OUKA
The University of Osaka
Institutional Knowledge Archive

Title	NACSIS-ILLログ分析に基づく無料デジタル化資料に関する考察
Author(s)	藤江, 雄太郎; 小島, 由香; 長屋, 俊
Citation	大学図書館研究. 2015, 102, p. 34-43
Version Type	VoR
URL	https://hdl.handle.net/11094/55627
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

NACSIS-ILL ログ分析に基づく無料デジタル化資料に関する考察

藤江 雄太郎，小島 由香，長屋 俊

抄録：今日世界中で数多くのデジタル化資料が無料公開されている。本稿では，これらの資料の発見性を高める方策を検討するため，NACSIS-ILL のログに注目し，分析を行った。ILL ログのうち，無料で公開されているデジタル化資料が利用できるという理由で謝絶に至ったログを抽出，案内されている資料を調査した。その結果，無料デジタル化資料は Web 上の広い範囲に散在し，特に国内では 60% 程度が NDL-Search・CiNii Articles・J-GLOBAL・J-STAGE の各サイトに加えて検索エンジンで検索すれば発見可能な資料であることがわかった。また，これらのサイトはメタデータ連携が部分的で，収録状況の詳細が明示されていないことがわかり，発見性低下の一つの要因になっていることが示唆された。

キーワード：無料デジタル化資料，NACSIS-ILL，国立情報学研究所，学術情報システム総合ワークショップ，機関リポジトリ，図書館間相互貸借，ログ分析，オープンアクセス，デジタル化

1. はじめに

ウェブの爆発的な広がりとともに，従来は紙で流通していた図書，学術論文，雑誌記事，レポートなどの様々な資料がウェブ上で公開されるようになった。これらには，もともと紙資料として流通していたものを遡及デジタル化したものや，ボーンデジタルのものがある。また，有料で公開されているものだけでなく，無料で公開されているものも数多い。

従来の図書館では，OPAC や書誌情報データベースをはじめとした検索ツールやサービスを整備し，利用者へ資料の目録および所在情報を提供してきた。近年では紙資料の情報に加えて，電子ジャーナルや電子ブックなど，購入している電子資料の情報を統合的に提供する図書館も増加してきている。しかし，ほとんどの図書館において，現状では有料のデジタル化資料の情報提供が中心で，ウェブ上で無料公開されているデジタル化資料の情報を従来の検索ツールに結び付けて提供するまでには至っていない[1)]。これはすなわち，OPAC や書誌情報データベース上で検索した資料がウェブ上で無料公開されていても，OPAC や書誌情報データベースからは直接関連付けられていないため，利用者が気付きづらい状況にあるということを意味している。

このように，ウェブ上で無料公開されている学術情報については，効果的に利用者に情報が提供されていないという現況がある。本稿では，今後の図書館サービスにおける無料デジタル化資料の提供方法や利用方法の改善に資することを目的として，無料で提供されているデジタル化資料（以下，「無料デジタル化資料」という）に焦点をあて，現状の問題点の把握を目指す。

具体的な調査対象としては NACSIS-ILL のログに注目した。NACSIS-ILL は，NII（国立情報学研究所）が参加館同士の図書・論文等の複写・貸借のために提供しているサービスである。この NACSIS-ILL のシステム上のやりとりにおいて，受付館の担当者がウェブ上で公開されている無料デジタル化資料の存在を発見，依頼館に案内し，謝絶に至るというケース（以下，ILL においてウェブ上の無料デジタル化資料を案内する行為を「ナビゲート」という）がある。これらのケースは「ある資料を必要としていて，ILL の依頼者及び依頼館で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ことが受付館でわかった」という情報探索行動の過程と結果が分析可能な記録として残っている事例であるといえる。これらの事例を分析することで，探しづらい無料デジタル化資料の傾向を知り，現状の問題点の把握を行うことができると考えられる。具体的には，受付館が依頼館へナビゲートし謝絶になった ILL ログからナビゲートされている無料デジタル化資料を抽出，それらの資料の傾向の調査を行った[2)]。

なお，本稿は NII で開催された研修「平成 25 年度学術情報システム総合ワークショップ（試行）」の調査結果をまとめたものである。調査結果は，2013 年 12 月現在のデータに基づいている。状況は刻々と変化しており，現況は本稿の報告と異なる点もある。読者の皆様には最新の情報をその都度確認していただきたい。

2. 調査に使用したログと分析用ログの抽出方法

調査に必要な特定の項目について ILL ログの提供を NII に依頼した。これは，学術情報システム総合ワークショップという研修の枠組みにおいて提供を受けたものである。個人情報にあたる部分は対象から予め外した。

調査対象としたログの範囲は最終更新日が 2010 年 4 月 1 日から 2013 年 3 月 31 日の計 3 年分のログで，ILL が終了しているログ，具体的には，状態が“確認”/“返却確認”/“CANCEL”となっているログ[3] とした。提供されたログの件数は表 1 のとおりである。なお，NII 公開の NACSIS-ILL レコード件数[4] とは若干誤差がある。

これらのログから，ILL 受付館が依頼館に対してデジタル化資料の存在を案内し謝絶しているケース，つまりナビゲートしているコメントが含まれた事例のログを抽出した。抽出した手順を以下に示す。

（手順 1）NACSIS-ILL で，受付館が「無料のデジタル化資料がウェブ上で公開されている」というメッセージをコメントに記載し依頼館に送る際には，必ず“INQUIRE”もしくは“PARDON”のどちらかのコマンドが発行されるため，これらのコマンドが含まれるログ（以下，「ログ集合 1」という）の抽出を行った。

