

平成20年度CSI報告交流会(コンテンツ系)  
2009年7月9日; 一橋講堂

# 機関リポジトリのコスト分析

東北学院大学

佐藤 義則

# 構成

1. 目的(問題の所在)
2. 先行研究におけるコストモデル
3. 現状分析(集計)
  - i. 集計結果
  - ii. 件数と金額のマッピング
  - iii. 事例調査
4. まとめと今後の課題

# 問題の所在

- \* 機関リポジトリの担当者が、機関リポジトリへの長期の支援を望むならば、コストに関するデータの収集と分析が不可欠である
- \* しかし、現状ではデータはきちんと整理されているとはいえない
  - \* コンテンツ投入の所要費用はどのくらいか？
  - \* 適正な費用とは？
  - \* 作業が必要で、実際にどれだけの費用、人員、及び所要時間を要しているのか

# 目的

- \* コレクションのタイプ別（学術雑誌論文（海外）、学術雑誌論文（国内）、学位論文、紀要論文、研究成果報告書、会議発表論文、会議発表用資料、その他）、および処理方式（機関内処理、外注の別）別に、コンテンツ投入にかかる費用および所要時間をできる限り明確にし、各機関がベンチマークとして活用できる情報の提供を目指す

# 作業の概要

- \* 平成20年9月から
- \* 作業内容；
  - \* コストモデルに関する文献調査
  - \* 平成17年度から平成20年度までの各機関からの成果報告の取りまとめ
  - \* 訪問調査(4大学)
  - \* 検索可能性に関するサンプル調査

# 先行研究におけるコストモデル

- \* Swan. Sustainability of repositories.
  - \* In A DRIVER's Guide to European Repositories
- \* Berlin6 Open Access Conf. 「オープンアクセスのコストと便益：ビジネスモデルと公共投資」セッション
- \* Blue Ribbon Task Force
  - \* Sustaining the Digital Investment
  - \* 文献レビュー
- \* LIFEプロジェクトのライフサイクルモデル
- \* BCL (Beagrie-Chruszcz-Lavoie) モデル
  - \* "Keeping Research Data Safe"

# Swan等の調査

- \* 2004年、英国、アイルランド、米国、カナダの機関リポジトリを対象 (Swan et al., 2004; Swan & Awre, 2006)
- \* 設置費と運営費
  - \* 設置費;
    - \* 組織内構築のリポジトリ 平均9千250ユーロ(約120万円; ハードウェアとソフトウェア含む)
    - \* サーバを機関内に置き構築を外部委託したリポジトリ  
平均7千ユーロ(約90万円)
    - \* 構築とホスティングの両方を外部委託したリポジトリ  
平均3万8千ユーロ(約490万円)

# SHERPAプロジェクトの実際に基づく機 関リポジトリの設置費と運営費

初期設置費用(€)	技術サポート/保守(€)	年間の運営費(€)	論文投入にかかるコスト(€)
ソフトウェア	0 高等教育機関において標準的なWebサービスの保守(3年更新)	スタッフ給与 51,000	論文投入にかかるコスト(€) 週当たりの時間 17.7
サーバ 設定 カスタマイズ	ハードウェア 5,100 人件費 1,020		1時間当たりの (投入)論文数 4
6,630 (約85万円)	6,120 (約79万円)	(約655万円)	7.58 (約970円)



# Piorun & Palmer (2008)

- \* マサチューセッツ大学医学・生物医学大学院における博士論文のデジタル化で要するコスト
- \* 外注と館内処理の場合の経費の比較
- \* コスト算定において、スキヤニング、品質管理、抄録のOCR処理、登録、プロジェクト管理、ハードウェア／ソフトウェア等に分割して積算する方法

# LIFEプロジェクト

- \* UCL (University College London) と英国図書館のプロジェクト； JISCによる助成
- \* 目標： デジタル資料の管理と保存コストの原価計算に関する方法論の整備
- \* フェーズ1 (McLeod et al., 2006)
  - \* UCLの電子ジャーナル、英国図書館のWebアーカイビング活動、英国図書館のVDEP (Voluntarily Deposited Electronic Publications) デジタルコレクションのデータの検証
- \* フェーズ2 (LIFE<sup>2</sup> ; Ayris et al., 2008)
  - \* モデルを洗練させ、SHERPA-DP (デジタル保存プロジェクト)、SHERPA-LEAP (London E-prints Access Project ; 共同機関リポジトリ)、英国図書館の新聞コレクションの三つの事例研究
- \* フェーズ3 (プロポーザル)

