

電子化された日本語研究論文の流通実態と「つなぐ」文献目録によるアクセス支援

茂木俊伸*

本稿では、日本語研究（日本語学）分野における学術情報の流通の実態を明らかにするために、特定の研究トピック（「とりたて」）を扱った論文の電子化状況について、本文情報がどの程度電子化されているか、それらがどこで公開されているのかの調査を行った。また、前稿で提唱した「つなぐ」文献目録モデルの事例として作成した「「とりたて」関連研究文献目録」をデータとして用いることで、この目録が本文情報へのアクセス性の向上にどの程度貢献するものであるのかの評価も行った。

[キーワード：論文の電子化, CiNii Articles, 機関リポジトリ, 「つなぐ」文献目録]

1. はじめに

近年、人文科学系の研究分野でもオープンアクセスの動きが広がり、電子化された論文情報によって研究活動を効率的に進めることのできる環境が整いつつある。茂木（2010）では、日本語研究分野の論文の具体的な流通実態として、二次資料の年鑑に掲載されている2007年発行の大学紀要類の調査を行い、約37%の紀要の本文がウェブ上で入手可能であることを明らかにした。

このような学術情報の流通環境に対する日本語研究者の関与は、現状では、主としてその過程の始点（論文の作成と公開）と終点（論文情報の探索と入手）への部分的な関わりにとどまっている。しかし、情報発信の活性化や情報流通の円滑化といった環境整備にも、研究者自身がより主体的に関わるべきであると考えられる。

前稿（茂木2012）では、日本語研究におけるCiNii Articles（国立情報学研究所）（以下、本文中で言及する場合は「CiNii」と呼ぶ）を中心とした論文情報流通の現状の確認と、それを補完するために日本語研究者ができる貢献として「「つなぐ」文献目録」の提案を行った。

本稿は、実際に作成した文献目録のデータをもとに、特定のトピックに関する論文情報の流通実態の調査と、その結果に基づいた文献目録の評価を行うものである。

以下ではまず、前稿で示した文献目録モデルの概要（第2節）と、実際に作成した文献目録の概要（第3節）について述べる。次に、この文献目録のデータに基づき、ウェブ上で入手可能な本文情報がどれくらい、どこに存在するのか、調査を行う。さらに、CiNiiを利用したこれらの論文へのアクセス可能性の調査と、文献目録によるアクセス性の向上に関する分析を行う（第4節）。

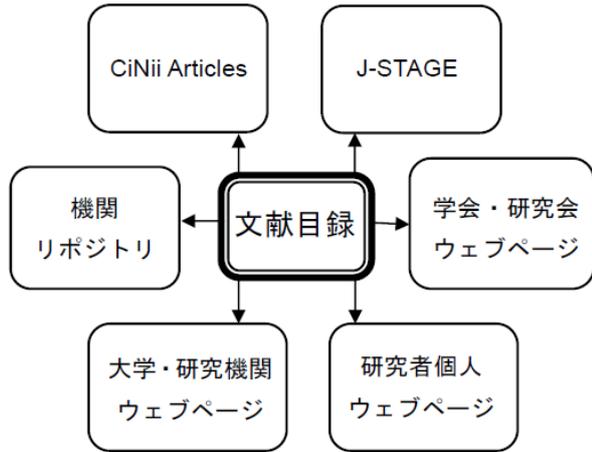
2. 「つなぐ」文献目録モデル

茂木（2012）では、人文科学系の研究者の情報入手行動に関する先行研究で明らかにされている「論文データベースを利用して論文情報を検索する場合、CiNiiの利用度が高い」という点をふまえたうえで、日本語研究者がCiNiiを利用して日本語研究論文を入手しようとする際、次のような要因によって本文へのアクセス性が低下する、もしくは確保されない可能性があることを指摘した。

- (1) a. CiNii未収録の紀要がある。
- b. CiNiiを経由せず、機関リポジトリや大学ウェブページに直接アクセスしなければ本文を入手できない紀要がある。
- c. CiNii経由では学会誌の本文が入手しにくい。
- d. 研究機関の論文データベースには本文が収録されておらず、CiNiiとの連携も取れていない。
- e. 機関リポジトリには「深層ウェブ」問題が見られる場合がある^{注1)}。