（手順 2）コメントに URL が記載されている場合，デジタル化資料へのナビゲートが行われている可能性が高いため，まずログ集合 1 から，コメントに文字列として「http://」もしくは「https://」を含むログを抽出し[5]，新たなログ集合（以下，「ログ集合 1-1」という）を作成した[6]。

（手順 3）コメントに URL ではなく，例えば「CiNii Articles で公開されている」など，サービス名等を示してナビゲートするケースもある。まず，コメントにどのような単語が含まれているかを調べるためログ集合 1 のコメント部分に対して形態素解析を行い，抽出された単語から，目視確認によりナビゲートしていると推定される単語を選定した。具体的には，「CiNii」や「リポジトリ」といったサービス名称や，「全文」や「無料」といったデジタル化資料へのナビゲートを連想する単語を拾い出し，単語リストを作成した。続けて，単語リスト中の単語がコメントに含まれるログを，ログ集合 1 から抽出し，新たなログ集合（以下，「ログ集合 1-2」という）を作成した[7]。なお，形態素解析には MeCab[8] を利用し，辞書には IPADIC[9] を用いた。

（手順 4）ログ集合 1-1 とログ集合 1-2 のログについて，URL やサービス名がコメントに含まれていても，ナビゲートではないコメントがあるため，コメントの文脈を目視確認し，ナビゲートしているコメントかどうか判定を行った。そして，ナビゲートをしていると判断されたログ（以下，「ナビゲートログ」という）を抽出した。

（手順 5）ナビゲートログのうち，ログ集合 1-2 由来のログについては URL の記載がないため，URL の記載のあるログと同じ扱いで分析するために，便宜的に，URL を調査・推定して追加した。（ただし，一部のサイトについては URL を推定せずサービス名のまま分析を行ったものもある。）

ナビゲートログの件数は，**表 1** のとおりであり，最終状態が CANCEL のものがログ全体の件数に占める割合は，2012 年度では 6.74％である。

表 1　ナビゲートログの件数と割合
（ ）内はログ全体に占める割合である。

年度	2010 年度	2011 年度	2012 年度
ログ全体の件数	946,140	911,504	873,810
最終状態が CANCEL のログの件数	61,572 (6.51%)	59,814 (6.56%)	58,866 (6.74%)
ナビゲートの件数	3,192 (0.34%)	3,344 (0.37%)	3,946 (0.45%)

3. ナビゲートログの妥当性の検証

ナビゲートログを用いた分析を行うにあたり，まずはナビゲートログを分析に使用する妥当性について検証を行った。以下に検証結果を報告する。

3.1　依頼館 / 受付館分析

ナビゲート行為が特定の参加館に偏っているとすると特定の参加館固有の問題と考えられる可能性があることから，ナビゲート行為が特定の参加館に集中していないかどうか検証を行った。

検証には，2010 年度～2012 年度の 3 年間のログのうち，ナビゲートログおよび成功した ILL のログ（最終的な状態が“確認”もしくは“返却確認”のログ）を用いた。成功した ILL のログについて，NACSIS-CAT/ILL の参加組織 ID 単位で，依頼館別および受付館別にログ件数を集計した。また，ナビゲートログについても，依頼館別および受付館別にログ件数を集計した。

3.1.1　ILL 依頼館単位の集計

3 年分のログから，依頼館ごとに被ナビゲートログの件数を集計した。その結果，3 年間で被ナビゲート 0 件の依頼館が 23%，0～3 件の依頼館が

53%であり，0〜21 件で 90%を越えた。8 割弱の依頼館がナビゲートされるという経験をしていることがわかる。

依頼館の参加組織 ID 単位における成功依頼件数および被ナビゲート件数の散布図は**図 1** のとおりであり，依頼館の参加組織 ID 単位における成功依頼件数および被ナビゲート件数の順位相関係数を算出したところ，$\rho = 0.815$（$p < 0.001$）と高い正の相関が確認された[10]。このことから，ナビゲートは特定の依頼館に偏ることなく全般に行われていると言える。

3.1.2 ILL 受付館単位の集計

3 年分のログから，受付館ごとにナビゲートログの件数を集計した[11]。その結果，ナビゲート件数が，0 件の受付館が 40%，0〜3 件の受付館が 67%，0〜9 件で 80%であった。ナビゲート件数が 3 年間で 1 桁という受付館が全体の 8 割を占める一方で，最も多い機関では 300 件を超えるナビゲートがあった。ただし，成功受付件数が多くてもナビゲート件数が多いとは限らず，成功受付数が少ないにも関わらずナビゲート件数が多い機関もあり，ナビゲートを ILL の業務フローに組み込んでいる機関とそうでない機関が分かれている可能性がある。

一方で，受付館の参加組織 ID 単位における成功受付件数およびナビゲート件数の散布図は**図 2** のとおりであり，受付館の参加組織 ID 単位における成功依頼件数および被ナビゲート件数の順位相関係数を算出したところ，$\rho = 0.651$（$p < 0.001$）と比較的高い正の相関が認められた[10]。このことから，ナビゲートは受付館毎の際立った偏りはみられないといえる。つまり，ナビゲートの業務フローへの組み込みについては機関によってやや散らばりがある可能性があるが，全般的にはどの機関においてもナビゲートは行われているようである。

