

平成25年度 学術情報システム総合WS 最終報告

グループ1:
名古屋大学 小島由香
日本原子力研究開発機構 長屋 俊
大阪大学 藤江雄太郎

本日の発表内容

1. 当初の課題設定とアプローチ
2. 第2回WSでの講評を受けて
3. 調査報告の前に
4. 調査報告①NACISIS-ILLログの計量的な分析
5. 調査報告②ナビゲートログによる分析の有効性検証
6. 調査報告③NII/NDL/JST以外のナビゲートの分析
7. 調査報告④NII/NDL/JSTナビゲートの分析
8. まとめと今後に向けて

1.当初の課題設定とアプローチ

第1回WSでの問題整理

デジタル化資料について問題は大きく以下の2つだと考えました。

A) そもそも存在を「知らない」

- ポータルの存在や、資料種別の存在などさまざま次元で「知らない」

B) あるはずだけど「たどり着けない」

- 探しているものがWeb上にあるという情報は持っているが、「たどり着けない」

当初の課題設定とアプローチ

A)「知らない」について

目標	国内外の主要データソース/ポータルの見取り図の作成
背景	図書館員がそもそも「知らない」。知らないので、利用者に情報提供もできない。
行った調査内容	NACSIS-ILL謝絶のログ解析 —潜在的ニーズの把握
	国内外の状況調査(基本情報と関係性を中心に) —ポータルの収録内容について知る —国内デジタル化資料の標本調査。全体像把握のあしかりに。

当初の課題設定とアプローチ

B)「たどり着けない」について

目標	デジタル化資料のメタデータ/組織化についての提言の作成
背景	データソース/ポータルによってメタデータが不揃いなことや、名寄せができていないことが原因で、探しているものに「たどりつけない」
行った調査内容	国内外のメタデータに関する報告書の収集 — 今後必要と思われる構成要素とメタデータ組織化の手法の抽出を目指す

2. 第2回WSでの講評を受けて

最終目標の設定

【最終目標】

発見できないデジタル化資料の傾向を分析し、発見できるような提案をする。

—見つかりづらいデジタル化資料の“Discovery”
に焦点をあてる

最終目標に向けた具体的なタスク

【具体的な調査・作業内容】

調査対象は国内デジタル化資料に絞る

- NACSIS-ILL謝絶ログの分析対象を増やし、より深く分析する
 - 無料デジタル資料にナビゲートしたILLの、ナビゲート先デジタル化資料を分類して、見つからない原因を探る
 - WS前半の標本調査の調査項目や調査結果をうまく絡める
 - 依頼館/受付館や分野の分析も行い、提言の材料として偏りがないかチェックする
- デジタル化資料が見つからない原因を明らかにすることで、学術コンテンツの発見性を高めるような手引き/ツールの作成を目指す

3.調査報告の前に 今回使用したデータの計量的情報について

調査のデータソース

- NII提供NACSIS-ILLログデータ(個人情報部分を除き、指定の項目に限って提供を依頼)
 - 最終更新日が2010.4.1～2013.3.31
 - この3年間で終了したレコード+最終更新日が上記で、レコードが最終状態に行っていないもの
- ※後者は今回分析の対象外

ILLログを調査に使用した理由

「ある情報を必要としていて発見できなかったが、実はそれが無料デジタル化資料として公開されている」という、情報探索行動と結果が記録として残っているものだから。

⇒「無料のデジタル化資料のうち発見しづらいもの」のサンプルとしてある程度有効と考えられる。

レコード件数

- NACSIS-ILL終了レコード件数（NII公表）

- ここでいう終了は最終状態が「確認」+「返却確認」

年度	2010年度	2011年度	2012年度
複写	789,397	760,947	721,349
貸借	94,813	90,180	92,877
計	884,210	851,127	814,226

- 今回の調査対象のレコード件数

- NII統計と若干誤差があるが抽出条件は同じため問題なしとした

年度	2010年度	2011年度	2012年度
確認	789,661	761,465	721,445
返却確認	94,907	90,225	92,970
計	884,568	851,690	814,415

ログの内訳

● A:ログ全体

- 最終状態が「確認」+「返却確認」→「成功したILL」
- 最終状態が「CANCEL」→「キャンセルしたILL」

年度	2010年度	2011年度	2012年度
A-1: 成功したILL	884,568	851,690	814,415
A-2: キャンセルしたILL	61,572	59,814	58,866
Z: まだ最終状態でないILL	-	-	529
計 (=A全体)	946,140	911,504	873,810

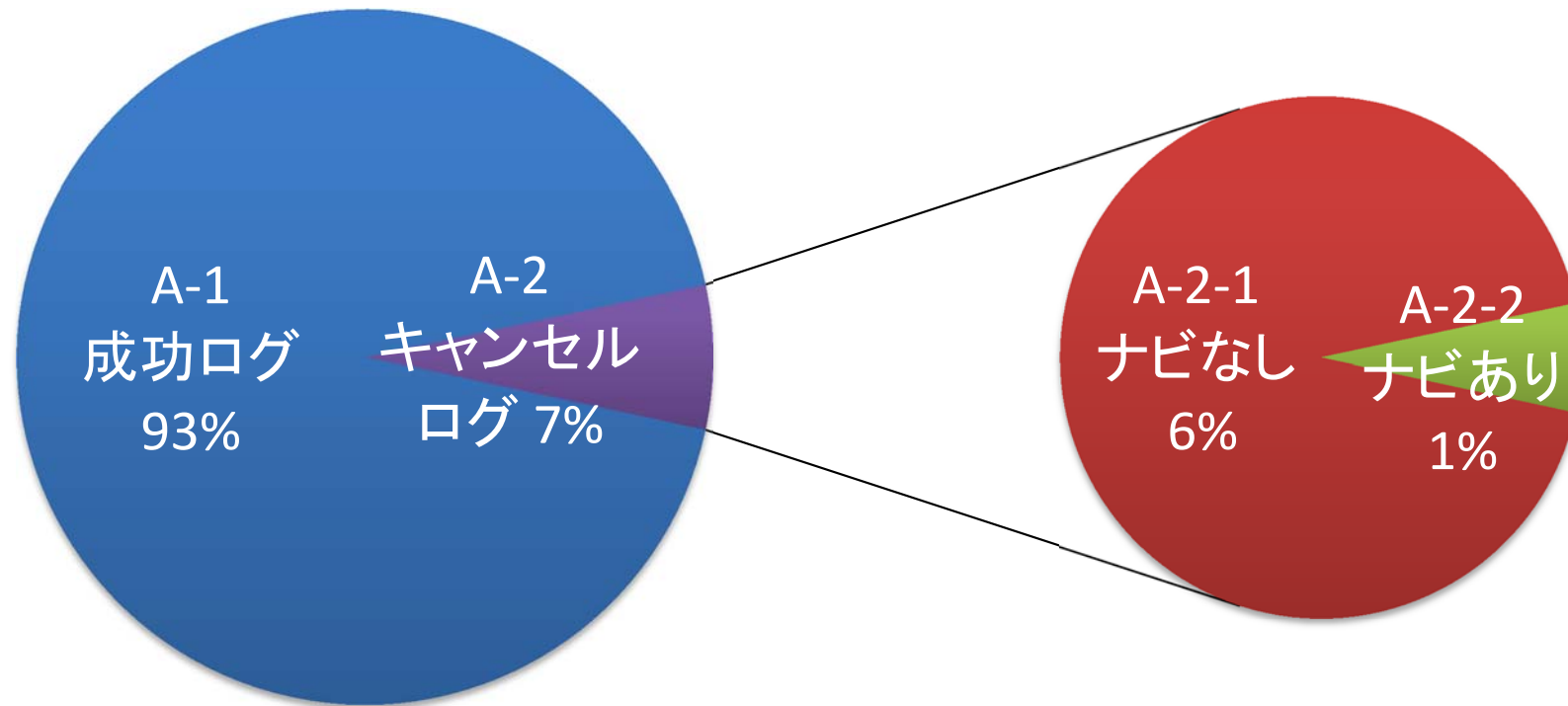
● A-2:キャンセルログ

- 無料デジタル化資料へのナビゲート有無で分類

年度	2010年度	2011年度	2012年度
A-2-1:ナビゲートなし	58,380	56,470	54,920
A-2-2:ナビゲートあり	3,192	3,344	3,946
計(A-2: キャンセルしたILL)	61,572	59,814	58,866

ログの内訳

2012年度ログ概要



デジタル化資料へのナビゲート有無 によるキャンセルログの仕分け

参加館間でやりとりされたコメントをもとに仕分けた

1. キャンセルログのやりとりコメントのうち、コマンドが“INQUIRE”もしくは“PARDON”のものを対象
2. (1)のログから、正規表現にてURLを機械的に抽出、もしくは形態素解析(mecab利用)により抽出したキーワードを利用し振り分けたログを抽出。
3. (2)のログを手作業で確認、デジタル化資料へナビゲートあり/なしのログを確定。この段階で1レコードに対し、複数のコメントのあるものは重複除去した。
4. ナビゲートあり(以降、「ナビあり」)のログのうちURL記載のないログは、URLを調査・推定し、追加した。

ナビありログの内訳

年度	2010年度	2011年度	2012年度
JST	304	490	577
NDL	21	52	84
NII	286	192	266
リポジトリ	643	555	531
その他	1,940	2,070	2,568
計(A-2-2:ナビあり)	3,194	3,359	4,026