そのうえで、日本語研究者がその専門性を活かして論文情報の流通の円滑化や効率化に貢献する一つの方法として、本文情報へのリンクを伴った文献目録を作成する（あるいは既存の文献目録に本文情報へのリンクを加える）ことを挙げた。日本語研究論文の本文情報は、現在、ウェブ上に存在してはいるものの、さまざまなデータベースやウェブページに分散した状態になっており、このような「つなぐ」文献目録によって、ともすれば膨大な情報に埋もれてしまう論文情報にピンポイントでアクセスできるようにすることができる。次の〔図1〕は、その概念図である^{注2)}。

4. 調査と分析



〔図1〕「つなぐ」文献目録モデル

以下では、事例研究として、このモデルに基づいて作成した文献目録をデータとして調査を行い、日本語研究の特定のトピックにおいて、論文本文の電子化はどの程度進んでおり、どこで入手可能であるのか、また、上のモデルのような文献目録がどの程度のアクセス性の向上に貢献するのか、といった点について分析を行う。

3. データの概要

本稿における調査では、筆者が2011年から作成し、ウェブ上に公開している「とりたて」関連研究文献目録（以下、「とりたて」文献目録とする）をデータとして使用した^{注3)}。

これは、現代日本語の研究において主要なトピックの一つとなっている「とりたて」（いわゆる「副助詞」「係助詞」に関わる問題）の研究を支援するために、文献の書誌情報と本文へのリンクを掲載しているものである。

論文情報の採録範囲は、現代日本語（共通語）に関する国内文献を中心にしており、少数の書評や目録を含む。また、古典日本語や方言に関する研究、他言語の対応する語に関する研究のほか、海外文献も一部含まれている。研究分野としては、日本語学・国語学、言語学の論文が中心となっているが、日本語教育、言語習得、自然言語処理等の関連領域の論文も一部含まれている。（詳細については、上記目録を参照。）

以下で報告する調査では、2012年12月末時点の「とりたて」文献目録を使用した。データ量は、2003～2012年の10年間に発行された文献、計500件である。

この500件のうち、国内の大学や機関、学会、出版社等によって発行されている文献は473件（94.6%）であり、この他に海外文献が27件（5.4%）あった。以下の分析では、国内分のみを分析対象としていく。

調査では、まず、473編の国内文献について、掲載されている媒体ごとに、書籍、雑誌（商業誌、学会誌、紀要）、報告書（科研費、その他）に分類した。これが次の〔表1〕である。

このうち「書籍」は、出版社により発行されているもので、単著、共著（論文集など）の両方を含む。単著の場合は全体で1編として、共著の場合は掲載されている関連論文の数をカウントした。単著には、雑誌論文を再録したのものも含まれる。

「商業誌」は、出版社から発行されている月刊・季刊等の雑誌である。日本語研究分野の学会誌や研究会誌には出版社から発行されているものがあるが、これは次の「学会誌」とした。

「学会誌」は、学会や研究会等の機関誌として発行されている雑誌である。ただし、学内学会によって発行されているものは、次の「紀要」に含めた。

「紀要」は、大学・機関や、その内部の組織（学部、研究室、学内学会等）が発行元の雑誌である。また、「科研報告書」は、科学研究費研究（科研費）の研究成果報告書として発行された冊子体の報告書である。最後の「その他」に含まれるのは、学内プロジェクト報告書、COEプログラム報告書、学内組織による論文集である。

〔表1〕「とりたて」研究論文の掲載媒体

掲載媒体	論文数	
書籍	89	18.8%
商業誌	10	2.1%
学会誌	66	14.0%
紀要	274	57.9%
科研報告書	25	5.3%
その他	9	1.9%
計:	473	100%

〔表1〕から分かるように、「とりたて」研究論文の掲載媒体としては「紀要」が最も多く、約6割を占める。これは、もともと紀要が日本語研究論文の主要な発表媒体として位置付けられていることと一致する（cf. 日詰・逸村2010）。

続いて「書籍」が多いが、これは「とりたて」をテーマとした論文集（1冊）の収録論文を個別にカウントしていることがある程度影響していると思われる。また、「科研報告書」の掲載論文が比較的多いのも、「とりたて」をテーマとする研究課題（3件）の研究成果報告書が含まれていることによる。