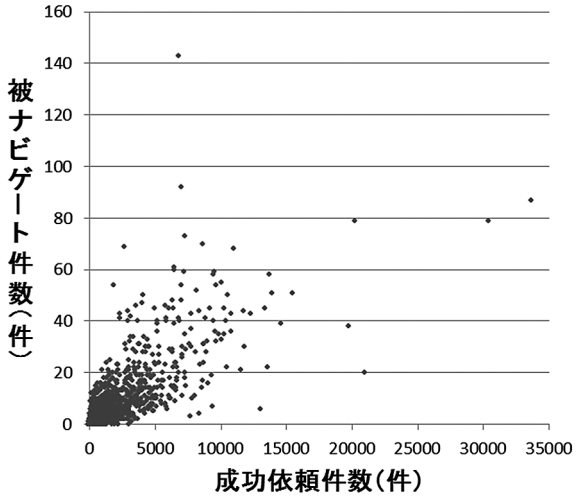

図 1　参加館ごとの成功依頼件数と被ナビゲート件数の散布図
2010〜2012 年度の累計で，該当参加館数は 1363 館である。

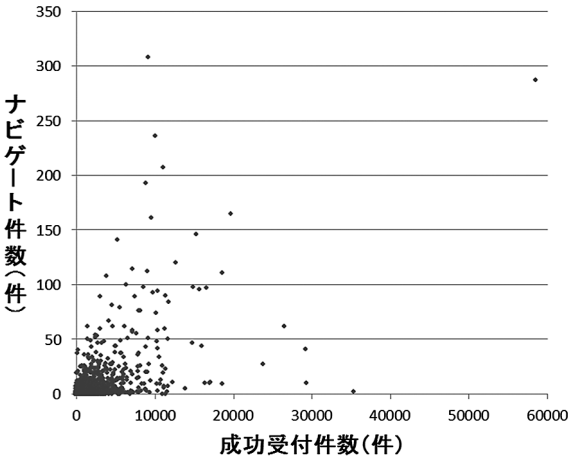

図 2　参加館ごとの成功受付件数とナビゲート件数の散布図
2010〜2012 年度の累計で，該当参加館数は 1230 館である。

3.2 BIBID による分析

ILL を依頼する際には，NACSIS-CAT の書誌 ID である NCID を，BIBID という項目に入力して依頼する。ナビゲートログについて，NCID の 3 年間の出現回数の多い順に「よく出現する BIBID リスト」[12] を作成した。（なお，ナビゲートログの BIBID に NCID が入力されている割合は，各年度とも 96%〜97%程度であった。）

全体的に，看護学に関するタイトルが多くランクインしている。佐藤（2007）[13] によると，ILL の和雑誌の複写依頼の件数が年々増え，2005 年度の「上位の和雑誌の多くが看護学の関連領域のタイトルで構成されている」とあるが，その傾向は現在に至るまで続いているようである[14]。ILL ログ全体に占める看護学関連領域のタイトルの割合が大きいため，ナビゲートログに看護学関連領域のタイトルが多く含まれることは，全体の傾向が反映された結果ともいえる。

このリストには，NII，JST（科学技術振興機構），機関リポジトリ（以下，「リポジトリ」という），大学等の各サイトで提供されているデジタル化資料が多数含まれている。後述するが，これらのサイトの収録資料を検索対象とするサービスは数多く，利用者及び依頼館は，これらの資料を探すことができなかったのではなく，探していないという可能性が推察される。ILL 業務では電子版がない資料

を多数扱っており，電子版資料があることを予め想定しているわけではない。そのため，探しやすいと思われるデジタル化資料についても，ILL の依頼前に電子版を検索するという業務フローがない場合，その存在に気づかず ILL の依頼をする可能性もある。

以上を踏まえ，本調査の分析対象としては必ずしも見つけづらい資料に限らず，ニーズが非常に高く，かつ，実際に多くの人に見逃された資料が含まれているという前提で調査を進める。受付館がデジタル化資料の存在に気がつかずに依頼館の依頼どおりに ILL 業務を滞りなく行うケースも多数あると考えられるが，これらのケースについては，本調査では分析を行っていない。

また，ILL システムにおける「複写 / 貸借区分」の「複写」「貸借」の両方のデータが含まれているが，ILL 全体に占める「複写」の依頼件数[15] が圧倒的に多く，今回の調査の分析対象のメインとなる資料は学術論文・雑誌記事であり，図書資料があまり含まれていないという点について，予めお断りしておく。

4. ナビゲートログの分析

4.1 全体の集計結果

ナビゲートログがログ全体に占める割合は，**表 1** のとおりである。ILL 全体の件数と最終状態が CANCEL のレコードの件数が減少傾向にも関わらず，ナビゲートログの件数は年を追うごとに増加傾向にあることがわかる。

ナビゲートされている URL について，ドメイン数を集計した。（以下、「URL の，“http://”もしくは“https://”から最初の“/”までの部分」を「ドメイン」と扱う）3 回以下しか出現しないドメイン数が全ドメイン数の 79.98% を占めており，ウェブ上の様々なサイトで無料デジタル化資料が公開されていることがわかった。

その一方で，何度も出現するドメインがあった。JST，NDL（国立国会図書館），NII のサイトにあたるドメインである。加えて，リポジトリと思われる URL が比較的多く見受けられたため，リポジトリの出現件数を別途集計することにした。なお，NII の「学術機関リポジトリ構築連携支援事業機関リポジトリ一覧」[16] に掲載されている URL と一致するサイトをリポジトリのドメインとして扱い，まとめて件数を集計した。JST，NDL，NII，リポジトリ，その他のドメインの件数は**表 2** のとおりである。

2012 年度については，JST，NDL，NII の 3 機関のドメインが占める割合は 23.03%，リポジトリの割合は 13.19% で，3 機関のドメインとリポジトリがナビゲートログに占める割合は，36.22% となった。

その他のドメインについては，トップレベルドメインによって集計を行った。（国別コードが「.jp」となっているものについては，セカンドレベルドメインで集計した。）その結果が**表 3** である。国別