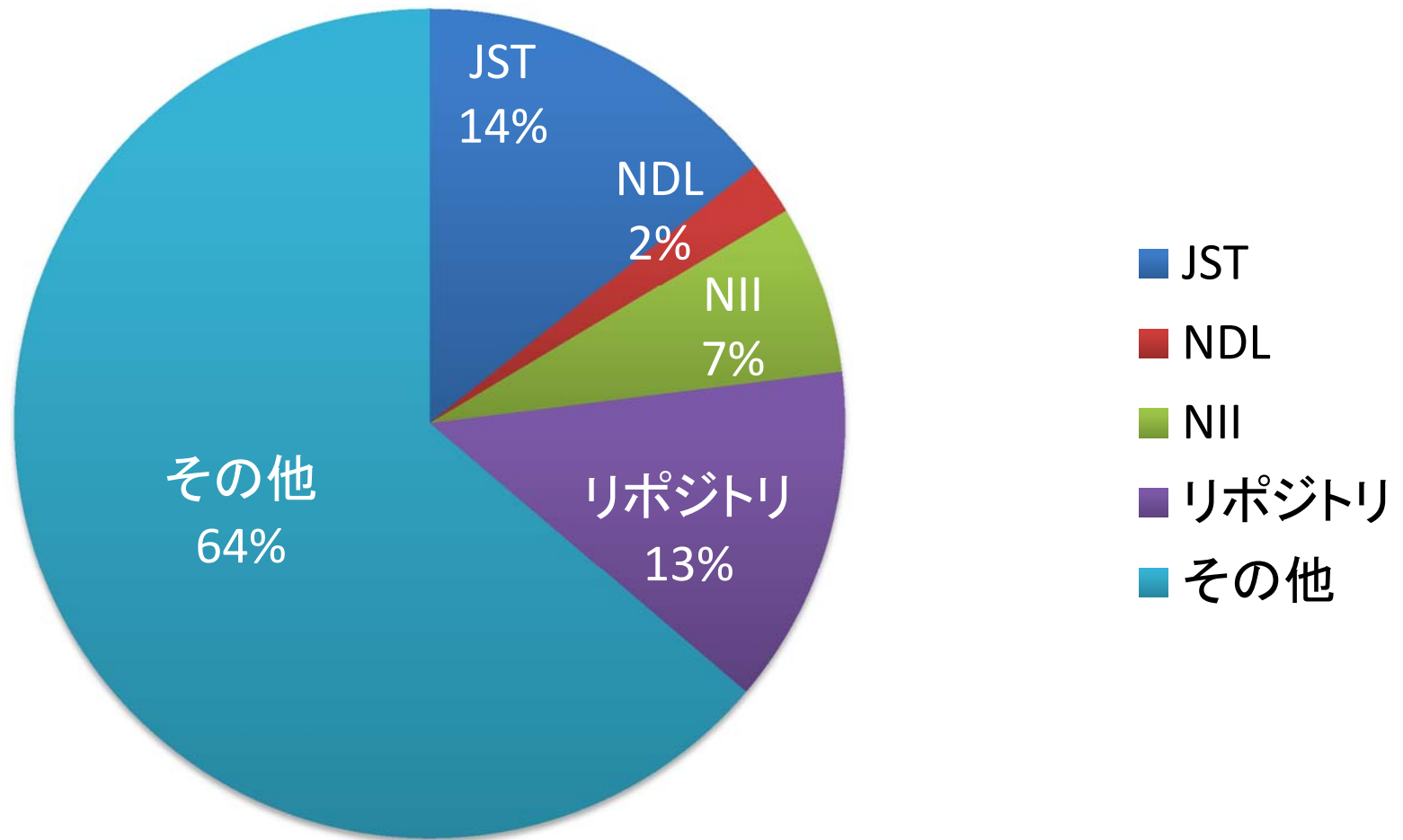
※ナビありログ数と差分あり

1つのコメントに複数URLがある場合は両方のURLをカウントしたため差分が発生

※URLなしログは、URLを推定し追加

ナビありログの内訳

2012年度ナビありログの内訳



この後の調査報告部分の 用語定義について

これらの用語は、以下のように定義します。

－「ポータル」

学術コンテンツのメタデータや本文(＝無料のデジタル化資料)を収集して、まとめて検索できるようにしているサービス。

－「BIBID」

NACSIS-ILLにおいて、依頼館が記入する書誌ID。ほとんどがNACSIS-CATの書誌ID(=NCID)にあたる。

4.調査報告①

NACISIS-ILLログの計量的な分析

BIBIDの入力割合

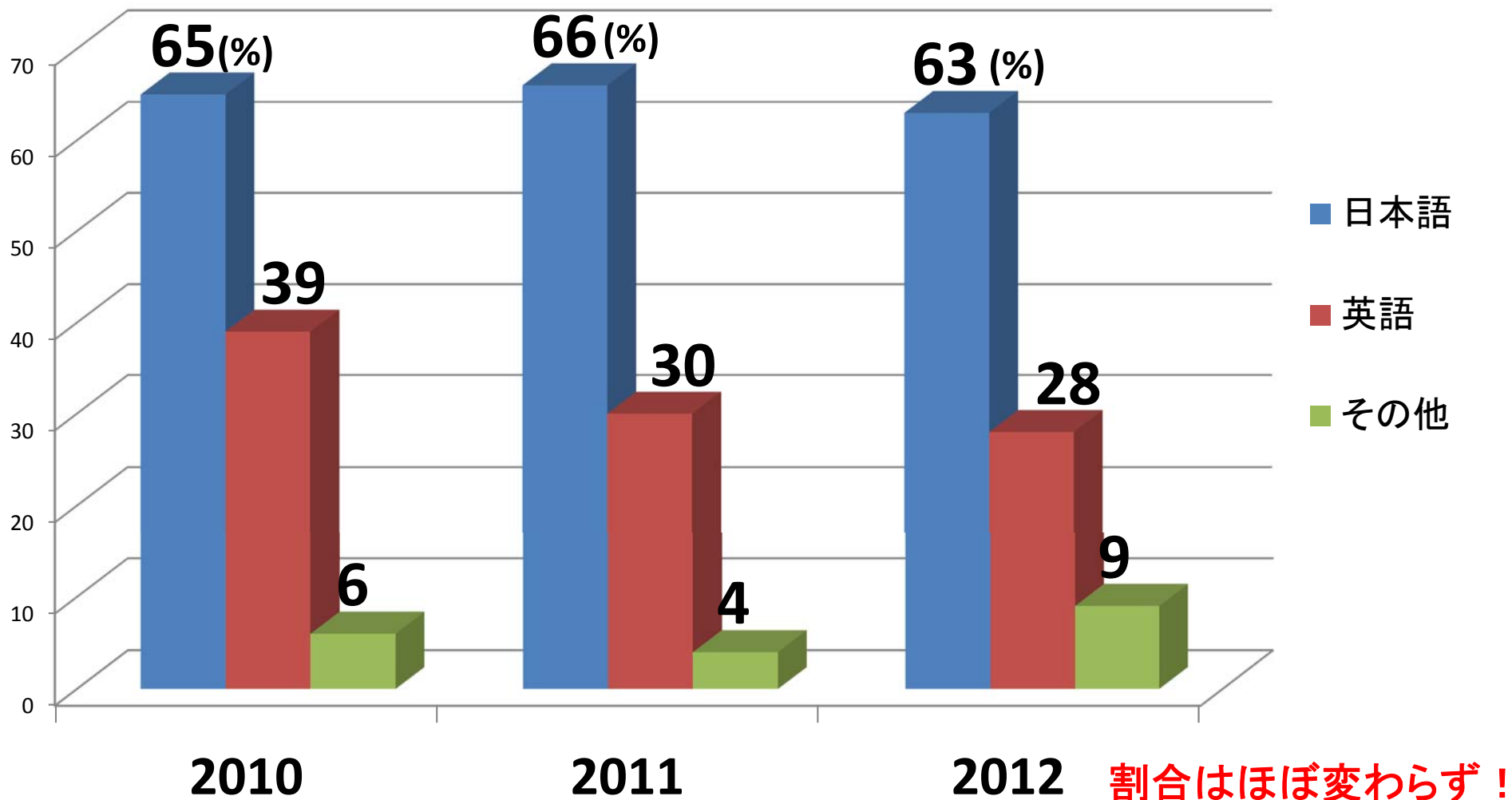
	2010	2011	2012
成功ILL (全体の割合)	98%	98%	98%
キャンセル (全体の割合)	74%	73%	72%
ナビゲート分 (全体の割合)	97%	96%	96%

キャンセルになるILL自体には1/4程度BIBIDの入力がないものがあるが、ナビゲート分はほとんどのデータに入っている

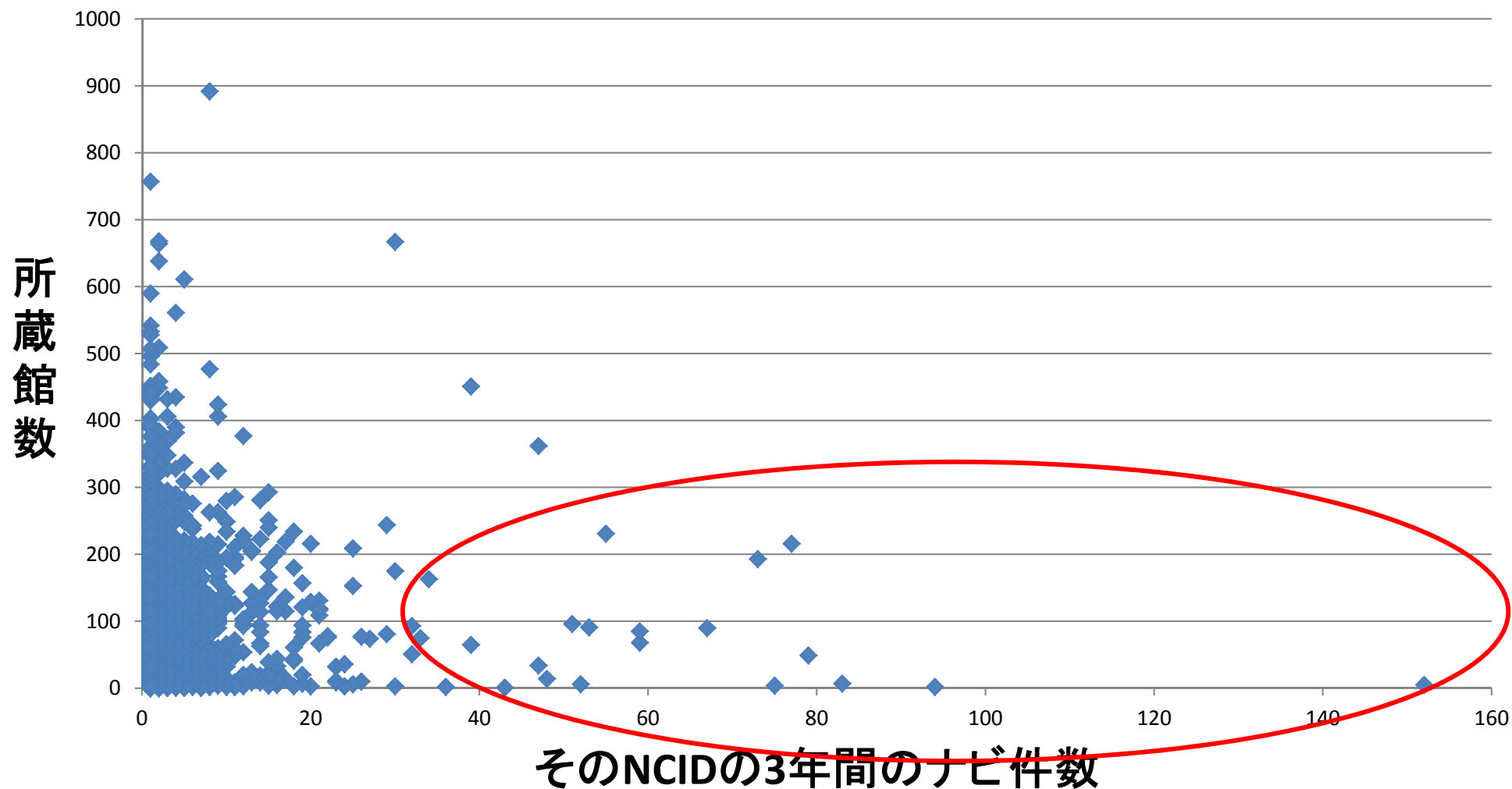
CiNii APIを利用した問い合わせ

- BIBIDをキーとしCiNii APIを利用して書誌データを取得し追加分析
 - 本文言語分布
 - 所蔵館分布

ナビあり分の本文言語分布



ナビあり分の所蔵館数分布



5.調査報告②

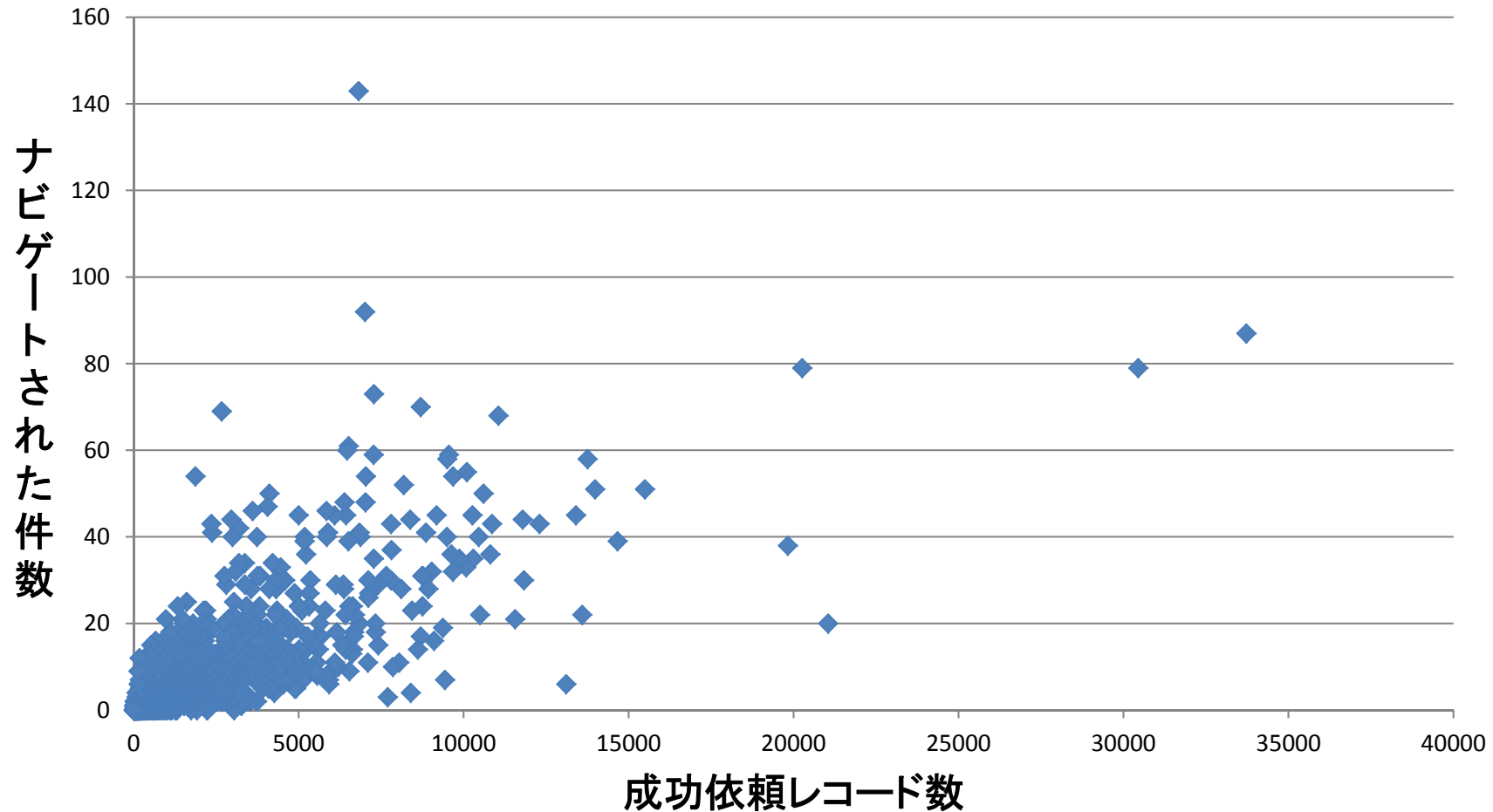
ナビゲートログによる分析の有効性検証

ナビゲートログを検討材料にすること の有効性の確認

- 依頼館/受付館に偏りがなにか確認する
- もし依頼館に偏りがあれば：
 - 一部の依頼館の確認ミスにより、受付館でのナビゲートが発生してしまっている(=大部分の依頼館には関係のない話)
- もし受付館に偏りがあれば：
 - ナビゲートを業務フローに組み込んでいる一部の受付館が気づいた無料のデジタル化資料の集合になってしまっている