# LIFE<sup>2</sup>のコストモデル

$$L_T = C + Aq_T + I_T + BP_T + CP_T + AC_T$$

$L_T$ : ライフサイクル(「生涯」にわたって対象物を維持するための費用)

$C$ : 創造または購入(対象物の購入または創造に要する費用)

$Aq_T$ : 取得(選定、寄託契約、知的所有権およびライセンス、発注、書類のやり取り、調達、検収)

$I_T$ : 受入(品質管理、メタデータ、寄託、所蔵更新、参照リンク)

$BP_T$ : ビットストリーム保存(リポジトリの管理、記憶装置の準備、システム更新、バックアップ、検査)

$CP_T$ : コンテンツ保存(保存状態の観察、保存計画、保存活動、再受入、廃棄)

$AC_T$ : アクセス(アクセスの提供、アクセス管理、利用者支援)

# Beagrie, Chruszcz, & Lavoie (2008)

- JISCによる助成 “Keeping Research Data Safe.”
- 高等教育機関による研究データのデジタル保存に関する長期的なコストモデルの整備
- コストを、アーカイブ以前、アーカイブ、サポート・サービスの三つの主要段階に分けることでデジタル・オブジェクトのライフサイクルを捉える \*

# コストモデル

- ✓ LIFEプロジェクトやBCLの包括的なコストモデルに見られるように、ライフサイクル全体にまたがって、オブジェクトの保存コストを含めるように関心は拡大している (Eakin et al., 2008)。これらのモデルは英国の高等教育セクターおよびその会計の枠組みに由来するものであり、特にその基盤にある活動基準原価計算 (Activity Based Costing) の考え方をそのままわが国の状況に適用できるかは明確ではないが、間接コストを含めたその枠組みは有用と考えられる。
- ✓ 最も金額の低い方式が必ずしも最良であるわけでない。費用対便益 (効果) を考慮する必要がある (Houghton et al. (2009))。

# 現状分析(集計)

- \* 平成17年度から平成20年度までの『次世代学術コンテンツ基盤共同構築事業委託事業成果報告書』(『CSI委託事業(コンテンツ系)成果報告書』)の数値、および聴き取り(訪問、電話による)
- \* 集計結果
  1. 件数と金額のマッピング
  2. 事例
    - \* 博士論文
    - \* 雑誌論文

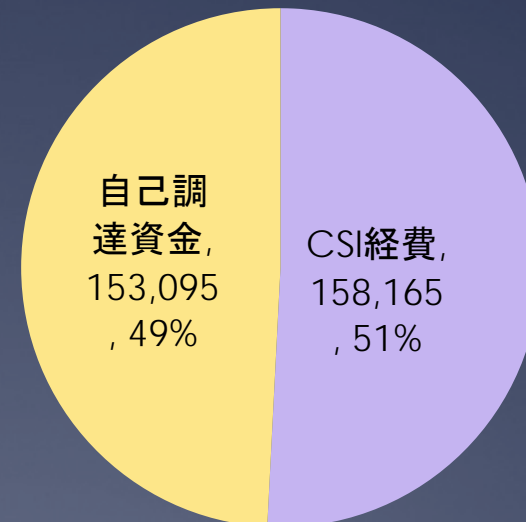
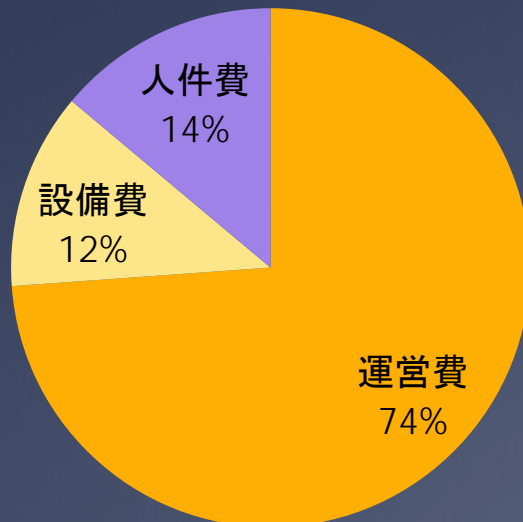
# システム導入経費