表 2　ナビゲートログにおけるカテゴリ別件数
（ ）内は合計に占める割合である。1 つのコメントに複数 URL があった場合は複数の URL をカウントしたため，表 1 の件数と差分が発生した。URL ではなくサイト名でのナビゲートを行ったログについては，複数のサイト名があった場合，任意に 1 つの URL を推定している。

年度	2010 年度	2011 年度	2012 年度
JST	304 （9.51%）	490 （14.59%）	577 （14.33%）
NDL	21 （0.66%）	52 （1.55%）	84 （2.09%）
NII	286 （8.95%）	192 （5.72%）	266 （6.61%）
リポジトリ	649 （20.31%）	559 （16.64%）	531 （13.19%）
その他	1,935 （60.56%）	2,066 （61.51%）	2,568 （63.79%）
計	3,195	3,359	4,026

表 3　「その他」カテゴリの内訳件数
表 3 のうち，単に「jp」と分類しているのは，国別コードが「.jp」のドメインのうち，「http://www.waseda.jp/」のように，セカンドレベルドメインでの分類ができないドメインである。

	2010	2011	2012	総計	2012 年の構成比率［%］
ac.jp	1,066	996	1,019	3,081	39.68
co.jp	26	26	43	95	1.67
com	99	150	234	483	9.11
edu	36	38	49	123	1.91
go.jp	91	104	109	304	4.24
google	22	33	40	95	1.56
gr.jp	9	4	25	38	0.97
jp	105	112	199	416	7.75
lg.jp			3	3	0.12
ne.jp	11	22	25	58	0.97
or.jp	78	123	178	379	6.93
org	157	182	287	626	11.18
その他	85	88	133	306	5.18
諸外国	150	188	224	562	8.72
計	1,935	2,066	2,568	6,569	100

コードトップレベルドメインがjp（日本）となっているものは，2012年度では62.34%となり，多数を占めていることがわかった。また，その他のうち，「ac.jp」の占める割合は，2012年度では39.68%となった。つまり，ナビゲートログのうち，NII，NDL，JSTと，リポジトリを含むセカンドレベルドメインが「ac.jp」のサイトが占める割合を合計すると，2012年度では61.53%となった。このことから，大学や研究機関等のサイトで公開されている無料デジタル化資料へのナビゲートが，ナビゲートログの約60%を占めているということがわかった。

4.2 カテゴリ別の集計結果

前述のとおり，ナビゲートされているサイトはJST，NDL，NII，リポジトリ，その他と，大きく5つのカテゴリに分けることができた。以下でカテゴリごとに見られた傾向の詳細を述べる。

（1）JST

ドメインがJSTとなっているサイト（サブドメインが異なるドメインも含む）の内訳はほとんどがJ-STAGE（旧 Journal@rchive を含む）であった。ナビゲートログのうち，2012年度では14.33%を占め，ナビゲートされる割合が比較的多いサイトという結果になった。

個別にタイトルを見ていくと，J-STAGEで公開されているタイトルには学会誌が多く含まれていた。中にはCiNii Articlesと両方でフルテキストを公開しているタイトルもあった。一方で，フルテキストはJ-STAGEでのみ公開し，CiNii ArticlesにはJ-STAGEへの外部リンクのみがあるというタイトルもあった。なお，CiNii ArticlesにはJ-STAGEで公開されている無料デジタル化資料への外部リンクがすべて網羅されているとは限らないという点は注意が必要である。

（2）NDL

ドメインがNDLとなっているサイト（サブドメインが異なるドメインも含む）の内訳は，国立国会図書館デジタル化資料（現：国立国会図書館デジタルコレクション）と近代デジタルライブラリーが全体の6割，NDLのホームページ上のコンテンツが2割弱，そのほかNDL SearchやNDL OPACの検索結果ページを示しているものなどであった。ナビゲートログに占める割合は，2012年度では2.09%であった。

（3）NII

ドメインがNIIとなっているサイト（サブドメインが異なるドメインも含む）の内訳は9割がCiNii Articlesのコンテンツであり，残り1割はJAIROなどであった。ナビゲートログに占める割合は，2012年度では6.61%であった。

個別にタイトルを見ていくと，紀要など，大学等の学内刊行物が公開されている場合が比較的多く見受けられた。CiNii Articlesにフルテキストを搭載しているものの，自機関のサイトには掲載しない大学等がある一方で，CiNii Articlesと大学等のサイトの両方にフルテキストを搭載するケースや，CiNii Articlesにフルテキストを搭載して大学等のサイトにCiNii Articlesへのリンクを作成するケースなど，自機関のサイトと連携していることもある。

同じ雑誌タイトルであっても，すべての巻号を1か所のサイトで公開しているとは限らない。大学等のサイトにはすべての巻号のフルテキストへリンクが用意されていることが多いが，CiNii Articlesには大学等のサイトで公開している巻号の論文情報が収録されていないことがある。また，NDLの雑誌記事索引などのほかの書誌情報データベースから提供を受けた情報をもとに作成されたレコードについては，CiNii Articlesにフルテキストを搭載せず，外部リンクも用意されていないことがある。このように，大学等のサイトでのみ公開されている学内刊行物については，CiNii Articlesからフルテキストへ繋がっていない場合があり利用者が見落としてしまうという可能性がある。

（4）リポジトリ

リポジトリと分類されたサイトの件数は，**表2**のとおりである。3年間を通じて件数は減ってきているが，ナビゲートログに占める割合は2012年度では13.19%であり，リポジトリへのナビゲート行為はある程度存在していることが明らかになった。リポジトリの書誌情報はJAIROに収録されていることが多く，検索エンジン対策もとられているので，筆者らは比較的見つかりやすい資料だと予想していたが意外な結果となった。