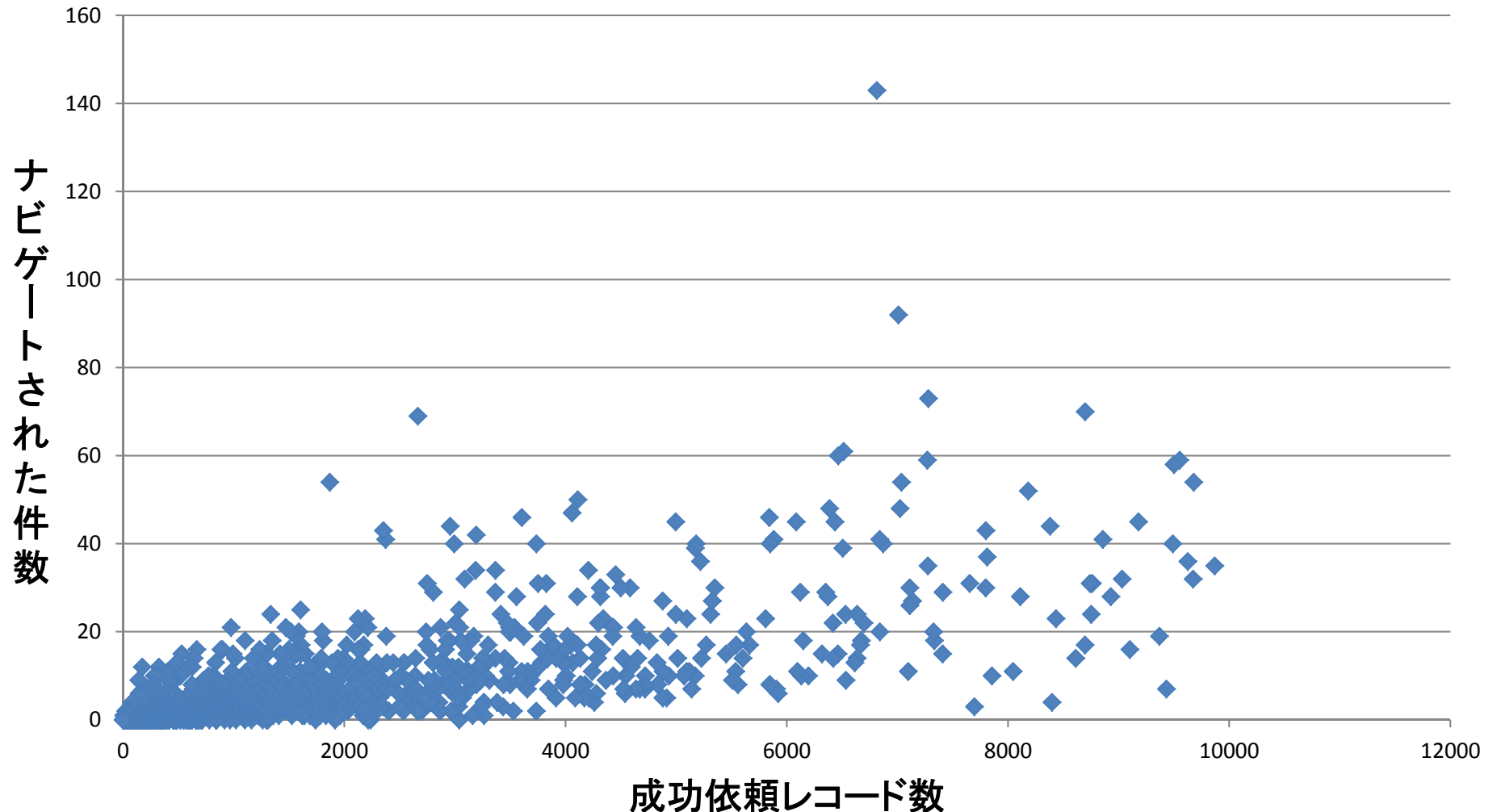
ナビゲートログ依頼館の傾向①

参加館ごとの成功/ナビゲート割合2010-2012

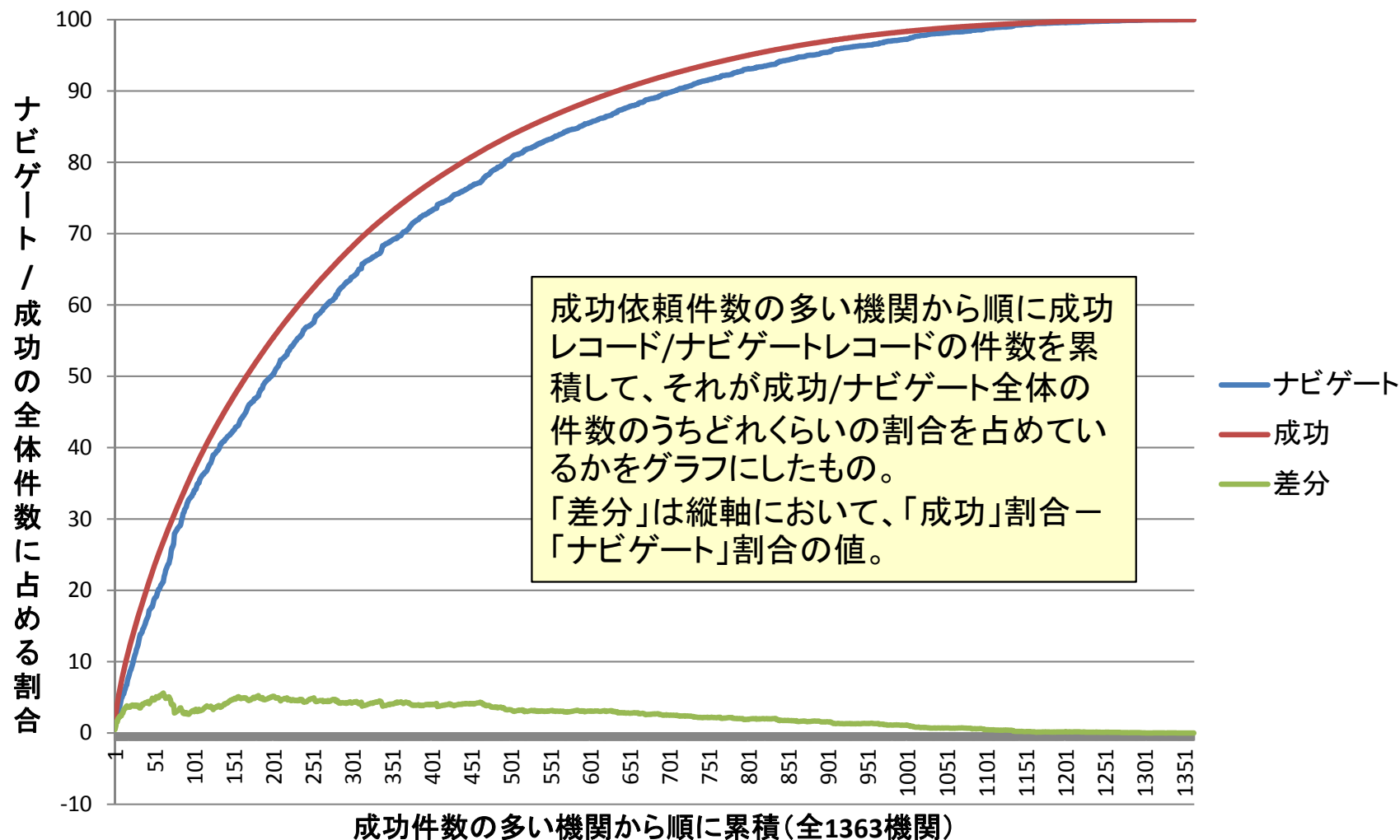


ナビゲートログ依頼館の傾向①

そのうち、成功記録1万件以下の参加館



ナビゲートログ依頼館の傾向②

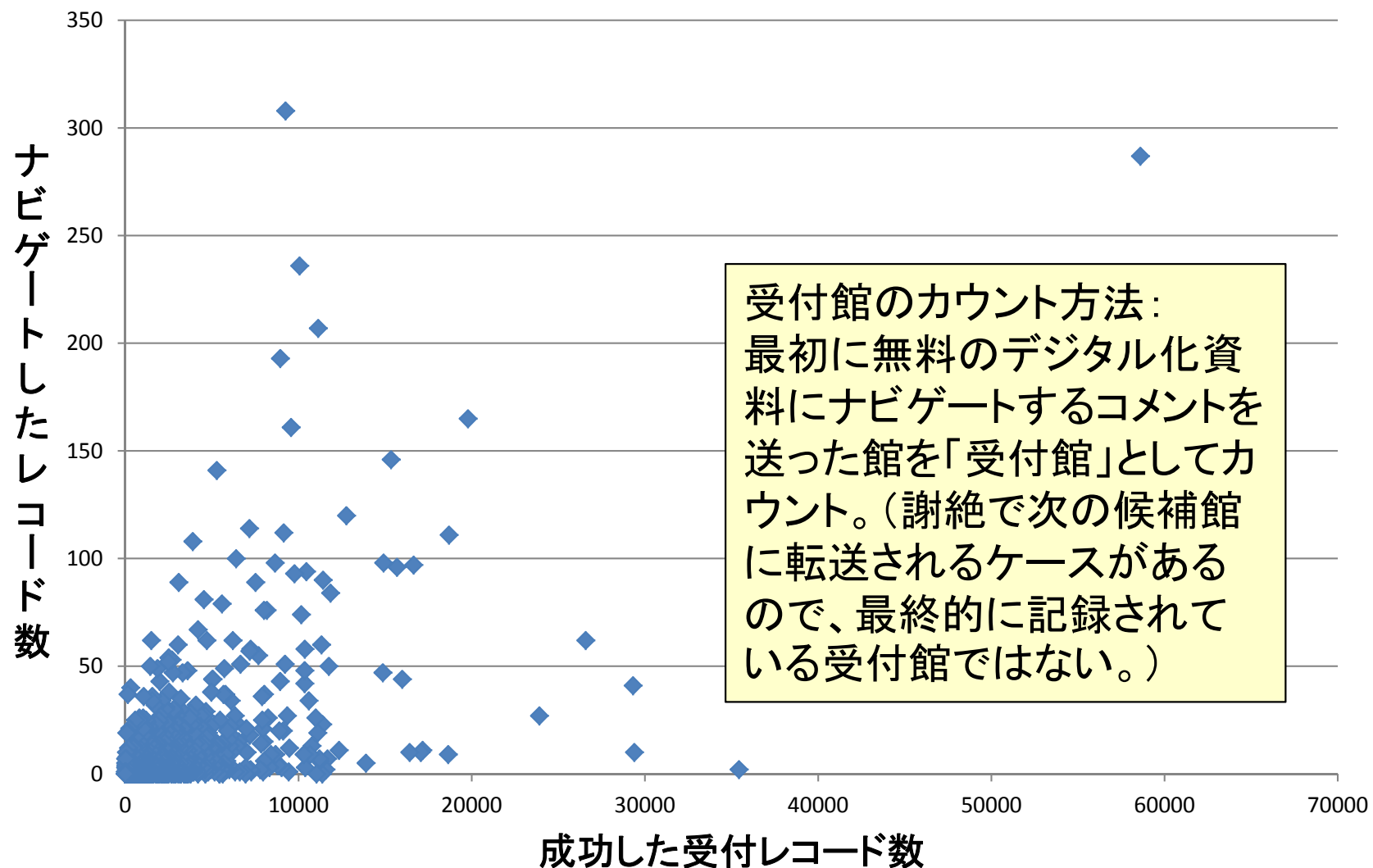


ナビゲートログ依頼館の傾向

- ナビゲートされた件数が、0件の参加館が23%、0～3件の参加館が50%、0～21件で90%
- どの参加館もある程度の件数はナビゲートされてしまっている
- ②のグラフの分布より、成功依頼とナビゲートされたもののグラフはおおよそ類似した形になっている＝依頼とナビゲートはおおよそ比例
→依頼館の偏りは考えなくてもよい

