

June 12, 2008

平成19年度CSI委託事業報告交流会

# 機関リポジトリの アウトプット分析

東北学院大学 佐藤 義則

# IRの評価手法

- **インプット(投入資源)**
  - 種類ごとのコンテンツ数
  - 費用(金額、人員、時間)
  - **コンテンツの投入状況**
    - Carr, Leslie et al. "Size isn't everything." *D-Lib Magazine*, 2007, 13(7/8)
  - **著者別の論文寄託件数(構成員の参加のパターン)**
    - Thomas, Chuck et al. "Measuring and comparing participation patterns in digital repositories." *D-Lib Magazine*, 2007, 13(9/10)

# アウトプット(産出)への注目

- 意義

- 有用性(有効性)を示す指標として

- (計画および機関の)アカウンタビリティ

- 現状分析および今後の(改善)計画のため

- ベンチマーキングのツールとして

- プロモーション(アドヴォカシー)における活用

“「協調問題」と呼ばれる状況では、他の人々も同じように参加する限りにおいて、自分もある共同的行動に参加したいと思っている。協調する一つの方法は、単純に「さあ、みんなで参加しよう」というようなメッセージを伝達することである。しかし、他の人々が参加する限りにおいて自分も参加するのであるから、そのメッセージが成功するためには、皆がそのメッセージについて知っているだけでなく、それぞれが、他の人々もそのメッセージについて知っているということを知っていなければならない。 実際、他の人々がそのメッセージについて知っていることを、他の人々が知っている……ことを知っていなければならない。

マイケル・Y・チウエ著、安田雪訳『儀式は何の役に立つか：ゲーム理論のレッスン』新曜社，2003，p. 8.



# アウトプット分析のツール

- ログ分析

- AWStats

- WebAlyzer

- Analog

- ...

- ビーコン方式

- Google Analytics

生のアクセスログ

HTTPステータスコードによる抽出

重複アクセスの制御

ファイル・タイプによる絞り込み

既知／新規のロボット等  
によるアクセスの除去

リンク先読みの除外

分割ファイルの扱い

内部利用の除去

真の  
アクセスログ？

COUNTER-2

COUNTER-3D

# データ

- 依頼： 14大学宛
- データ受取： 11大学分
- 分析対象： 9大学における平成19年7月1日から12月31日のログデータ(1大学については7月7日以降)； 提供ログデータの対象期間が3か月未満の2大学については、比較分析の対象から除外

# 9つの機関リポジトリの インプット・アウトプット

IR	教員数 (2007.5)	コンテンツ数 (2008.1)	訪問者の 純数	訪問者の 総数	HTMLまたは CGIファイル へのアクセス 数	PDFファイル への アクセス数	総ファイル アクセス数	トラフィッ ク量 (GB)
A	2,082	21,502	176,223	261,848	871,252	187,309	3,389,068	212
B	1,270	20,115	56,668	76,477	216,532	66,614	377,295	76
C	1,032	13,059	3,842	6,713	96,282	3,124	138,632	12
D	1,886	6,612	63,486	88,943	258,430	61,330	476,886	128
E	727	5,445	20,786	25,708	133,654	18,679	239,590	28
F	2,869	22,746	89,940	124,731	409,093	92,254	2,229,193	101
G	2,619	6,922	7,666	10,109	221,339	1,687	412,062	10
H	1,832	12,640	58,810	76,386	175,806	37,313	1,414,242	49
I	2,274	6,202	102,360	135,077	323,979	126,575	630,360	172