また，リポジトリ以外に，CiNii Articlesを含む複数のサイトで刊行物を公開している場合もある。この場合には，検索エンジンでの検索結果が，1か所のサイトに集約されないことがある。それ以外のケースとして，特定のタイトルについて発行元大学等でリポジトリへの掲載の運用をしていないが，著者の所属大学等のリポジトリに著者稿が掲載されて

おり，一部の論文のみがリポジトリに収録されている場合がある。

このように，大学等の事情によって様々なケースがあることを想定しておく必要がある。

（5）その他

ナビゲートログのうち，上述の（1）～（4）に含まれないサイトについてドメイン単位でグループ化を行い，出現回数を集計した。**表 4** にあるとおり，2012 年度では 82.48％が 3 回以下しか出現しないドメインとなり，全体の約 8 割を占めていることがわかった。そのため，ここではいずれかの年度内で 4 回以上登場するドメインのみを分析対象とし，3 年間の出現回数を集計した。

出現回数の多いドメインをリスト化したものが，「よく出現する URL リスト」[17] である。いくつかのカテゴリに分けられる傾向があったため，分類した結果が**表 5** である。約半分が「大学・研究所等」，四分の一が学会や政府機関を含む「各種団体」，残る四分の一がその他である。

そのうち，「大学・研究所等」のカテゴリに分類されたドメインについてのみ，3 年間の出現回数の多いドメインをリスト化したものが，「よく出現する大学・研究所等リスト」[18] である。刊行物や所蔵資料等を独自のデータベースで公開している大学がランクインしている点が特徴である。こうしたデータベースは，システムの構造やメタデータの形式，提供されているファイルの形式が様々であり，検索エンジンのクローラーが検索した場合には，検索結果が上位に表示されないことがあり，見つけづらくなっている可能性があると考えられる。

表 4　「その他」カテゴリのドメイン単位の出現回数
（ ）内は合計に占める割合である。

年度	2010	2011	2012
4 回以上	104 （14.94%）	117 （15.12%）	147 （17.52%）
3 回以下	592 （85.06%）	657 （84.88%）	692 （82.48%）
計	696	774	839

表 5　4 回以上出現するドメインのカテゴリ別種類数
（ ）内は合計に占める割合である。

年度	2010	2011	2012
大学・研究所等	61 （60.40%）	66 （56.90%）	71 （49.31%）
各種団体	23 （22.77%）	30 （25.86%）	38 （26.39%）
その他	16	18	33
不明	1	2	2
計	101	116	144

4.3　各カテゴリ間の関係

NII，NDL，JST は，雑誌論文や図書の書誌情報などを収集して提供するサイト（以下，「ポータル」という）をそれぞれ運営している。また，これらの機関は多くの無料のデジタル化資料を収集および作成しており，それぞれの運営するポータルに搭載して無料公開している。そのため，収録されている無料デジタル化資料は発見性も高いと考えられるが，調査の結果，ナビゲートログに数多く含まれているということが判明した。その背景として，各ポータルにどのような学術コンテンツが収録されているのかを十分に理解できていないことや，各ポータルは書誌情報を共有しているが，その共有の範囲を十分に把握できていないことが一因として考えられる。そのため，これらの機関の運営しているポータルについて，収録範囲の確認を行った。調査方法としては，ポータルで公開されている情報[19]～[24] を参照し，不足している情報について各機関のポータル担当者にメールで照会した。調査対象は，無料デジタル化資料の収録書誌情報の量が各機関において大きい CiNii Articles，NDL Search，J-GLOBAL，J-STAGE とした。なお，JST については 4.2（1）にあるように J-STAGE へのナビゲートが大半であったが，より多くの書誌情報や外部リンク情報を収録しているポータルとして J-GLOBAL についても調査を行った。

調査結果をまとめたものが**表 6** である。また，全体像について，**図 3** にて模式的に示した。なお，ここで示した 2 つの図表はあくまで主要な無料デジタル化資料収録データベースの状況についてのみまとめたもので，各ポータルは他にも多くの外部データベースと連携し，書誌情報やリンク情報を収集・提供している。2 つの図表から，CiNii Articles，NDL Search，J-GLOBAL，J-STAGE はそれぞれ他 3 つのポータルと部分的に連携しているが，特定のポータルから他の主要なポータルの収録資料を全て検索することはできないということがわかった。特に，J-GLOBAL からは，J-STAGE 収録コンテンツの一部分しか検索できないという点と，CiNii Articles からは，JAIRO 収録コンテンツのうち，博士論文や研究報告書など“非論文系”の学術コンテンツを検索できない[20] という点は，見落としやすい点と思われる。