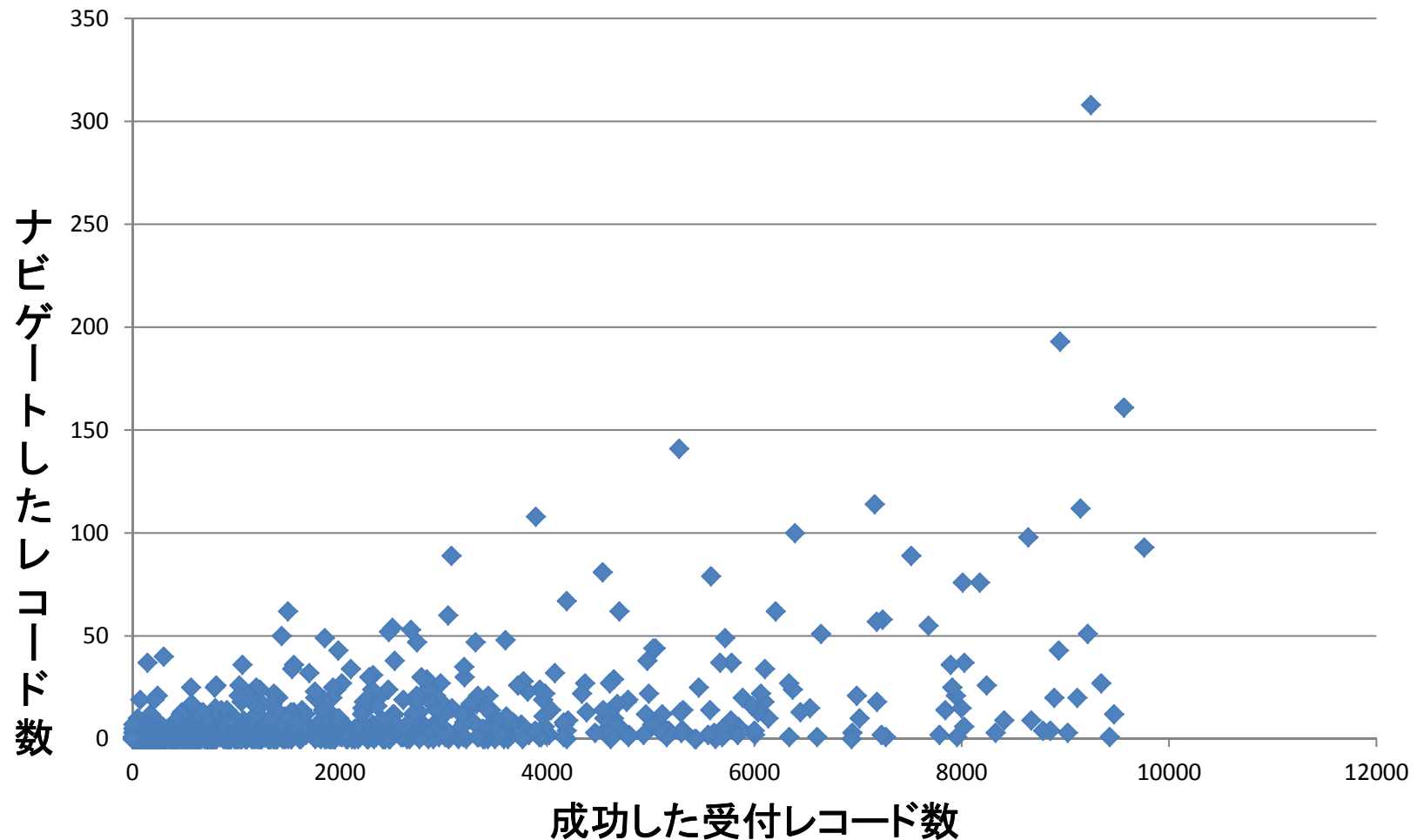
ナビゲートログ受付館の傾向①

参加館ごとの成功/ナビゲート割合2010-2012

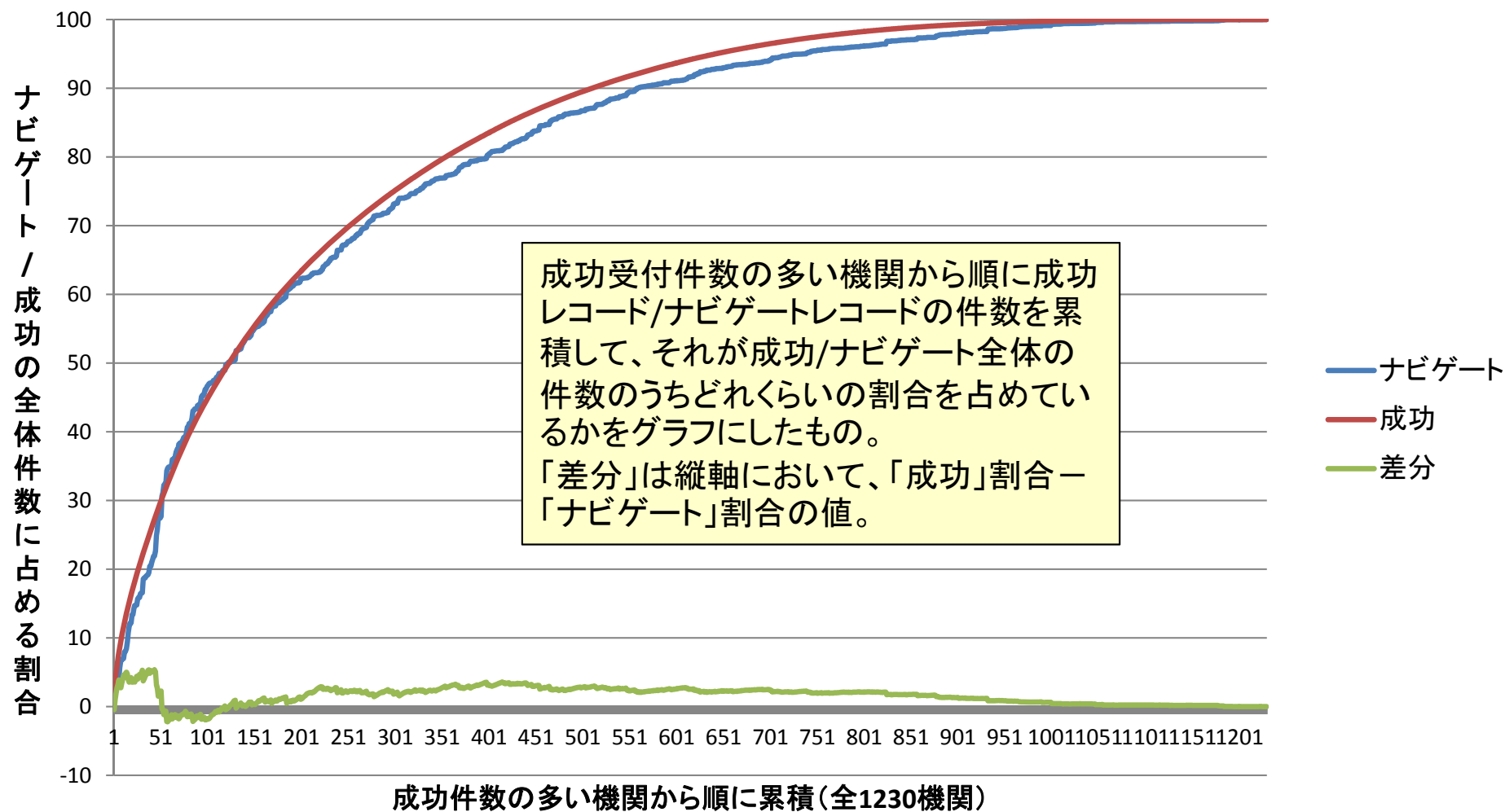


ナビゲートログ受付館の傾向①

そのうち、成功記録1万件以下の参加館



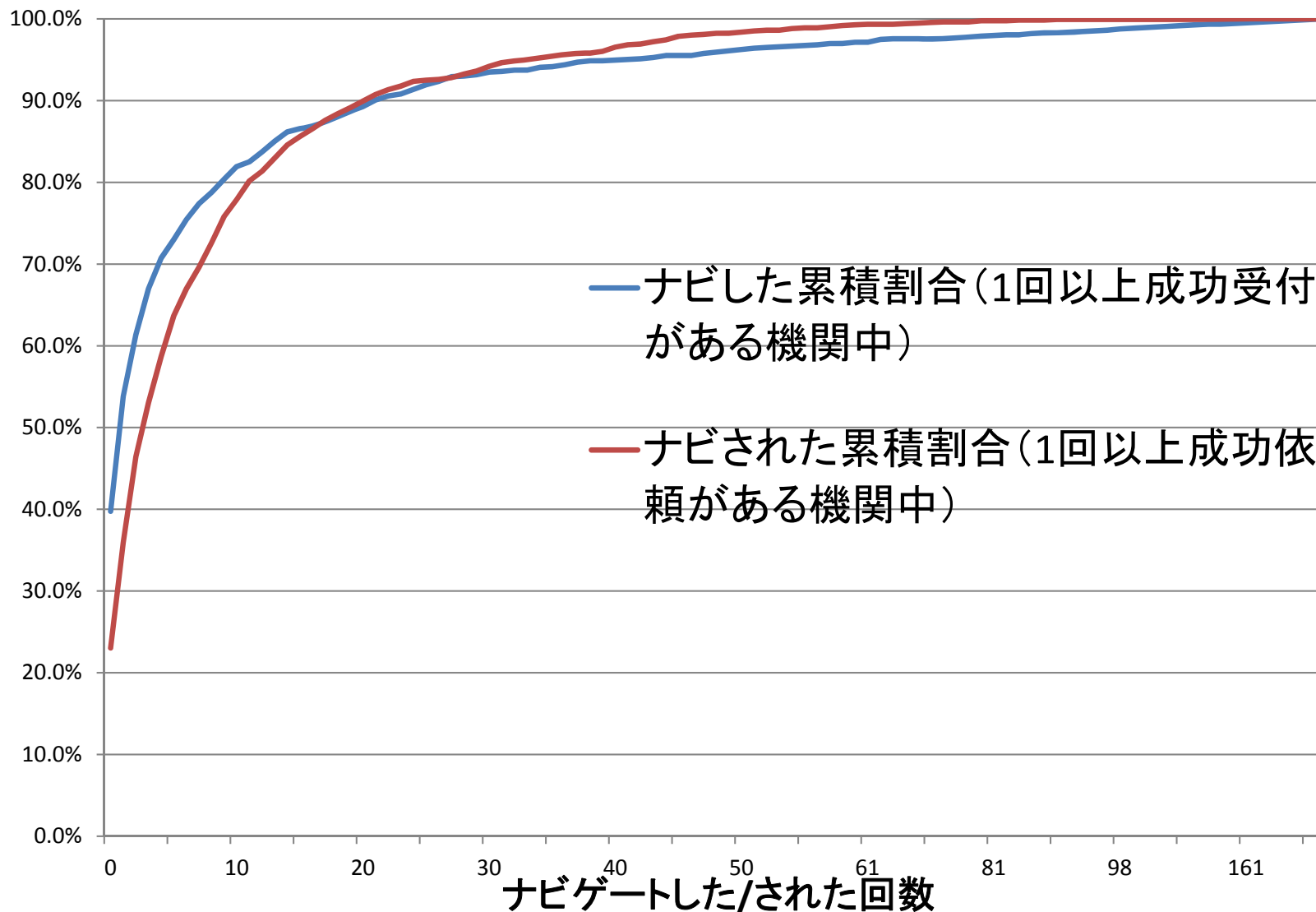
ナビゲートログ受付館の傾向②



ナビゲートログ受付館の傾向

- ナビゲートした件数が、0件の参加館が40%、0～3件の参加館が67%、0～9件で80%
- ナビゲートした数が1ケタ台の参加館が8割。その中には受付件数上位の機関も。
- ナビゲートという行為をILL受付フローに組み込んでいるところと、そうでないところがあると考えられる。
- ただ、②のグラフの分布より、成功受付とナビゲートした件数のグラフはおおよそ類似した形になっており、割合の差も小さいので、全体の傾向としては大きな偏りはないと考えられる

