくお願いしますm(__)m**

アクセスログ分析の限界

- ログからわかるのは「利用」の諸側面の**一部**
 - 「**投稿者**」の利用は？
 - ダウンロードした人は**本当に読んでいる？**
 - 読んだ後の**活用状況は？**
- **別の観点からの研究も必要**
 - Ex)被引用数との関係分析(**ZSプロジェクト**)

研究 ⇔ 実践の
協同でより良い
学術情報流通を！

Acknowledgements (1)

- 本発表は「科学研究費補助金(基盤研究(C)機関リポジトリへの登録が学術文献流通に及ぼす効果についての定量的分析」および国立情報学研究所次世代学術コンテンツ基盤共同構築事業委託事業(領域2)「機関リポジトリへの登録が学術文献流通に対して及ぼす効果についての定量的解析のための文献蓄積及びデータ整理」による支援を受けた研究に基づいています。
- ロボットリストの一部は東北学院大学の佐藤義則先生(ROATプロジェクト)からご提供いただきました。

Acknowledgements (2)

- ログを提供いただいた機関/共同リポジトリ
 - 北海道大学学術成果コレクション
 - 京都大学学術情報リポジトリ
 - つくばリポジトリ
 - アジア経済研究所学術研究リポジトリ
 - 一橋大学機関リポジトリ
 - 金沢大学学術情報リポジトリ
 - 福井県地域共同リポジトリ
 - 九州大学学術情報リポジトリ
 - 学術成果発信システムやまがた ゆうキャンパスリポジトリ