なお，今回調査対象とした ILL ログについては，

表 6　各ポータルの連携状況のまとめ

対象データベース		3 つのポータルの各データベースへの連携状況		
名称	収録コンテンツの件数	CiNii Articles	J-GLOBAL	NDL Search
NII-ELS（学協会刊行物・研究紀要）	一部有料コンテンツ込みで，458 万件（2013.9 時点）	○全て収録【一部有料コンテンツ込みで，458 万件（2013.9 時点）/日次更新】	×	○基本的に全て。更新頻度未定のためタイムラグあり。【420 万件（2013.11 時点）/現在初期投入分のみ。2013 年度末更新予定】
JAIRO	167 万件（2013.12 時点）	△機関リポジトリ側の CiNii 連携申請＋論文の NII-Type による対象制限。【95 万件（2013.9 時点）/週次更新】	×	○全て収録【週次更新】
J-STAGE	一部有料コンテンツ込みで 258 万件（2013.12 時点）	△コンテンツ提供学協会の CiNii 連携許諾（オプトアウト）。【一部有料コンテンツ込みで 70 万件（2013.9 時点）/年数回更新】	△ JDreamIII 採択記事となっている論文のみ。【一部有料コンテンツ込みで 35 万件（2013.12 時点）】	△コンテンツ提供学協会の NDL Search 連携許諾（オプトアウト）。【15 万件（2013.11 時点）/現在初期投入分のみ，2013 年度末より定期更新予定】
国立国会図書デジタルコレクション（電子書籍・電子雑誌）	18 万件（2013.12 時点）	×	（NDL Search へ検索をかけるアイコンあり）	○全て収録【9 万件（2013.11 時点）/週次更新】
国立国会図書デジタルコレクション（所蔵資料）	Web 無料公開は 47 万件（2013.11 時点）	×	（NDL Search へ検索をかけるアイコンあり）	○全て収録【228 万件のうち Web 無料公開 47 万件（2013.12 時点）/週次更新】

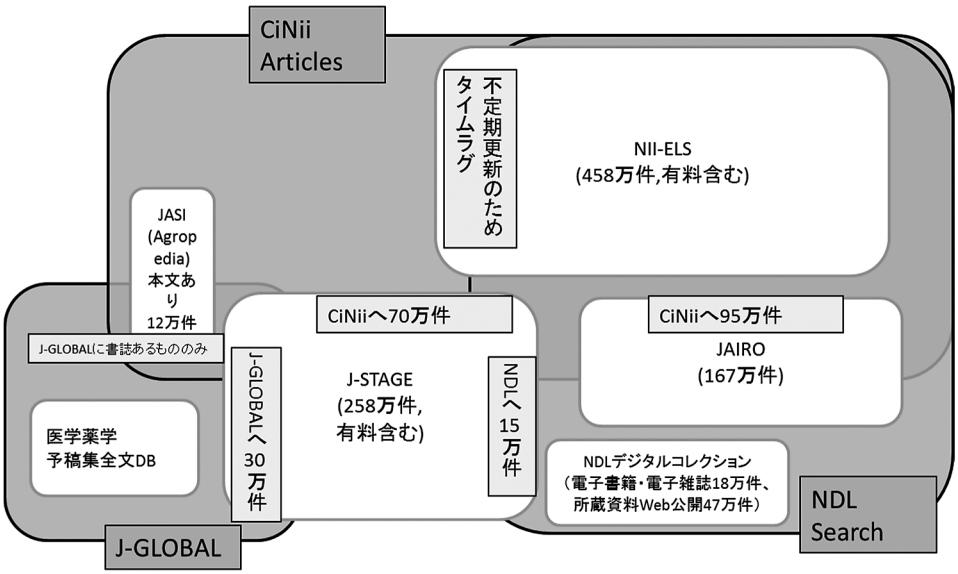

図 3　NII，NDL，JST の主な国内無料デジタル化資料の収録模式図

ナビゲートされている URL は必ずしも全文が公開されている URL であるとは限らず，外部リンクを持つ書誌データ等の URL やポータル等がナビゲートされている場合がある。また，リポジトリについては大半が JAIRO に書誌が収録されている。大学等のサイトで公開されているデジタル化資料についても JAIRO や CiNii Articles で公開されているものがある。つまり，どのサービスの URL へのナビゲートとカウントされるかはあくまでもコメント欄に入力されていたとおりである。また，ナビゲートログの約 60% 程度を占める NII，NDL，JST，リポジトリ，大学等のサイトで公開されている資料は複

数のサイトで公開されているもの，複数のポータルで検索が可能なものが含まれている。以上の点により各カテゴリの割合はかなり振れ幅があると推定され，各カテゴリが相互に重なり合っていると考えられる。さらに，これら 60％の資料群は，ILL の依頼前に CiNii Articles，NDL Search，J-GLOBAL，J-STAGE に加えて検索エンジンを利用すれば，発見可能な資料群であるという結果になった。

5. 考察

2012 年度のナビゲートログのドメインの種類数を集計した結果，3 回以下しか出現しないドメインが全体の約 8 割を占めていた。ドメインは組織ないしはサービスレベルで使い分けられることが一般的であり，様々な組織・サービスにおいてデジタル化資料が公開されていると同時に，ILL 参加館である大学等でニーズのある資料が広くウェブ上に存在していることがわかる結果となった。一方で，ナビゲートされているドメインの約 6 割は，NII，NDL，JST，リポジトリ，大学等のサイトで公開されているデジタル化資料であった。ただし，これらの資料は，一か所のみではなく，複数のサイトで公開されているものや複数のポータルで検索が可能なものが含まれており，複雑に関連していた。さらに，NII，NDL，JST の 3 機関の各種ポータルの相関関係を調査した結果から，どのポータルもワンストップではなく，3 機関で提供しているデジタル化資料を漏れなく検索するためには，各機関のポータルをそれぞれ検索する必要があるということがわかった。

しかしながら，言い換えれば，今回の調査でナビゲートされていた無料デジタル化資料の 6 割は，CiNii Articles，NDL Search，J-GLOBAL，J-STAGE，検索エンジンを全て検索すれば，いずれかには収録対象となっているという点で発見可能な資料であった。このように，国内の主要なポータル数か所と検索エンジンを利用すれば見つけられる資料が ILL で依頼されてしまった原因を考えることで，以下のような課題とその解決策が見えてきた。