【参考】依頼館のちらばりと受付館のちらばり



6.調査報告③

NII/NDL/JST以外のナビゲートの分析

再掲：ナビありログの内訳

年度	2010年度	2011年度	2012年度
JST	304	490	577
NDL	21	52	84
NII	286	192	266
リポジトリ	643	555	531
その他	1,940	2,070	2,568
計(A-2-2:ナビあり)	3,194	3,359	4,026

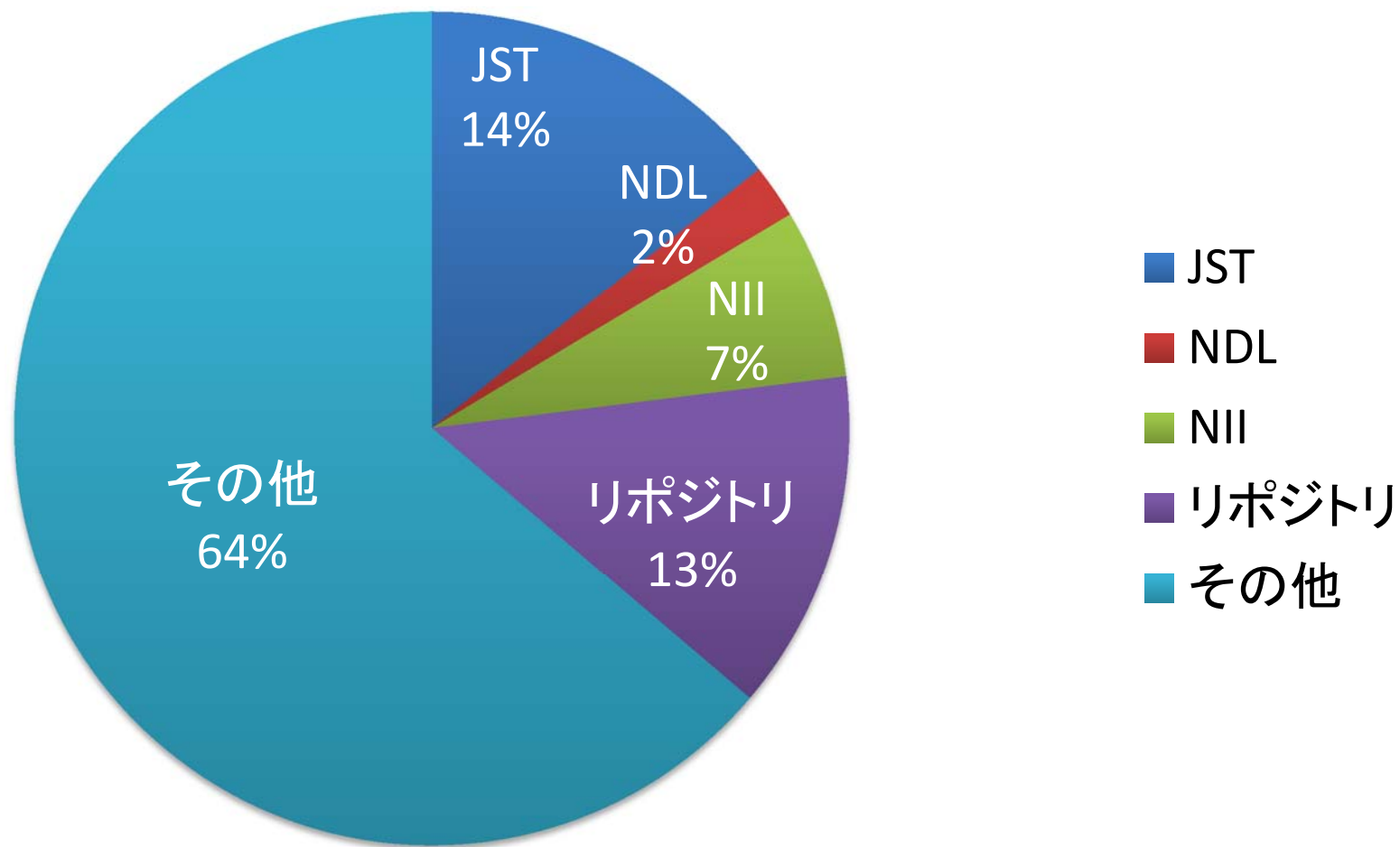
※ナビありログ数と差分あり

1つのコメントに複数URLがある場合は両方のURLをカウントしたため差分が発生

※URLなしログは、URLを推定し追加

再掲：ナビありログの内訳

2012年度ナビありログの内訳



6-1. リポジトリ以外のナビゲートロ グ

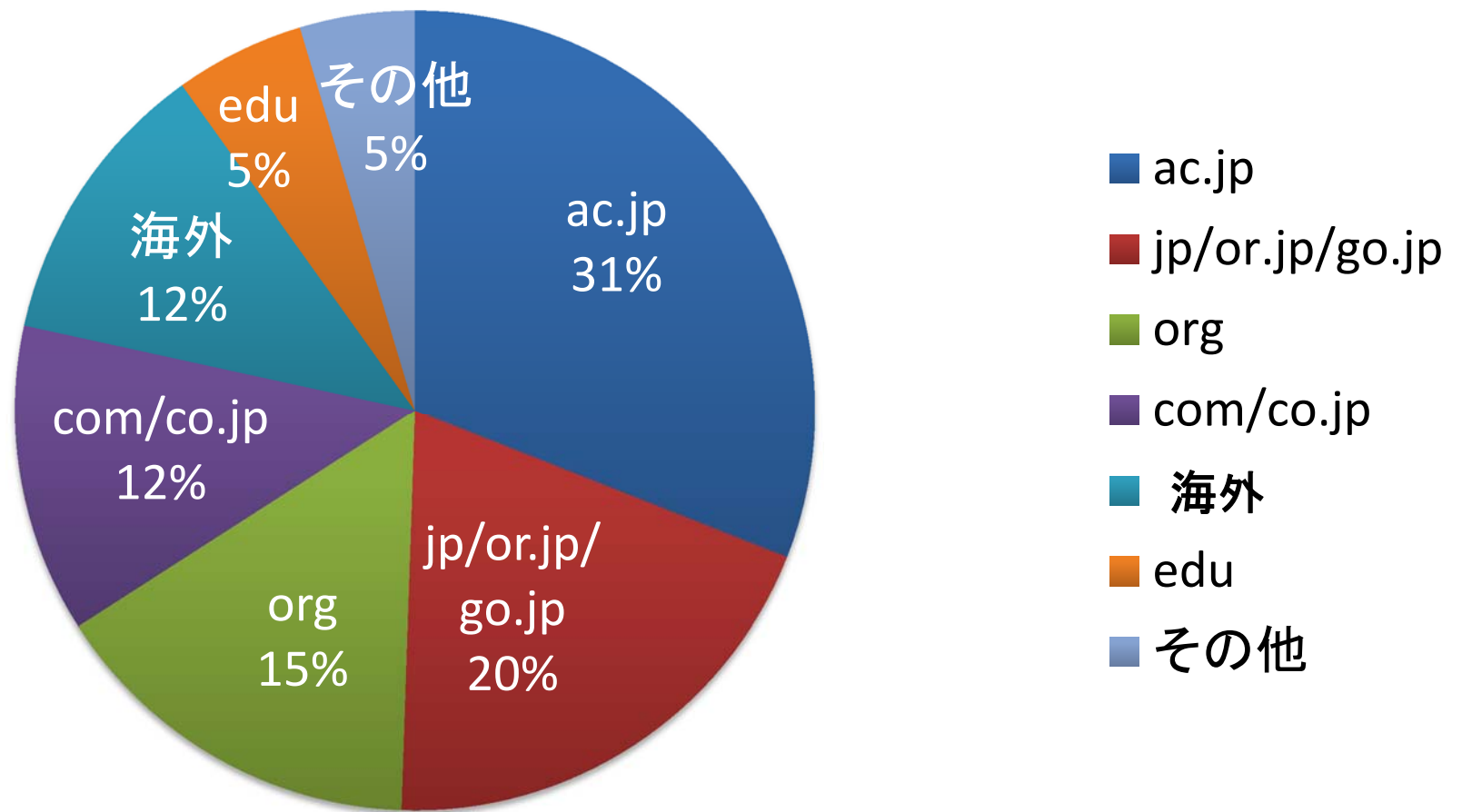
ログの内訳

年度	2010	2011	2012
ac.jp	275	261	260
jp/or.jp/go.jp	137	155	164
org	90	107	129
com/co.jp	70	91	105
海外	70	99	98
edu	33	33	44
その他	21	28	39
計	696	774	839