4.1 「とりたて」研究論文の電子化の状況

以下では、[表1]の媒体のうち、電子化されて公開されることは少ないと思われる「書籍」と「商業誌」を除き、これ以外の媒体に掲載されている「とりたて」研究論文374編を対象として、これらの電子化された本文がどの程度ウェブ上で公開されている（入手可能である）かについて見ていく。

この調査は2013年1月上旬に行ったもので、論文ごとに先の[図1]に示したデータベース類やウェブページを確認し、本文ファイルが公開されているかを調べた。

この結果を示したものが、次の[表2]である。なお、電子化された論文にはテキスト形式で公開されているものも3編あったが、それらも含めてすべてPDF形式の本文データが入手可能であった。

〔表2〕「とりたて」研究論文の電子化

掲載媒体	論文数	本文あり	電子化率
学会誌	66	19	28.8%
紀要	274	141	51.5%
科研報告書	25	0	0.0%
その他	9	8	88.9%
計:	374	168	44.9%

[表2]のように媒体別に見ると、本文が電子化されている論文の割合は「その他」で最も高く、続いて紀要論文が約半数、学会誌論文が3割弱となっている。科研報告書の掲載論文に関しては、ウェブ上で本文が得られなかった。

以下では媒体別に電子化率を見ていくが、議論の便宜上、紀要以外の媒体の掲載論文について先に見ていく。

4.1.1 学会誌論文の電子化の状況

[表2]で「学会誌」論文に分類した66編のうち、19編がウェブ上で本文公開されている。これらがどのデータベースやウェブページで得られるのか、その公開場所を示したものが次の[表3]である。

〔表3〕学会誌論文の本文の公開場所

公開場所	論文数	電子化率
CiNii Articles	10	52.6%
学会・研究会	9	47.4%
J-STAGE	3	15.8%
出版社	1	5.3%
著者	1	5.3%

なお、この19編のうち、5編は複数の場所(例えば、CiNiiと学会ウェブページの2か所)で全文が公開されていた。

このようなケースは、[表3]ではそれぞれの場所に独立してカウントしている(このため、論文数の合計は24になる)。パーセンテージについては、19編のうちどれだけの論文の本文がそこで入手可能であるかを示している。

学会誌論文の公開場所として最も多いのはCiNiiであるが、うち4編は有料コンテンツであった。これに学会・研究会ウェブページが続くが、無料アクセスが可能な本文という点では、CiNiiより多い。ただし、この9編の発行元となっている学会・研究会は4団体であり、日本語研究分野のすべての学会や研究会が独自に本文を公開するという形態をとっているわけではない。

学会誌論文のみに見られる本文の公開場所としては、J-STAGE(科学技術振興機構)が挙げられる。J-STAGEで公開されている3論文の発行元はそれぞれ異なり、学会ウェブページとJ-STAGEの2か所で公開されている論文も1編あった。

一方、著者本人のウェブページによる学会誌論文の公開は少ない。著者による公開が行われているケースでは、学会誌論文以外の場合も含めて、Read&Researchmap(科学技術振興機構)のようなプラットフォームではなく、すべて個人作成のウェブページで本文が公開されているのが特徴的である。

また、学会誌論文を個人単位で機関リポジトリを通して公開している例は見られなかった。この要因としては、日本語研究分野の多くの学会誌が出版社から発行・市販されているため、権利関係の処理がハードルとなっていることが考えられる^{注4)}。

4.1.2 報告書類の電子化の状況

先にも述べたように、科研費の研究成果報告書に掲載されている「とりたて」研究論文は、今回の調査の範囲では全文公開されていなかった。

これらの論文の約半数は、書籍や学会誌、紀要等での既発表論文が再録されたものであるか、あるいは報告書刊行後に別の媒体で発表されているものであった。このような再録のあり方は著作権処理に影響するため、積極的な電子化の対象になっていないとも考えられる。

一方、[表2]で見たように、「その他」に分類した学内プロジェクトの報告書やCOEプログラム報告書、学内組織による論文集の掲載論文は、かなり高い割合でウェブ上で本文が入手できた。これらの公開場所を示したのが、次のページの[表4]である。