- \* 初期導入費(平成17年度～20年度)
- \* 保守費用(平成17年度～20年度;単年度ベース)
- \* 開発費(平成18年度～20年度累計)

	(1) 初期導入費	(2) 保守費用	(3) 開発費
最大値	26,250	12,537	28,045
平均値	2,209	512	2,042
中央値	1,764	160	966
標準偏差	3,211	1,819	1,972

# 運営費、設備費、人件費（平成20年度）

	運営費(千円)			設備費(千円)			人件費(千円)		
	CSI経費	自己調達資金	小計	CSI経費	自己調達資金	小計	CSI経費	自己調達資金	小計
平均値	1,984	1,404	3,362	112	445	557	230	402	632
中央値	2,000	645	2,590	0	0	0	0	0	0
最大値	3,300	11,726	13,926	2,400	9,144	9,144	1,451	3,750	4,375
最小値	87	0	0	0	0	0	0	0	0
合計	134,883	95,501	228,584	7,634	30,245	37,879	15,648	27,349	42,997





# コンテンツ作成経費と人件費

## (平成19年度、20年度)

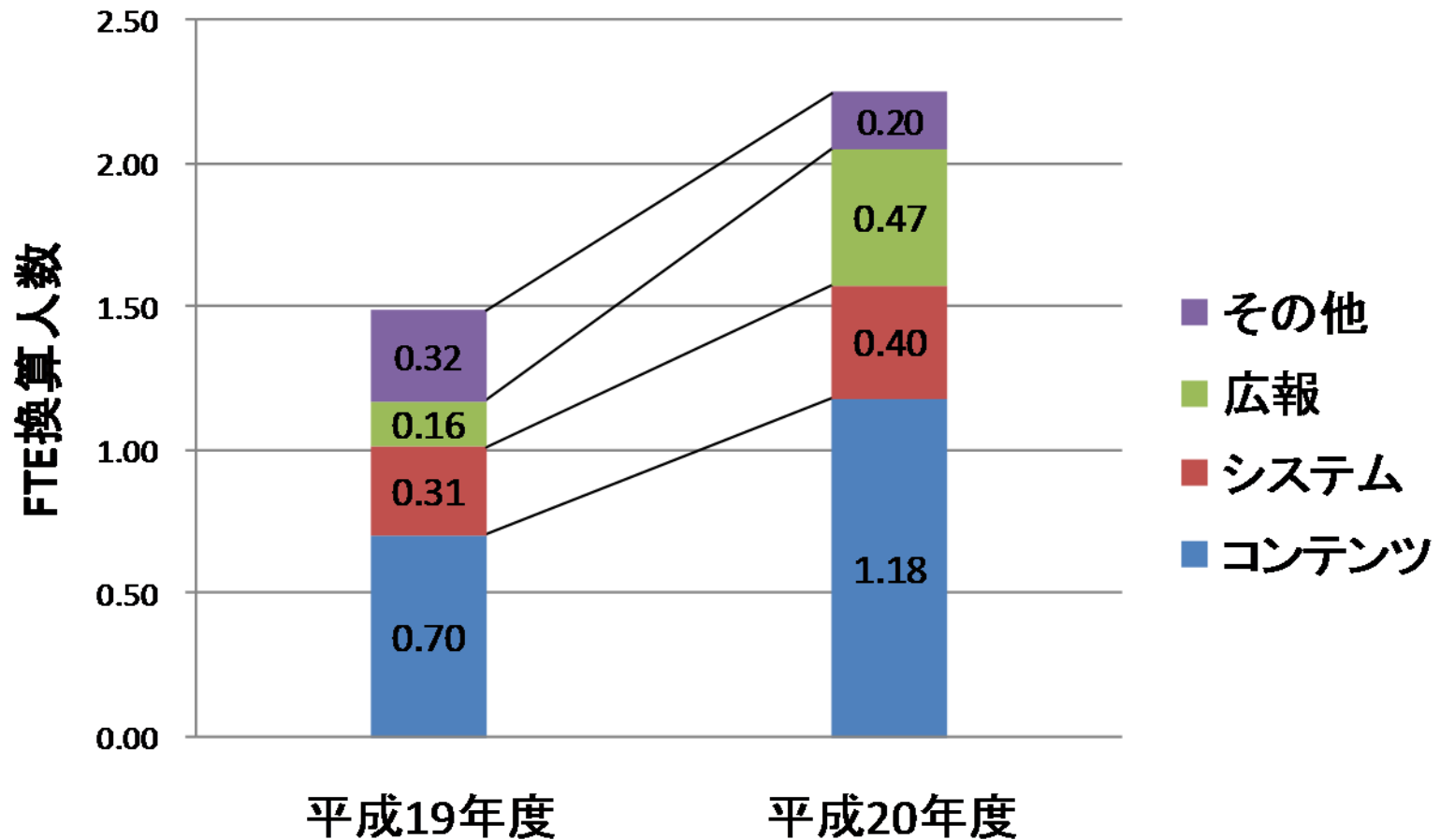
	コンテンツ作成経費 (CSI経費分のみ)(千円)		その他人件費(千円)	
	平成19年度	平成20年度	平成19年度	平成20年度
平均値	2,209	1,461	392	527
中央値	1,235	1,414	0	0
最大値	18,749	3,300	5,209	9,247
最小値	0	0	0	0
標準偏差	2,900	956	843	1,514