一点目が，図書館員を含めた利用者が，無料デジタル化資料を十分に検索していないという点である。4.2 で言及したように，複数のサイトに書誌情報があるがフルテキストへのリンクは全てのサイトにない場合があり，書誌のフルテキストへのつながりが不十分なケースがある。また，雑誌資料は同一タイトルでも巻号によって別のサイトで公開されていたり，同じ資料が複数のサイトで公開されたりするケースもある。このような複雑な状況が原因の一つとして考えられる。利用者としては，主要なポータルと検索エンジンを併用するなど，検索スキルの向上が課題になるといえる。

二点目が，国内のポータル間の関係を理解するに当たって利用者にとって十分な情報が提供されていないという状況である。NII，NDL，JST の 3 機関の提供するポータルサイトについて，無料デジタル化資料の収録範囲を調査したが，十分な情報が提供されていない部分があり，別途メールにて担当者に問い合わせを行った。情報検索時に，そのポータルにおけるデジタル化資料の収録対象を知ることができなければ適切に探すことは難しい。この課題については，各種ポータルの提供者側で情報を公開することが重要であり，適切な情報開示を求めたい。なお，書誌の相互連携の場合には，機械的な連携に限界があり，完全な網羅が難しい場合があるため，利用者の側もその点については念頭に置いておく必要がある。

三点目として，学内刊行物の公開方法が機関によって様々であったという点である。3.2 で報告した「よく出現する BIBID リスト」[12] によると，ILL で依頼された無料デジタル化資料の中に，大学等で発行されている学内刊行物が数多く含まれていた。学内刊行物の公開方法については，国内で統一された基準があるわけではなく，書誌データも自館のサイトで様々な形式で作成している場合がある。近年では，リポジトリの構築が進んでいるが，今回の調査で，リポジトリへのナビゲートも比較的大きな割合を占めているという結果も明らかとなった。学内刊行物を公開する際には，リポジトリへのコンテンツ登録の重要性とともに各種ポータルでの書誌作成や連携を考慮したうえで，リポジトリのコンテンツの発見性についても考えていく必要があるのではないかと思われる[25]。

6. 今後に向けて

今後，ウェブ上のデジタル化資料がますます増加し，それを追いかけるように各種ポータルや検索エンジンはその勢力図を塗り替え新旧交代をしていく中で，効率的にデジタル化資料を探せるポータルや検索エンジンは時代とともに変わっていくと推測される。しかし，ウェブ上のデジタル化資料の全てのインデックスを持っていてそれら全てを検索可能なツールがない以上は，複数の検索ツールを併用することがこれからも必要になると考えられる。

そこで，無料デジタル化資料の発見性向上のために，まずポータルの提供者に，ポータルの収録範囲の明示を求めたい。このことによって，図書館員を

含め，検索者が思い込みを持ったまま検索することを防ぐことができる。今日では多くのポータルがWeb APIを公開するなどしていて，利用の自由度が増して来ている。こういった技術をうまく用いることで，検索スキルや検索の対象範囲を意識せずともデジタル化資料の発見性を高めるようなサービスやツールを作成できる可能性も広がっている。このような既存のポータルを動的に組み合わせてサービスやツールを作成する場合にも，そのポータルから何を検索できるのかという情報が改めて重要になってくるだろう。

また，本調査結果から，上記のようなポータルや検索ツールそのものの改善に加えて，無料デジタル化資料の発見性向上のためには，図書館員が各種ポータルの検索対象範囲を認識し，そしてその知識を更新していくことが必要なこともわかった。そして，そういったポータルの収録範囲と利用方法を利用者にも認識してもらえるよう利用者教育を行うことも必要であるといえる。

一方で，無料デジタル化資料のうち，ポータルやリポジトリで提供されていない資料や検索できない資料について考えた場合，デジタル化資料を提供する組織あるいは個人が多岐にわたっているため，別途アプローチを考えていく必要がある。

最後に，本稿そのものがデジタル化資料をとりまく課題共有の参考となれば幸いである。

謝辞

本稿執筆にあたり，順位相関係数の算出には，佐藤義則教授（東北学院大学）のご助力をいただきました。また，佐藤教授，大向一輝准教授（国立情報学研究所），国立情報学研究所学術基盤推進部学術コンテンツ課の高橋菜奈子副課長をはじめとするスタッフの皆様，ともに受講したグループ2の皆様には，平成25年度学術情報システム総合ワークショップ（試行）の研修開始から本稿作成まで長きにわたり，多くの知識や示唆を与えていただきました。ここに感謝の意を表します。また，メールでの聞き取り調査に丁寧に対応してくださった塩崎亮氏（国立国会図書館），坂内悟氏（当時，科学技術振興機構）に御礼申し上げます。

注記・参照文献

1）リンクリゾルバやウェブスケールディスカバリーサービス等のナレッジベースには，一部の無料公開デジタル化資料の情報が搭載されている。

2）ILLの謝絶状況及び謝絶理由を分析した先行研究としては以下のものがある。石山夕記．“NACSIS-ILLの理念と実態：参加館における「謝絶」の状況とその理由の解明に向けて”．Library and information science. 2010, no.64, p.81-107.

3）NACSIS-ILLの具体的な状態遷移については，以下のNIIのサイトにて公開されている。国立情報学研究所．“状態遷移図2010年12月版”．国立情報学研究所目録所在情報サービス．（オンライン），http: //www. nii. ac. jp/CAT-ILL/archive/pdf/Diagram.pdf,（参照2014-03-28）.

4）国立情報学研究所．“NACSIS-ILL統計情報”．国立情報学研究所目録所在情報サービス．（オンライン），http: //www. nii. ac. jp/CAT-ILL/archive/stats/ill/endrecord.html,（参照2014-03-28）.