なお、ここで「大学・機関」に分類したのは、機関リポジトリ以外の大学・機関のウェブページ(例えば、図書館や附属施設の紀要目次のページ、研究室のページ等)で公開されているケースである。

〔表4〕 「その他」論文の本文の公開場所

公開場所	論文数	
機関リポジトリ	5	62.5%
大学・機関	2	25.0%
著者	1	12.5%

このタイプの報告書論文の公開場所は、大学・機関内が大部分を占めている。今回の調査の範囲では、学内プロジェクト等の報告書は、科研報告書と比べ、学内成果物として機関リポジトリの収録対象になりやすいという傾向があるようである。

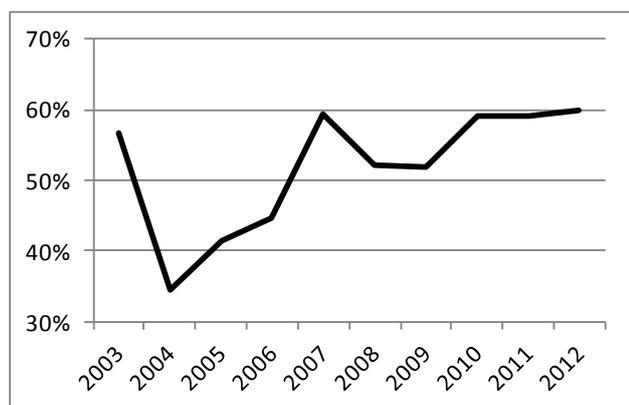
4.1.3 紀要論文の電子化の状況

紀要に掲載されている「とりたて」研究論文は、〔表2〕で見たように、274編のうち141編（51.5%）が本文公開されている。まず、年別に見た紀要論文の数と、その電子化率を示したものが、次の〔表5〕である。

〔表5〕 紀要論文の電子化率

年	2003	2004	2005	2006	2007
論文数	30	26	29	38	32
本文あり	17	9	12	17	19
電子化率	56.7%	34.6%	41.4%	44.7%	59.4%
年	2008	2009	2010	2011	2012
文献数	23	27	22	22	25
本文あり	12	14	13	13	15
電子化率	52.2%	51.9%	59.1%	59.1%	60.0%

データが少ないため確実なことは言えないが、グラフ化した次の〔図2〕からは、紀要論文の電子化率は出版年が後のものほど安定しており、全体的にはゆるやかな上昇傾向にあるように見える。新しい紀要論文ほど、ウェブ上で本文を入手しやすい状況にあると言える。



〔図2〕 紀要論文の電子化率（出版年別）

次に、141編の紀要論文の本文の公開場所を見ると、次の〔表6〕のようになる。

なお、141編のうち32編は、複数の場所で本文入手が可能であり、最も多い組み合わせは「CiNii+機関リポジトリ」で20編である。続いて「CiNii+大学・機関」が5編、同一大学・機関内の複数の場所で公開されている「機関リポジトリ+大学・機関」が同じく5編、「CiNii+著者」、「機関リポジトリ+著者」が各1編であった。