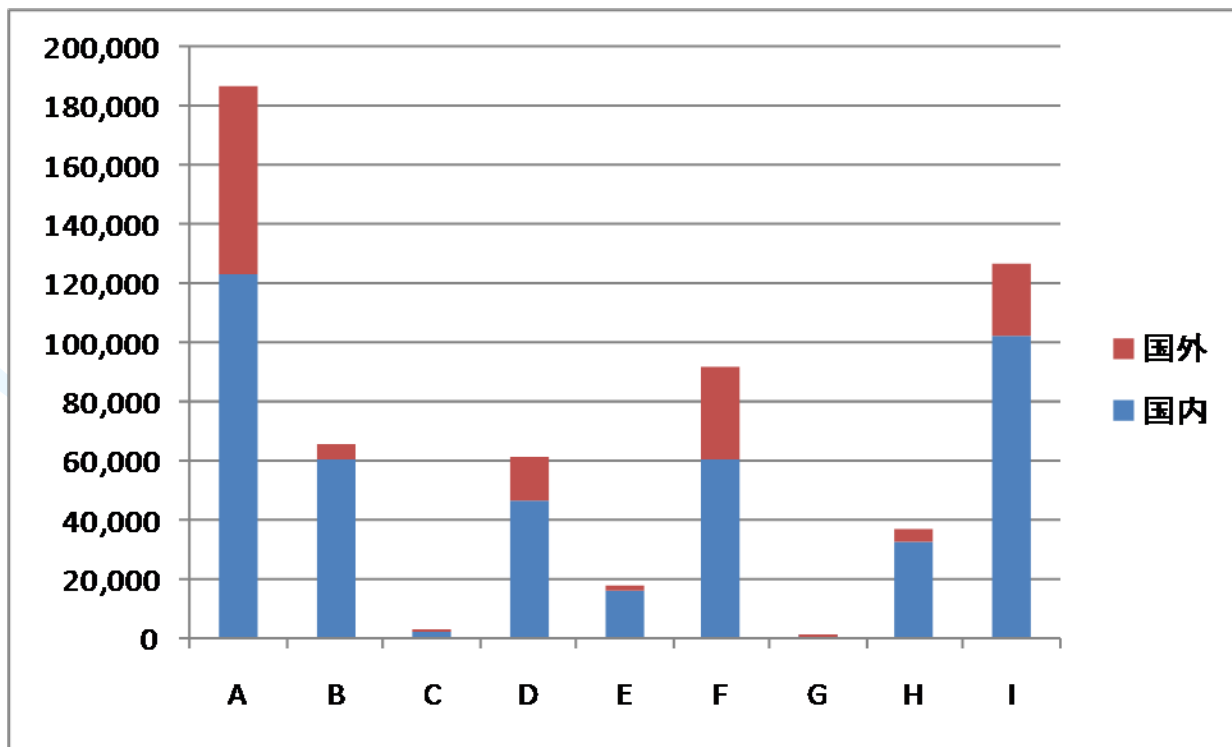
集計期間: 2007年7月1日~12月31日

# 9つの機関リポジトリの単位あたり利用状況

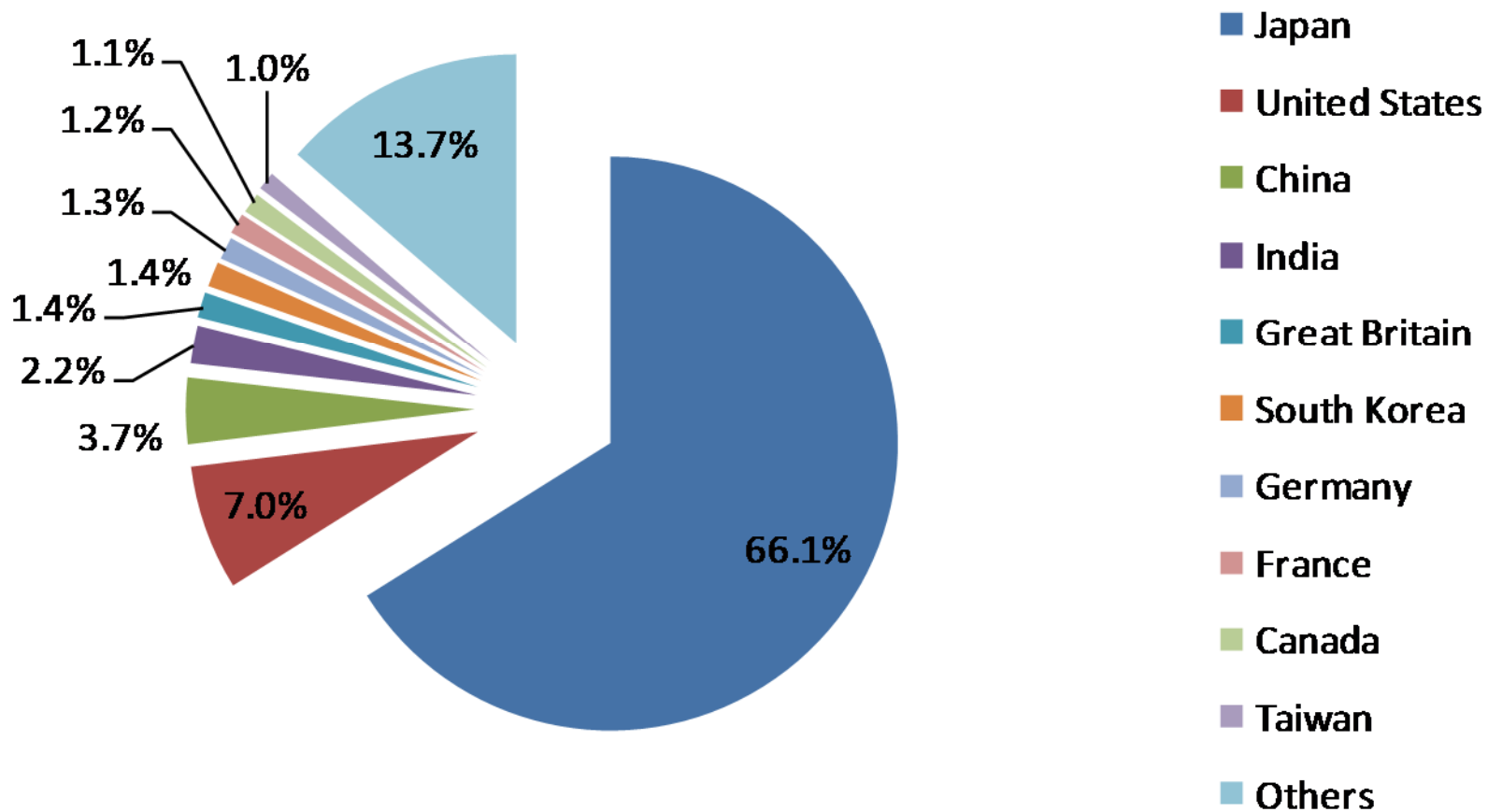
IR	教員一人あたりのHTML/CGIアクセス回数	教員一人あたりのPDFアクセス回数	コンテンツあたりのHTML/CGIアクセス回数	コンテンツあたりのPDFアクセス回数
A	418.47	89.97	40.52	8.71
B	170.50	52.45	10.76	3.31
C	93.30	3.03	7.37	0.24
D	137.03	32.52	39.08	9.28
E	183.84	25.69	24.55	3.43
F	142.59	32.16	17.99	4.06
G	84.51	0.64	31.98	0.24
H	95.96	20.37	13.91	2.95
I	142.47	55.66	52.24	20.41