# 投入人員(FTE換算)

	平成19年度				
	コンテンツ	システム	広報	その他	計
平均値	0.70	0.31	0.16	0.32	1.49
中央値	0.5	0.2	0	0.2	1.1
最大値	3.2	2	1.8	2	6
最小値	0	0	0	0	0
標準偏差	0.73	0.41	0.33	0.41	1.26
	平成20年度				
	コンテンツ	システム	広報	その他	計
平均値	1.18	0.40	0.47	0.20	2.25
中央値	1	0.3	0.3	0	1.8
最大値	7	2	5	1.8	14
最小値	0	0	0	0	0.05
標準偏差	1.04	0.42	0.70	0.36	1.99

※本表の人員費には、非常勤職員、パート職員を含む。

# 投入人員(FTE換算平均値)



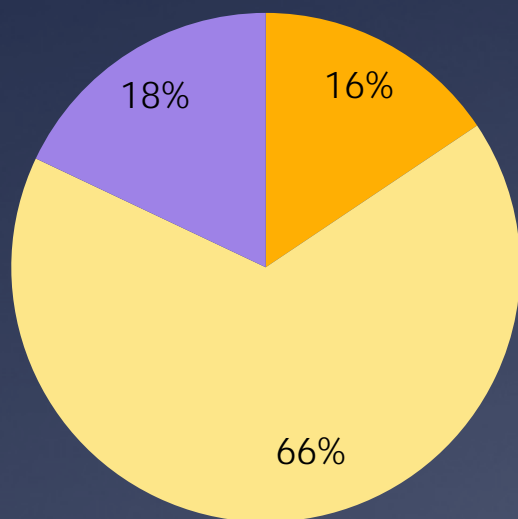
# 処理件数とワークフロー

## (平成20年度)

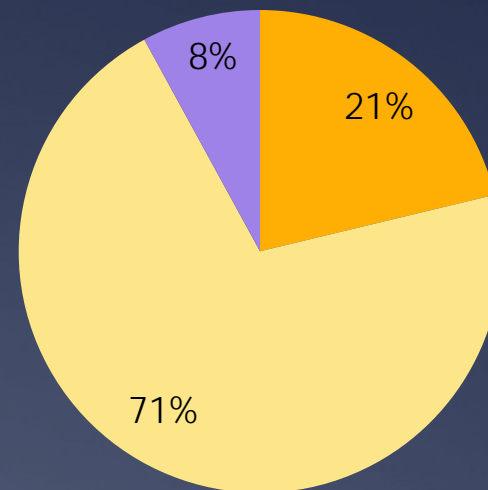
平成20年度実績	ポーンデジタル		印刷物の電子化		既存システムからの移行		小計	全体数	内訳の報告率
	件数	割合	件数	割合	件数	割合			
学術雑誌論文	4,325	15.6%	18,424	66.4%	4,981	18.0%	27,730	32,387	85.6%
学術雑誌論文(海外)	2,462	72.4%	557	16.4%	382	11.2%	3,401	4,930	69.0%
学術雑誌論文(国内)	1,124	4.8%	17,738	75.6%	4,599	19.6%	23,461	26,589	88.2%
学術雑誌論文(海外、国内込み)	739	85.1%	129	14.9%	0	0.0%	868	868	100%
学位論文	675	21.2%	2,250	70.8%	253	8.0%	3,178	5,144	61.8%
研究成果報告書	1,400	42.6%	1,809	55.0%	79	2.4%	3,288	4,126	79.7%
紀要論文	5,988	9.1%	44,908	68.2%	14,982	22.7%	65,878	86,100	76.5%
会議発表論文	327	5.5%	5,469	92.4%	126	2.1%	5,922	9,068	65.3%
会議発表用資料	317	88.8%	40	11.2%	0	0.0%	357	884	40.4%
図書・図書の章	18	2.8%	625	97.2%	0	0.0%	643	708	90.8%
テクニカルレポート、ワーキングペーパー	248	29.1%	54	6.3%	550	64.6%	852	877	97.1%
一般雑誌記事	42	23.5%	113	63.1%	24	13.4%	179	952	18.8%
プレプリント	10	14.3%	33	47.1%	27	38.6%	70	76	92.1%
教材	103	26.8%	32	8.3%	249	64.8%	384	524	73.3%
教材(OCW)	13	100%	0	0.0%	0	0.0%	13	13	100%