※NII/NDL/JST/リポジトリを除く

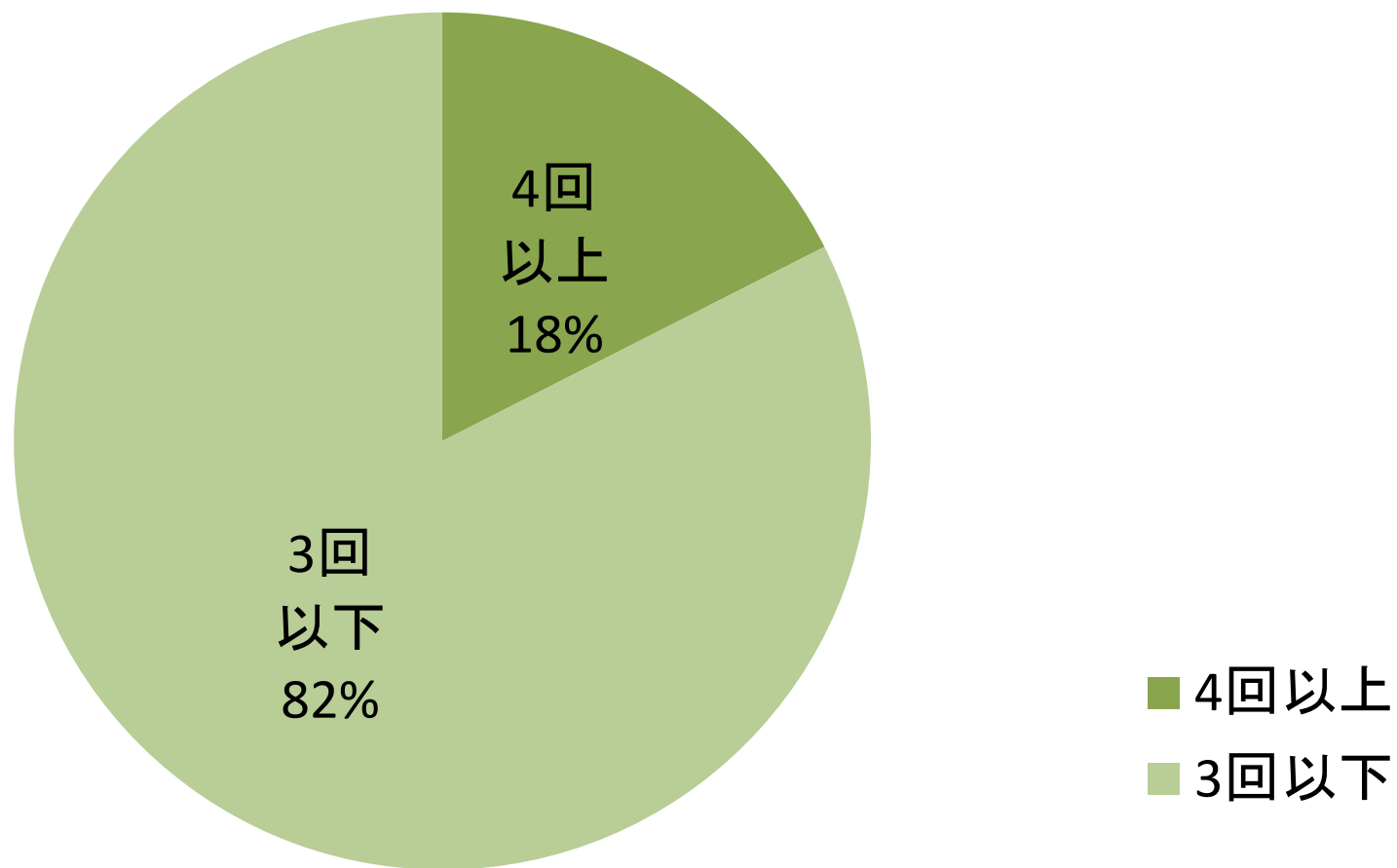
ログの内訳

2012年度

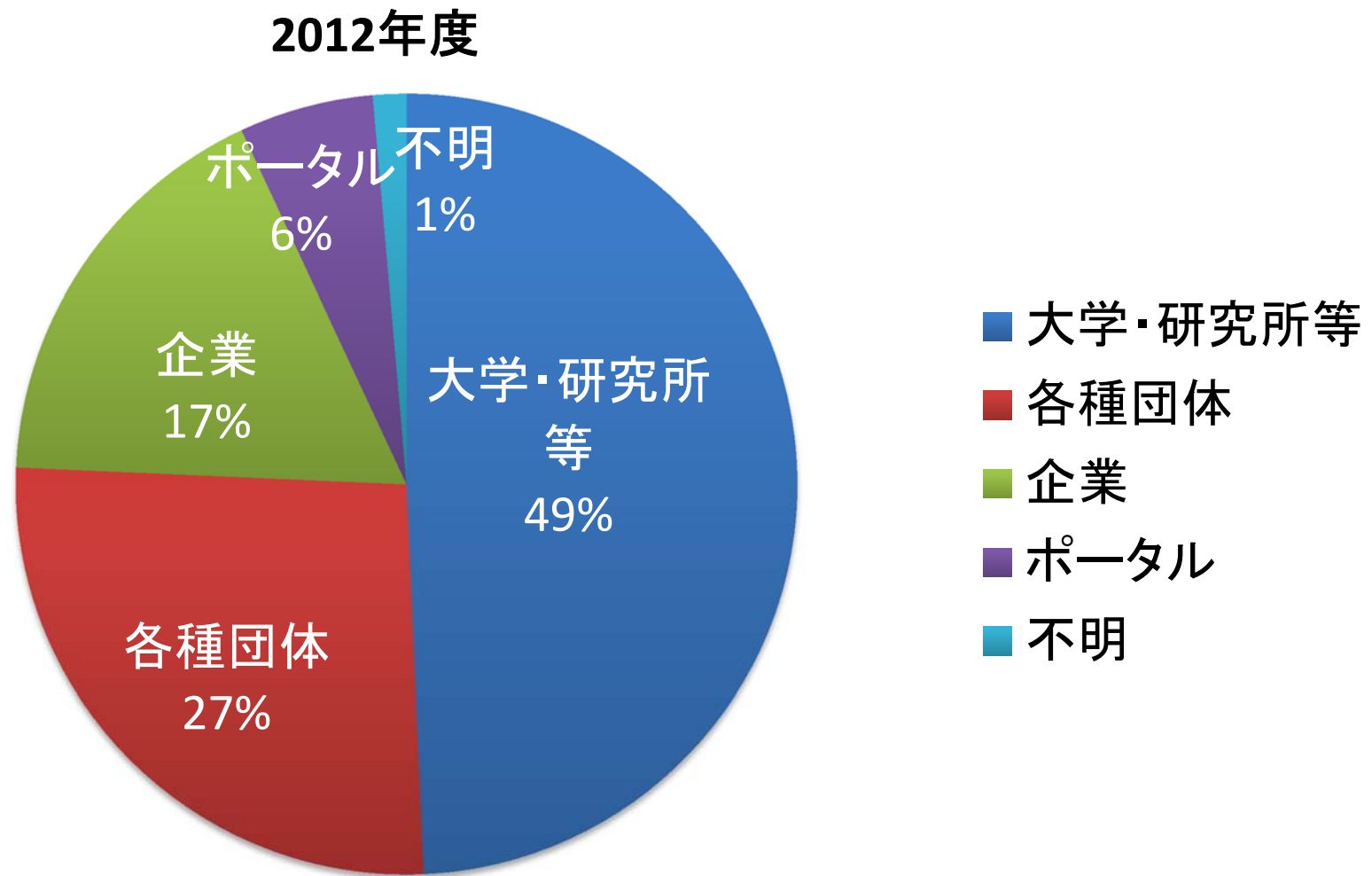


4回以上出現するURLの割合

2012年度



4回以上出現するURLのカテゴリー別割合



よく出現するURLリスト

- 4回以上出現するURLを、3年間の出現回数
の合計が多い順に、リスト化

サイト名	URL	ドメイン	カテゴリー	各年における出現回数				順位
				2010	2011	2012	計	
handle.net	http://hdl.handle.net	その他	ポータル	72	55	92	219	1
弘前大学医学部保健学科	http://www.hs.hirosaki-u.ac.jp	ac.jp	大学・研究所等	61	59	77	197	2
大阪市立大学学術情報総合センター データベース ※DB HPからの入口がわかりづらい	http://dlisv03.media.osaka-cu.ac.jp	ac.jp	大学・研究所等	75	65	37	177	3
NCBI(PubMed)	http://www.ncbi.nlm.nih.gov	諸外国	ポータル	38	34	47	119	4
InternetArchive	http://archive.org	org	ポータル	24	18	69	111	5
青山学院大学図書館女子短期大学図書館 ※リポジトリのアドレスがHPドメインと同じ	http://www.agulin.aoyama.ac.jp	ac.jp	大学・研究所等	37	34	26	97	6
google	google	google	ポータル	22	33	40	95	7
山梨大学附属図書館 ※図書館HPの雑誌の目次からリポジトリへリンク ※リポジトリはOPACのアドレスと同じ	http://www.lib.yamanashi.ac.jp	ac.jp	大学・研究所等	28	24	23	75	8
東洋大学	http://www.toyo.ac.jp	ac.jp	大学・研究所等	20	18	37	75	
筑波大学附属図書館 ※リポジトリのアドレスがHPドメインと同じ	https://www.tulips.tsukuba.ac.jp	ac.jp	大学・研究所等	50	9	14	73	9
関西学院大学	http://www.kwansei.ac.jp	ac.jp	大学・研究所等	33	31	8	72	10
大阪国際大学	http://kiyo.oiu.ac.jp	ac.jp	大学・研究所等	29	17	24	70	11
Wiley	http://online.library.wiley.com	com	企業(出版社を含む)	10	32	27	69	12
Agriknowledge/Agropedia	http://agriknowledge.affrc.go.jp	go.jp	ポータル	18	18	29	65	13

よく出現するポータルサイト

- 「よく出現するURLリスト」の中で、“ポータル”と分類されたサイトは9サイト

注)分類の便宜上、handleやDOIの管理機関も含めた。

- この3年の間に、どのサイトも出現回数が増えている。



ILLの全体のログ数が減っており、キャンセルログの数も減っているにも関わらず、ナビゲートログは増えており、ポータルサイトへの案内も増えている。

よく出現するポータルサイト

よく出現するポータルサイト

順位	memo	2010年度 出現回数	2011年度 出現回数	2012年度 出現回数	出現回数 3年間計
1	handle.net	72	55	92	219
2	NCBI(PubMed)	38	34	47	119
3	InternetArchive	24	18	69	111
4	Google	22	33	40	95
6	Agropedia	18	18	29	65
5	Gallica	6	6	9	21
7	HathiTrust			9	9
8	Japan Link Center			9	9
9	DOI (International DOI Foundation)			7	7

よく出現するポータルサイト

- デジタルアーカイブサイトが増加

[InternetArchive/Gallica/HathiTrust](#)

- パーマネントリンクが浸透

[handle.net/Japan Link Center/DOI](#)

- NDL/NII/JST以外によく出現するポータルサイト

[PubMed/Agropedia\(AgriKnowledge\)](#)

よく出現する大学・研究所等リスト

- 「よく出現するURLリスト」から、大学・研究所等のURLのみを抽出し、3年間の出現回数合計が多い順に、リスト化

サイト名	カテゴリー	各年における出現回数				順位	"よく出現するBIBIDリスト"に出現するNCID
		2010	2011	2012	計		
弘前大学医学部保健学科	大学・研究所等	61	59	77	197	2	AA12306943,AA11603444,AA12520987
大阪市立大学学術情報総合センター データベース ※DB HPからの入口がわかりづらい	大学・研究所等	75	65	37	177	3	AA11847609,AA12017447,AN00029633,AN00069015,AA12147442,AA12092227,AN00029688,AN00029724,AA11397750,AA11831493
青山学院大学図書館女子短期大学図書館 ※リポジトリのアドレスがHPドメインと同じ	大学・研究所等	37	34	26	97	6	AA1164957X,AN00057173,AN00008906,AN00008870,AN00009024
山梨大学附属図書館 ※図書館HPの雑誌の目次からリポジトリへリンク ※リポジトリはOPACのアドレスと同じ	大学・研究所等	28	24	23	75	8	AA1180065X
東洋大学	大学・研究所等	20	18	37	75		AN00170380,AA11905975,AA1121883X
筑波大学附属図書館 ※リポジトリのアドレスがHPドメインと同じ	大学・研究所等	50	9	14	73	9	AN1043864X,AN0014895X,AN10302400,AN00121658
関西学院大学	大学・研究所等	33	31	8	72	10	AN00047522
大阪国際大学	大学・研究所等	29	17	24	70	11	AN10110523
神戸女子大学図書館 ※DB 図書館HPに「神戸女子大学紀要検索」システムがあり	大学・研究所等	22	26	14	62	14	AA12385305,AN10482153

よく出現する大学・研究所等リスト

- 非常に需要あり、かつ、探しづらいタイトル持つ大学・研究所のURLのリスト。
- 探しづらいタイトルが複数ある大学も多い
 - 特定のタイトルのみ需要があるわけではなく、全体的に探しづらい場合もあり？
- 出現回数の上位のURLは、非常に頻繁にILLの受付館にナビされている。

独自DB

- 独自のDBが多くランクインしている。独自DBに収録のデジタル化資料は探しづらい？

例)大阪市立大(2位) 神戸女子大学(14位)

自治体国際化協会(17位) 国文研 (28位)

※その他、DBやデジタルアーカイブ等が意外と登場している

→国文研/アジ研/日文研/早大など、デジタルアーカイブを積極的に行っている

6-2. リポジトリのナビゲートログ

リポジトリについて

- ・リポジトリへのナビは13%

→リポジトリは、JAIROに収録されており、検索エンジン対策もとられている。探しやすいはずだが、意外と探せていない。

※リポジトリへの分類について

NII公開の「国内の機関リポジトリー一覧」によった。

このリストにないものは、名称に“リポジトリ”の単語を含むものをリポジトリとして扱った。

リポジトリとOPAC

- リポジトリに収録のデジタル化資料をOPACから探せるようにしている大学もあり。

例) 山梨大(8位) 筑波大(9位)

→学内の人にはわかりやすいが、外部の人はOPACを検索しない？

→山梨大はOPAC以外に、HPからリポジトリへのリンクを別途作成している。

どこからでも探せるようにしている努力が、逆に外部の人にはわかりづらい？

リポジトリとCiNii Articles

- リポジトリを新たに構築した大学もいくつか見受けられた。リポジトリ構築前のデジタル化資料は図書館HPで掲載している場合もあり。

例) 上越教育大学など

- リポジトリのない大学は図書館HPとCiNii Articlesで公開している例も。

例) 健康科学大学など

→大学HPで公開している紀要へは、CiNii Articles (ポータル)で検索しても本文へのリンクがないので注意！

複数のサイトでの公開

- 複数のサイトでデジタル化資料を公開している紀要も多い。
例) 山口医学
リポジトリ+J-STAGE+CiNii Articles
 - 学会誌はJ-Stageでの公開が多い。CiNii Articlesの両方で本文を公開している場合も。
例) Chemical & pharmaceutical bulletin
J-STAGE+CiNii Articles
- 本文の公開は1か所で、他のサイトからはリンクのみという場合もあるので注意。
- ログではどのURLを案内しているかばらつきがある。CiNii Articles/NDL/J-STAGE/リポジトリ/その他の割合は、実際には振れ幅があると思われる。