5）文字列の抽出には正規表現を用いた。

6）1つのコメントの中に2つもしくはそれ以上のURLが記載されている場合，複数のURLを抽出し，集計した。

7）1つのログの中に，ナビゲートしていると思われるコメントが複数あった場合には，時系列上最初に出現するコメントを採用した。

8）MeCab: Yet Another Part-of-Speech and Morphological Analyzer.（オンライン），http://mecab.googlecode. com/svn/trunk/mecab/doc/index. html,（参照2014-03-28）.

9）taku-ku.“mecab-ipadic”. Sourceforge.（オンライン），http://sourceforge.net/projects/mecab/files/mecab-ipadic/,（参照2014-03-28）.

10）$p < 0.001$ の“p”は有意確率（probability）である。この場合では，帰無仮説「2変数間は無相関である」を，99.9%の確率で棄却できるということを意味する。

11）受付館としてカウントしたのは，最初に無料のデジタル化資料にナビゲートするコメントを送った館である。謝絶（PARDON）コマンドで次の候補館に転送されるケースがあるので，最終的にILLログに記録されている受付館とは異なる場合がある。

12）小島由香，長屋俊，藤江雄太郎．“よく出現するBIBIDリスト”．学術情報システム総合ワークショップ平成25年度成果物．（オンライン），http://www. nii. ac. jp/hrd/ja/ciws/report/h25/nii1_ref5.pdf,（参照2014-03-28）.

13）佐藤義則．近年のNACSIS-ILLにおける看護文献の需要と供給：ログ分析の結果から．看護と情報：看護図書館協議会会誌. 2007, no.59, p.1-15.

14）佐藤義則．“これからのNACSIS-CAT/ILLの運用体制について”．NACSIS-CAT/ILLワークショップ平成23年度講義資料．（オンライン），http://www. nii. ac. jp/hrd/ja/ciws/h23/txt3. pdf,（参照2014-03-28）.

15）今回の調査では2012年度では12.88%が貸借であった。

16）国立情報学研究所．“機関リポジトリ一覧”．学術機関リポジトリ構築連携支援事業．（オンライン），http://www.nii.ac.jp/irp/list/，（参照 2014-03-28）．
17）小島由香，長屋俊，藤江雄太郎．“よく出現するURL リスト”．学術情報システム総合ワークショップ平成 25 年度成果物．（オンライン），http://www. nii. ac. jp/hrd/ja/ciws/report/h25/nii1_ref4.pdf，（参照 2014-03-28）．
18）小島由香，長屋俊，藤江雄太郎．“よく出現する大学・研究所等リスト”．学術情報システム総合ワークショップ平成 25 年度成果物．（オンライン），http://www. nii. ac. jp/hrd/ja/ciws/report/h25/nii1_ref6.pdf，（参照 2014-03-28）．
19）国立情報学研究所学術基盤推進部学術コンテンツ課．“NII コンテンツサービスアラカルト―CiNii”．学術情報システム総合ワークショップ平成 25 年度講義資料．（オンライン），http://www. nii. ac. jp/hrd/ja/ciws/h25/2-05.pdf，（参照 2014-03-28）．
20）国立情報学研究所．“システム情報”．学術機関リポジトリ構築連携支援事業．（オンライン），http://www. nii. ac. jp/irp/archive/system/irdb_harvest.html，（参照 2013-12-10）．
21）国立国会図書館．“検索対象データベース一覧”．国立国会図書館サーチ．（オンライン），http://iss.ndl.go.jp/information/target/，（参照 2013-12-10）．
22）国立国会図書館．“資料デジタル化について”．（オンライン），http://www.ndl.go.jp/jp/aboutus/digitization.html，（参照 2013-12-10）．
23）科学技術振興機構．“J-GLOBAL について”．J-GLOBAL．（オンライン），https://jglobal.jst.go.jp/footer.php?page = aboutus，（参照 2013-12-10）．
24）科学技術振興機構．“既存リンクサービスのご利用にあたって”．J-STAGE．（オンライン），https://www.jstage.jst.go.jp/pub/html/AY04S250_ja.html，（参照 2013-12-10）．
25）Digital Repository Federation．“IRcuresILL”．（オンライン），http://drf.lib.hokudai.ac.jp/drf/index.php?IRcuresILL，（参照 2014-03-28）．

＜2014.4.30　受理　ふじえ　ゆうたろう　大阪大学附属図書館吹田地区図書館サービス課理工学図書館班，こじま　ゆか　名古屋大学理学部・理学研究科・多元数理科学研究科図書掛，ながや　しゅん　日本原子力研究開発機構研究技術情報部情報メディア管理課収集調整係＞

Yutaro FUJIE, Yuka KOJIMA, Shun NAGAYA

Some considerations about free online academic resources based on the analysis of NACSIS-ILL logs

Abstract：There is now a multitude of digitized materials that have been made available for free around the world. In order to consider measures to improve the discoverability of these free resources, the authors have analyzed the transaction logs from NACSIS-ILL logs. The authors surveyed the interlibrary loan transactions where the request was cancelled because an item was freely available to see what types of materials were requested. As a result, the authors found that the digitized materials are widely-scattered on the web and of particular significance approximately 60% of the domestic digitized materials are discoverable using the search engines and following services: NDL Search, CiNii Articles, J-GLOBAL, and J-STAGE. While there is a partial linkage between the metadata existing in each of these sites, it is not possible to see detailed information about the contents and that fact contributes to the decreased discoverability.

Keywords：free digital resources / NACSIS-ILL / National Institute for Informatics / Workshop on Scientific Information Systems / institutional repositories / interlibrary loan / data log analysis / open access / digitization