# アクセス元の国内、国外の比率

	A	B	C	D	E	F	G	H	I	合計
国内	123,787	60,713	2,940	47,134	16,293	60,864	1,534	33,033	102,274	448,572
	66.1%	92.3%	94.1%	76.9%	87.2%	66.0%	90.9%	88.5%	80.8%	75.5%
国外	63,522	5,053	184	14,196	2,386	31,390	153	4,280	24,301	145,465
	33.9%	7.7%	5.9%	23.1%	12.8%	34.0%	9.1%	11.5%	19.2%	24.5%
総アクセス数	187,309	65,766	3,124	61,330	18,679	92,254	1,687	37,313	126,575	594,037



# アクセス国別 (A大学)



# A大学のIRに対する 国内のアクセス元別回数

		国内での比率	全体での比率
大学内	41,396	33.4%	22.1%
*.ac.jp (学外)	17,541	14.2%	9.4%
*.go.jp	2,037	1.6%	1.1%
*.pref.jp	1,502	1.2%	0.8%
*.co.jp	7,537	6.1%	4.0%
*.ne.jp	48,946	39.5%	26.1%
others	4,828	3.9%	2.6%
小計	123,787		

# 9つの機関リポジトリごとの アクセス経路(リファラ)

サイトへの接続元	A		B		C		D		E	
	アクセス数	%	アクセス数	%	アクセス数	%	アクセス数	%	アクセス数	%
アドレスの直接指定 / ブックマーク	20,958	11.10%	8,110	12.10%	160	5.10%	7,402	12%	1,462	7.80%
インターネット検索エンジンから	94,655	50.50%	47,348	71%	2	0%	37,528	61.10%	10,897	58.30%
Google	<b>82,839</b>		<b>17,316</b>		<b>1</b>		<b>32,952</b>		<b>9,977</b>	
Yahoo.com	<b>6,660</b>		<b>27,510</b>				<b>2,629</b>		<b>68</b>	
その他	<b>5,156</b>		<b>2,522</b>		<b>1</b>		<b>1,947</b>		<b>852</b>	
外部ページから (IR内のメタデータページを含む)	71,595	38.20%	11,091	16.60%	2,954	94.50%	16,312	26.50%	6,281	33.60%
不明	101	0%	65	0%	8	0.20%	88	0.10%	39	0.20%

サイトへの接続元	F		G		H		I	
	アクセス数	%	アクセス数	%	アクセス数	%	アクセス数	%
アドレスの直接指定 / ブックマーク	10,459	11.30%	207	12.20%	4,070	10.90%	9,396	7.40%
インターネット検索エンジンから	54,557	59.10%	66	3.90%	17,887	47.90%	102,179	80.70%
Google	<b>47,851</b>		<b>40</b>		<b>10,378</b>		<b>96,313</b>	
Yahoo.com	<b>2,041</b>		<b>23</b>		<b>6,237</b>		<b>533</b>	
その他	<b>4,665</b>		<b>3</b>		<b>1,272</b>		<b>5,333</b>	
外部ページから (IR内のメタデータページを含む)	27,157	29.40%	1,414	83.80%	15,312	41%	14,977	11.80%
不明	81	0%	0	0%	41	0.10%	23	0%