よく出現するBIBIDリスト

- 3年を通じて4回以上出現するBIBIDを、3年間の出現回数の合計が多い順に、リスト化

BIBID	タイトル	各年における出現回数				順位	TXTL	YEAR	所蔵館数 (2013.8月 現在)	公開URL (複数ある場合は、任意に 1つのURLを選択)
		2010	2011	2012	計					
AA12227144	看護・保健科学研究 ※全国看護管理・教育・ケアシステム研究会	65	50	37	152	1	jpneng	2001	5	高知大学学術情報リポジトリ
AA12306943	弘前大学大学院保健学研究科紀要 = Bulletin of Health Sciences Hirosaki	36	31	27	94	2	jpn	2008	2	"よく出現する大学・研究所等リスト"のURL参照
AN10558109	新潟県厚生連医誌	27	37	19	83	3	jpn	1984	7	新潟県厚生連HP http://www.niigata-kouseiren.jp/
AA11354292	千葉看護学会誌 = Journal of Chiba Academy of Nursing Science	33	23	23	79	4	jpn	1996	49	"よく出現する大学・研究所等リスト"のURL参照
AN10110523	国際研究論叢：大阪国際大学紀要 = OIU journal of international studies	36	17	24	77	5	jpneng	1989	216	"よく出現する大学・研究所等リスト"のURL参照
AA1180065X	山梨大学看護学会誌 = Yamanashi nursing journal	34	23	18	75	6	jpn	2002	4	"よく出現する大学・研究所等リスト"のURL参照
AN00047522	関西学院大学社会学部紀要	33	32	8	73	7	jpn	1960	193	"よく出現する大学・研究所等リスト"のURL参照
AA11603444	弘前大学医学部保健学科紀要 = Bulletin of Health Sciences Hirosaki	25	20	22	67	8	jpn	2002	90	"よく出現する大学・研究所等リスト"のURL参照
AA11289598	聖路加看護学会誌	20	17	22	59	9	jpn	-199	68	聖路加看護大学リポジトリ
AA11818837	滋賀医科大学看護学ジャーナル = Journal of nursing, Shiga University of Medical Science	29	17	13	59		jpn	2003	85	"よく出現する大学・研究所等リスト"のURL参照
AN10375470	川崎医療福祉学会誌 = Kawasaki medical welfare journal	21	18	16	55	10	jpn	1991	231	"よく出現する大学・研究所等リスト"のURL参照
AA12290923	島根県立大学短期大学部出雲キャンパス研究紀要	8	15	30	53	11	jpneng	2007	91	"よく出現する大学・研究所等リスト"のURL参照

よく出現するBIBIDリスト

- 看護学に関するタイトルが多い。
→全体的に看護学に関する文献依頼が多いため、相対的にナビも多くなっている？
- 探しづらいものと探しやすいものが混在しているように見受けられる。探しにくいタイトルだけがランクインしている訳ではない？
- NII/NDL/JST/リポジトリ/それ以外から万遍なく登場している。

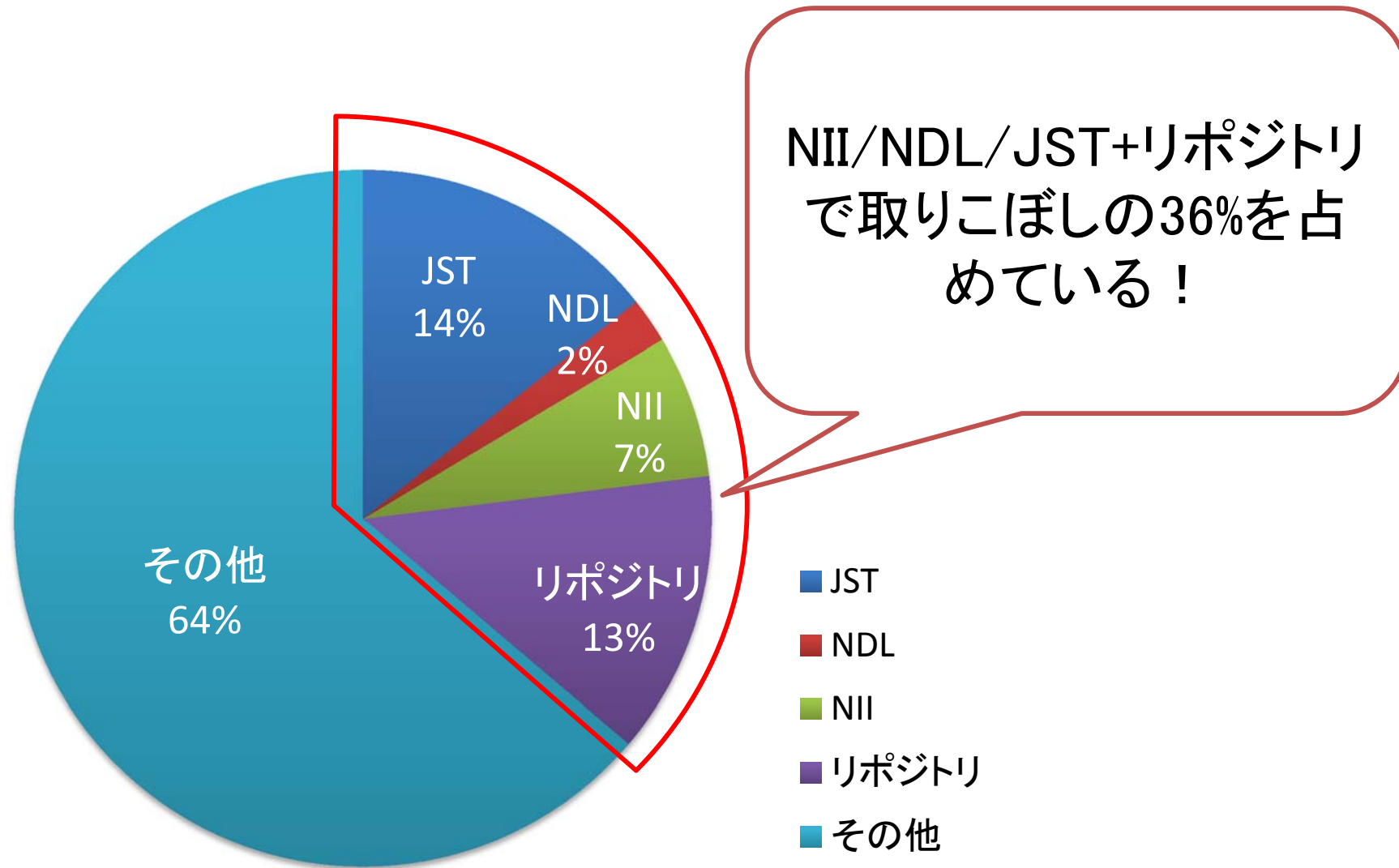
よく出現するBIBIDリスト

- リストを見て、即座にどのサイトを探すべきか答えるのはかなり難しいのでは？
- 電子版がないものと交じってILL業務をしている場合、電子版があることを想定できないと思われる。
→実際の有無に関わらず、依頼前に電子版を確認する業務フローが必要？

7.調査報告④

NII/NDL/JSTナビゲートの分析

再掲：2012年度ナビありログの内訳



取りこぼしの原因を考える

- リポジトリがサービス担当者にまで十分に認知され、無料Web公開資料の存在も相当浸透してきている2012年度においても、NII/NDL/JST+リポジトリで**36%**も取りこぼしてしまっている。
- きちんと探せば見つかるものが多数→それぞれのポータルで何を探せて何を探せないのかを理解できていないのが原因か？
- NII/NDL/JSTのそれぞれが直接つながる無料のデジタル化資料について整理する。

調査方法

- 基本的に、NII/NDL/JSTの各機関のHP上の情報をもとにまとめた。
- 調査上必要でHPの情報では不足している部分については、CiNii Articles/NDL Search/J-GLOBALの各ポータル担当者にメールにて質問した。

【NII】CiNii Articlesから直接つながる 無料デジタル化資料

- NII-ELS学協会/研究紀要収録論文
 - 学協会は有料or定額アクセスあり、アイコンで分かる
- JAIRO
 - CiNii連携申請していて、かつ、論文タイプがJournal Article/Departmental Bulletin paper / Articleのもの
- J-STAGE
 - 提供元学協会が連携拒否していないことが前提。有料/無料両方にリンク、アイコンは同じ。
- JASI(日本農学文献記事索引)
 - JASI採録論文のうち本文無料公開されているものに対してのみリンク
- CrossRef
 - リンク先が無料かどうかは関係ない

【NII】CiNii Booksから直接つながる 無料デジタル化資料

現時点での可能性としては、NDLサーチへのリンクアウトからの発見くらいか。

(グループ2の今回のWSでの検討が今後に大きく寄与しそう)

【NII】GeNiiから直接つながる 無料デジタル化資料

- CiNii Articles
- JAIRO
 - 横断検索だからJAIROすべてヒット
- KAKEN

→検索対象としてはCiNii Articlesと比較して、
JAIROの差分とKAKENが追加

【JST】J-GLOBALから直接つながる 無料デジタル化資料

まず、前提としてJ-GLOBALの作りをとらえる

「文献」カテゴリーの情報:検索結果レコードとしてヒットするもの(外部連携DBのコンテンツは直接はヒットしない)

1. J-DREAM3を構成するファイルのうち、JSTが作成している文献情報JSTPlus,JMEDPlus,JST7580, JSTChinaから、NTIS提供情報を抜いたもの
2. Web of Scienceの日本人著者の情報

⇔CiNii Articlesは文献情報自体を外部からももらってマージしている

J-GLOBALからたどれる無料デジタル資料はこの2つの条件を両方満たすものと考えられる

- ・J-Global「基本情報」の「文献」カテゴリーに収録(大前提)
- ・外部連携対象になっているDBに含まれるもの

【JST】J-GLOBALから直接つながる 無料デジタル化資料

- J-STAGE
- Agropedia(JASI)
- 医学薬学予稿集全文DB
- PubMed
 - リンク先が無料かどうかは関係ない
- CrossRef
 - リンク先が無料かどうかは関係ない
- そのほか特許情報も多数

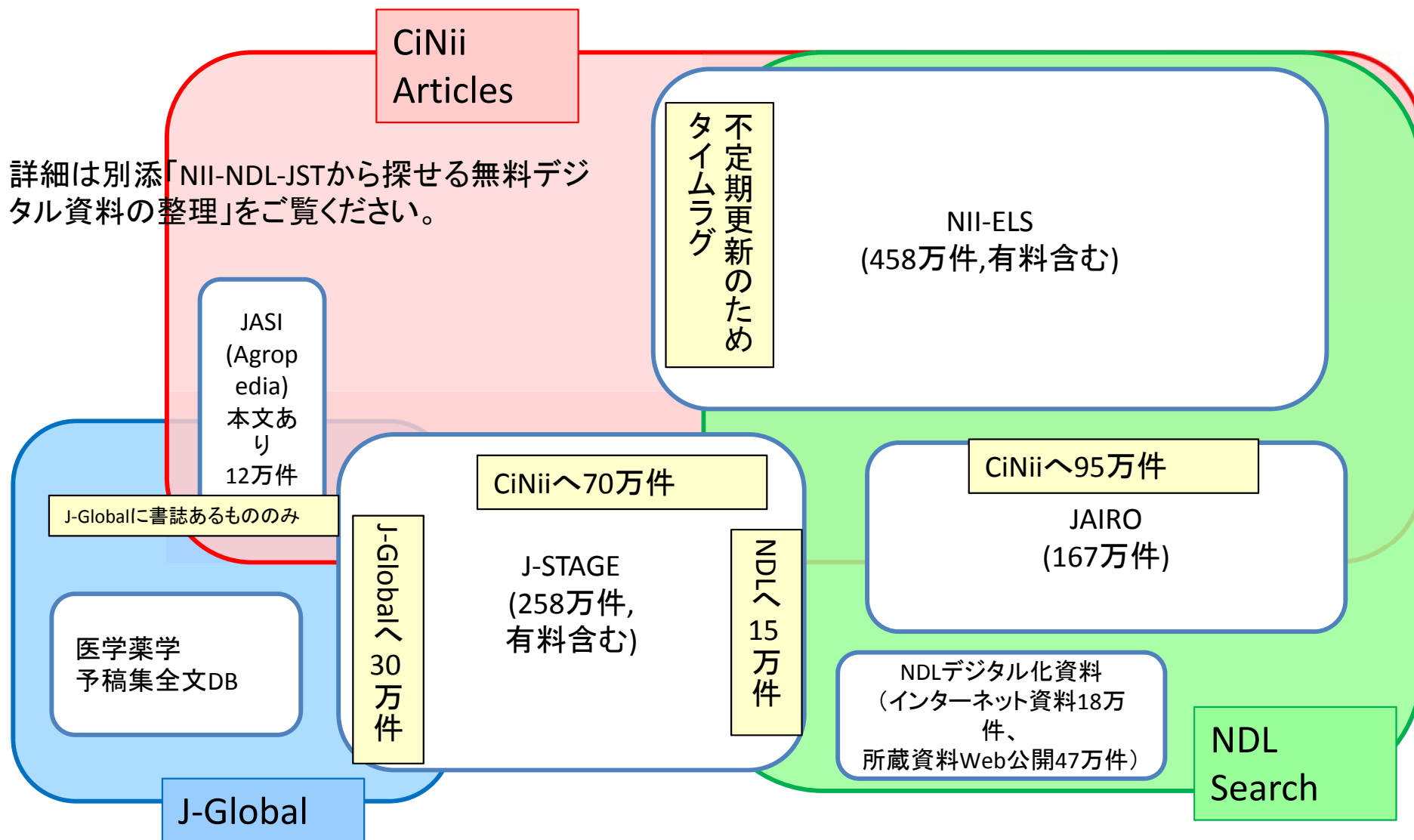
※J-STAGEにもつながるが、一部分であることに注意

※GeNiiとは連携しているが、文献情報(CiNii Articles)は連携なし

【NDL】NDL Searchから直接つながる 無料デジタル化資料

- NDLホームページ掲載刊行物、CA
- デジタル化資料：インターネット資料
 - WARPで集めたもののうち、図書/雑誌/論文記事
- デジタル化資料
 - 2013.11現在、47万件がWebで公開
- NII-ELSおよびJAIRO
 - 基本的に全て。NII-ELSは不定期更新のためタイムラグあり。JAIROは週次更新
- J-STAGE
 - 提供元学協会が連携拒否していないことが前提。今年度末より定期更新になる予定。
- その他さまざまなデジタルアーカイブと連携。横断検索分も調べたいときはチェックを入れるのを忘れずに
- 国会の議事録や法令などNDLならではのコンテンツも多数