データ・データベース	6	100%	0	0.0%	0	0.0%	6	11	54.5%
ソフトウェア	0	0.0%	0	0.0%	0	0.0%	0	2	0.0%
その他	350	16.9%	1,621	78.4%	97	4.7%	2,068	5,321	38.9%
合計	13,822	12.5%	75,378	68.2%	21,368	19.3%	110,568	146,193	75.6%

# 処理件数とワークフロー (平成20年度)

## 学術雑誌論文



## 学位論文



- ボーンデジタル
- 印刷物の電子化
- 既存システムからの移行

# 外注入力の単価と方式 (学位論文の場合)

	A大学			B大学					
	単位	単価	仕様	単位	単価	仕様	単位	単価	仕様
				(解体・再製本)			(解体不可の場合)		
PDF作成作業									
モノクロ	頁	35		頁	25	400dpi	頁	39	400dpi
グレースケール				頁	70	300dpi	頁	180	300dpi
カラー	頁	50		頁	100	200dpi	頁	220	200dpi
PDFの透明テキスト化	件	500	校正なし	頁	12	校正なし	頁	12	校正なし
PDF編集作業 (カラー画像差込作業他)	件	500							
書誌情報作成	文字	0.8							
リポトリ登録作業	件	300			なし			なし	
抄録入力	文字	0.8			なし	入力しない		なし	入力しない
基本料金 (運搬費、打合せ費等)	式	0							
再製本(ハードカバー)				件	1,450				
再製本(ソフトカバー)				件	400				

# 外注入力の単価と方式 (学位論文の場合)

	C大学			D大学			E大学		
	単位	単価	仕様	単位	単価	仕様	単位	単価	仕様
				(解体・再製本)			(解体不可の場合)		
PDF作成作業									
モノクロ	頁	58		頁	26.25		件	4,700	モノクロ
グレースケール									
カラー				頁	84				
PDFの透明テキスト化				頁	15.75	校正なし			校正なし
PDF編集作業 (カラー画像差込作業他)									
書誌情報作成				件	315		件	220	内部作業 (人件費を換算)
リポジトリ登録作業		なし							
抄録入力		なし	入力しない	文字	1.785			なし	入力しない
基本料金 (運搬費、打合せ費等)									
再製本(ハードカバー)		なし		件	1,995			なし	
再製本(ソフトカバー)		なし						なし	

# 今後の課題

- \* LIFEプロジェクトやBCLのモデルを、CSI事業における機関リポジトリに適用するための評価と検討
- \* 成果報告書における、人件費、FTE換算の投入人員等に関する概念および項目説明の明確化
- \* 遡及的PDF作成のような作業における、品質面での標準の設定
- \* 活動基準原価計算を行うための、タイプの異なる数機関における詳細なケーススタディ(参加機関の協力が必要)