# 利用頻度の高いコンテンツ

- 上位50件のうち
  - 紀要論文24件
  - 教材8件
  - 図書6件
  - 雑誌論文5件
  - 博士論文3件
  - 講演資料2件

機関リポジトリのコンテンツの多様性を反映

	回数	タイプ	件数	資料名	掲載誌等
1	3,546	雑誌論文	1	LAYER GROWTH OF CdSb PHASE IN THE Cd-Sb DIFFUSION COUPLE AT HIGH PRESSURE	The Review of Physical Chemistry of Japan Vol. 48 No. 2 (1978)
2	2,827	教材	1	解剖学学習指針	
3	2,433	紀要論文	1	スウェーデン福祉政策の動向と展望	公共研究 Vol1 no.2 page.5-16 (20050329)
4	2,025	教材	1	統合・基礎神経学 - 神経系の構造を中心に	
5	1,793	紀要論文	1	高齢者ケアの日本・スウェーデン比較	公共研究 Vol1 no.2 page.25-31 (20050329)
6	1,444	教材	1	力学入門(講義ノート)	
7	1,283	教材	1	基本顕微構造解析法 I-II	
8	1,193	展示目録	1	読み継がれる『源氏物語』-広島大学図書館収蔵物語コレクション	
9	1,050	雑誌論文	1	自衛隊生徒の発足 -1955年の少年兵-	日本の教育史学, 45: 162-180
10	972	図書	1	磁気モーメントとg値	漁火書店「疑問の発生シリーズ」(1981年創刊)のPDF版
11	949	紀要論文	1	要介護高齢者の口腔のケアから考える保健・医療と介護・福祉政策の連携	公共研究 Vol2 no.1 page.226-259 (20050628)
12	902	図書	1	化学ポテンシャルと平衡定数	漁火書店「疑問の発生シリーズ」(1981年創刊)のPDF版
13	873	紀要論文	1	公共哲学とは何か	公共研究 Vol1 no.1 page.29-46 (20041227)
14	843	図書	1	衝突頻度と平均自由行程	漁火書店「疑問の発生シリーズ」(1981年創刊)のPDF版
15	795	紀要論文	1	高校生球児の野球留学とキャリア形成の諸課題	『生涯学習・キャリア教育研究』第2号(2006年) 39~44頁
16	783	図書	1	電磁気学における単位系	漁火書店「疑問の発生シリーズ」(1981年創刊)のPDF版
17	770	紀要論文	1	「行政国家」から考える公共性論	公共研究 Vol2 no.1 page.152-191 (20050628)
18	754	紀要論文	1	グローバリゼーションと福祉国家の後退: アメリカの場合	公共研究 Vol1 no.2 page.99-118 (20050329)
19	743	紀要論文	1	中国の経済発展と社会保障: アジアの社会保障と国際協力の展	公共研究 Vol3 no.1 page.254-261 (20060628)
20	736	紀要論文	1	熱帯の木材	名古屋大学博物館報告 No. 21, 183-237, 2005
21	727	紀要論文	1	アダム・スミスの道徳哲学と政治論: フランシス・ハチスンを媒介と	公共研究 Vol2 no.2 page.177-237 (20050928)
22	719	図書	1	歳差運動の物理学	漁火書店「疑問の発生シリーズ」(1981年創刊)のPDF版
23	702	紀要論文	1	中国における社会変動と社会保障制度改革	公共研究 Vol2 no.2 page.5-19 (20050928)
24	697	紀要論文	1	ビルマ(ミャンマー)の開発と人権・環境問題	公共研究 Vol2 no.1 page.6-22 (20050628)
25	660	教材	1	2004年度 グラフ理論講義ノート	
26	656	雑誌論文	1	Molecular epidemiology of adenoviral conjunctivitis in Hanoi, Vietnam.	American Journal of Ophthalmology, 142(6): 1064-1066, 2005
27	648	紀要論文	1	ヨーロッパ福祉国家改革の可能性	公共研究 Vol1 no.2 page.90-98 (20050329)
28	647	紀要論文	1	精神障害者の社会的自立について	公共研究 Vol3 no.1 page.232-240 (20060628)
29	629	教材	1	ジェンダーって何?(情報の道しるべ)	
30	617	教材	1	Quantum Mechanics (講義ノート) : 量子力学(講義ノート)	

# 今後に向けて

- 安定的に運用可能な分析方式の確立
  - 各大学が自主的に評価できる環境の創出
    - システム構築／監査方式による統計データの保証
  - メタデータとログデータのマッチング
    - そのためのリポジトリシステム要件の検討
  - さまざまなプラットフォームへの対応
    - IRStats for ePrints
    - DSpace, e-Repository, XooNips, Fedora...
  - 全国的な統計収集機能の検討