NII/NDL/JSTの 主な国内無料デジタル化資料の収録モード図



NII/NDL/JSTの収録関係から 読み取れること

- どのポータルもワンストップポータルではない。あるDBの一部を収録している、というのが勘違いのもとか。
- JAIRO搭載コンテンツをまとめて検索できるのはJAIRO自体もしくはNDL Search。機関リポジトリコンテンツを探すときの要注意ポイント。CiNii Articlesで全部探せるわけではない。
 - IRからCiNii Articlesへの3段階があることを知る
 - IRDBにうまくハーベストされる→CiNii連携申請がされている→コンテンツタイプ(NII-Type)が条件に合致する

調査結果③との関連性

- 無料デジタル化資料へのナビゲートURLの分類は、受付館がナビゲートしたURL・サイト名をもとに分類している
- 無料のデジタル化資料は複数の場所で提供されているケースも多い。大学HPやリポジトリへナビゲートしているデジタル化資料も、実はNII/NDL/JSTに収録されている部分があると推測される。

簡易標本調査

下記のような方法で標本調査を行った

【対象】

- NII/NDL/JST/リポジトリ以外に分類されたもの
- そのうち、国内のもので、かつ、個々の大学や学協会サイトで公開されているもの
- 2010/2011/2012のどこかの年度でドメイン単位の延URL数TOP20に入ったもの

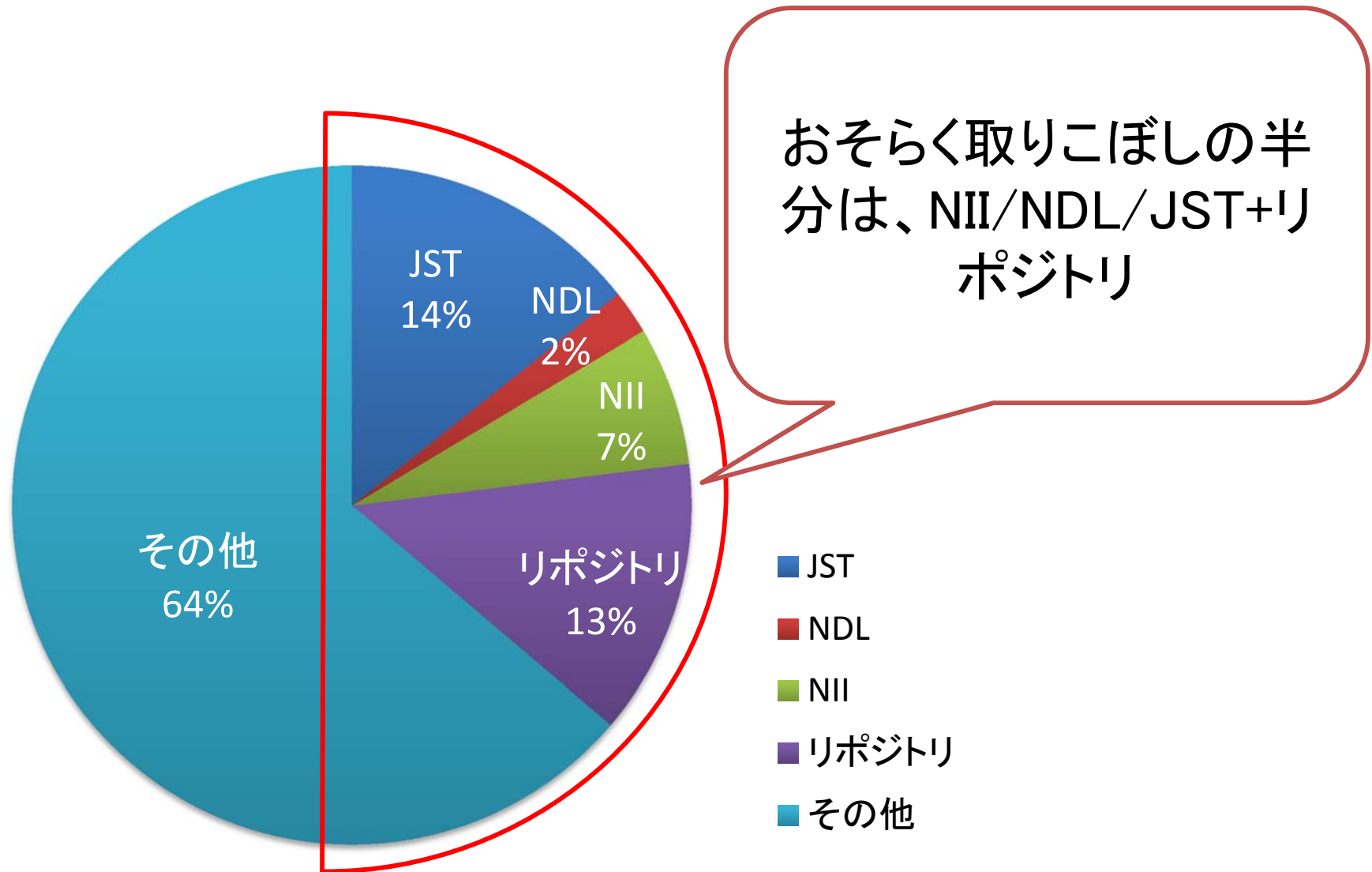
【調査方法】

- 3年分からランダムに5個URLを抽出し、CiNii Articles, JAIRO, J-Globalで検索
- 無料の本文へ直接リンクするものをカウントした

標本調査の結果

- ナビゲート当時と現在ではポータルへの収録状況も違うことはあらかじめ念頭におきます
 - 各ポータルでの本文あり無料デジタル化資料への到達状況
 - CiNii Articles 32.5%
 - JAIRO 35.8%
 - J-Global 0%
- 現状では3割強がCiNii Articlesから見つかる

標本調査の結果から推測されること



小括

- J-Global/NDL-Search/CiNii Articles＋J-STAGEを全て検索すれば、国内の多くの無料のデジタル化資料の取りこぼしをかなり無くせそう。
- それに加えて検索エンジンで検索すれば、上記の漏れもかなり拾えるのでは。
- とはいえ、時間がないときなどはそんな悠長なことはしてられない。J-Global/NDL-Search/CiNii Articles＋J-STAGEをリアルタイムでまとめて検索できる仕組みがあるとよさそう。さらに誌名・巻号年頁など基本的な項目で条件づけられるインタフェースがあればなおのことよさそう。

8.まとめと今後に向けて

ポータルを知る・理解する必要性

- ポータルの「罨」: 検索者側の思い込みが原因
 - いくつかの大規模ポータルをきちんと探す必要がある
 - ポータルの収録内容やポータル間の関係を理解せず、「ここで全部探せる」と思いこむと、検索漏れが出る
 - 書誌データと本文リンクは別。検索結果の書誌からすべての本文にリンクアウトできるわけではない。
- 海外のポータルやデジタルアーカイブサイト
 - InternetArchive/Google/Gallica/HathiTrustなど、多数登場している
 - そもそもポータルの存在を知らないことも。そして知っていても収録内容やポータル間の関係が分からない
 - 今回はできなかったが、きちんとした整理が必要。整理すれば連携の道が見えてくるのでは。

ポータル提供者のみなさまへの提案

- 「収録している」DBのコンテンツのうち、何がどこまで入っているのかの**情報公開**を。
 - 連携方法(データ取得方法)
 - 収録DBと検索結果の関係性
 - 更新頻度
 - 部分的にしか収録していない場合は、どういう条件で絞られているのか明示する
- ポータルのWeb APIの公開などで利用の自由度が向上している。収録内容を詳しく公開することで「**何を検索できて、何を検索できないのか**」が明確化され、さらに利用の幅が広がることが期待できる。

個々のデジタル化資料の提供者の みなさまへの提案

- メタデータやDB/サイト構造の改善の必要性
 - googleにハーベストされているけれど順位が低い、というメタデータの問題
 - 雑誌1号や図書1冊がまるごとPDFになっているというファイル単位の問題
 - 外への開放性のないDBにコンテンツを入れているというサイト構造の問題
- 提供ページの画面構成や表示の改善の必要性
 - せっかくデジタル化資料を公開したのにどこから本文にいけるかわからない。一目でわからないとユーザーは探索行動をあきらめてしまう。せっかくの公開を無駄にしないように、**サイトのユーザビリティは重要ポイント**。

図書館員のみなさまへの提案

- NII/NDL/JSTの各ポータル[○]の収録コンテンツはよく理解しておいたほうがよさそう
 - 何となくの把握では、思わぬ検索漏れの可能性あり。
- googleなどの検索エンジンでも1回調べるのがよい
 - ポータルの収録対象になっていないコンテンツや、出版者サイト内でも論文単位OAなど、電子ジャーナルリストやリンクリゾルバのKBでは網羅できてないコンテンツがまだまだたくさんある。
- これはデジタル化資料無さそう、という勝手な思い込みは禁物
 - 意外なところで意外なデジタル化資料が！

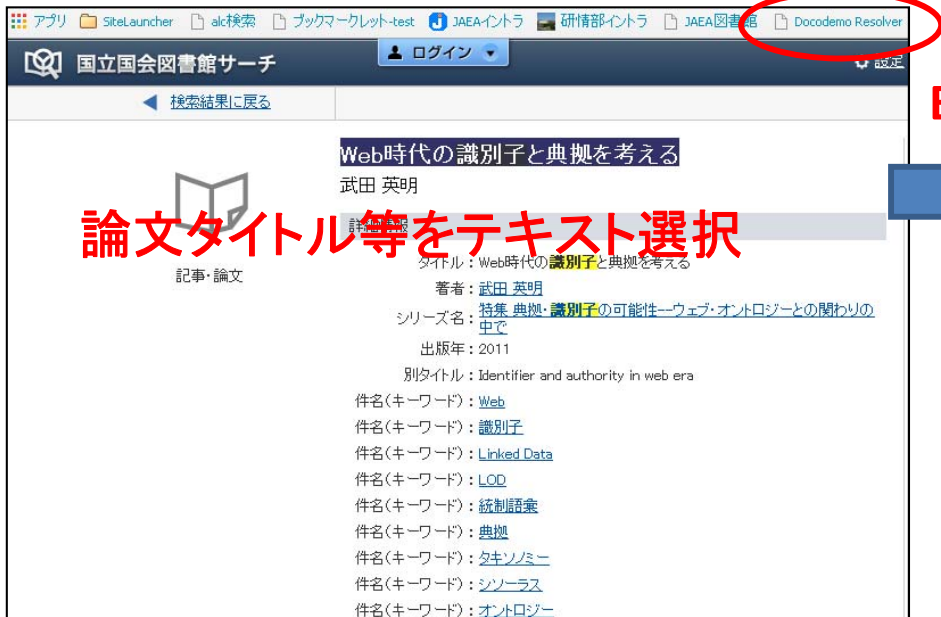
今後の展望

—調査からつなげられる具体的なアクション—

1. 具体的な項目とともに、ポータルに収録している情報の公開を、NII/NDL/JSTに依頼。また、3者の収録関係をまとめたものを“国内ポータルコンテンツ収録マップ(仮)”として公開。
2. 見つかりにくい文献情報(ナビゲートログによく出現する書誌・URLリスト)のデータを公開。ブラウザ用リストと機械可読用リスト2種類。
3. (可能であれば) NII/NDL/JST以外のナビゲート部分のパターン分けを精査し、コンテンツ提供者への改善提案を作成し公開する。

試行：課題解決のための具体的な仕組み例

- ポータルを横断する仕組み (Docodemo Resolver β)
 - CiNii Articles, NDL Search, J-STAGE, arXiv, JAIRO
 - Bookmarkletで起動しPHPで処理
 - iPadやiPhoneでも



Bookmarkletを起動



中間窓に各種ポータル
の検索結果が表示