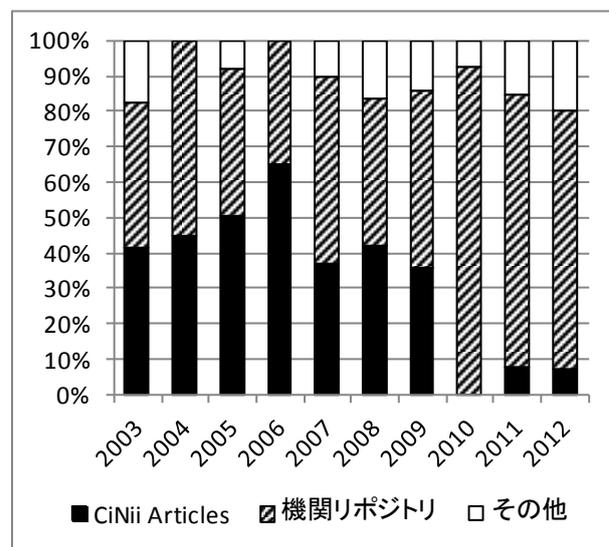
〔表6〕 紀要論文の本文の公開場所

公開場所	論文数	
機関リポジトリ	98	69.5%
CiNii Articles	47	33.3%
大学・機関	24	17.0%
著者	4	2.8%

〔表6〕のとおり、紀要論文の本文の公開場所として最も多いのは機関リポジトリで、全体の約7割に及ぶ。続いてCiNiiが全体の3分の1である。

機関リポジトリが運営されていない大学・機関が組織的に紀要の本文を公開しているケースは「大学・機関」に分類される。ただし、機関リポジトリを運営している大学でリポジトリ未収録の論文がこのタイプのページで公開されている例も見られた。

ここで、〔表5〕に示した各年の論文を、これらの公開場所別に分類したものが次の〔図3〕である。



〔図3〕 紀要論文の本文の公開場所（出版年別）

なお、公開元が複数の場合、CiNiiとの組み合わせはすべて「CiNii Articles」に、機関リポジトリと大学・機関、著者との組み合わせは「機関リポジトリ」にカウントしている。

〔図3〕からは、2009年までの紀要に掲載されている「とりたて」研究論文の本文は、CiNiiと機関リポジトリを併用する形で公開されているのに対して、2010年以降の論文では、機関リポジトリが公開の中心的役割を果たしていることが分かる。

機関リポジトリのコンテンツの中心が紀要であることはよく知られており（例えば、国立情報学研究所2011, 若松2011）、このデータからは、新規に刊行される紀要が、学内成果物として比較的早く機関リポジトリに登録される流れになっていることが推察される。

4.2 アクセス経路としての「とりたて」文献目録

前稿では、日本語研究論文の情報がCiNiiを中心として流通していることを見た。一方で今回の調査では、〔図3〕のように、2010年以降に発行された紀要論文の本文の多くが、CiNiiに直接収録されていないことが分かった。

したがって、これらの新しい紀要論文に関しては、CiNiiの書誌情報から機関リポジトリ等へのリンクによる誘導がなされていなければ、アクセス性が低くなると考えられる^{注5)}。そこで、次に、このようなCiNiiを経由した「とりたて」研究論文本文へのアクセスの状況について調査した。

まず、ウェブ上で本文が入手できる紀要論文141編のうち、CiNiiを利用してアクセス可能な本文の割合を出したものが、次の〔表7〕である。

〔表7〕 CiNiiを利用した本文入手

本文入手	形式	論文数	
可	本文掲載	47	33.3%
	リンク	72	51.1%
不可		22	15.6%
計:		141	100%

紀要論文のうち、CiNiiに本文が収録されているもの47編（33.3%）に加え、CiNiiの論文書誌情報からのリンクによって機関リポジトリ等から本文を入手できるものが72編（51.1%）あり、計119編（84.4%）がCiNiiの利用によって本文へのアクセスが可能であった。

この数字は、少なくとも今回の調査の範囲では、論文の本文入手のための行動として第一にCiNiiを利用するという方法が、実態としてもかなり有効であることを示している。

一方で、CiNiiを利用した経路では本文が入手できなかった論文が22編（15.6%）あった。このうち6編（全体の4.3%）は、そもそもCiNiiに書誌情報が登録されていないものであった。

なお、これらのうちの一部は、CiNiiと機関リポジトリとのデータ連携のタイムラグによって、調査時点では本文が入手できなかったと思われるものである（実際、本稿執筆時には新たに機関リポジトリへのリンクが設けられた論文がある）。

さて、前稿で提案した〔図1〕のような「つなぐ」文献目録の主な目的は、CiNii以外の場所に分散している論文の本文情報へのアクセス性を向上させることであった。

〔表7〕の結果から、紀要論文に関して言えば、CiNiiのみを使用して本文を探す場合と比べて、「とりたて」文献目録に設けた本文情報へのリンクは、15.6%の論文へのアクセス性を向上させる効果を持っていることになる。

さらに、〔表2〕の学会誌や報告書類まで含めた全体では、CiNiiの利用で入手できない本文情報を「とりたて」文献目録でカバーしている論文は39編（うち8編はCiNii未収録）となり、これらの論文全体の23.2%になる。

したがって、「つなぐ」文献目録の事例として作成し、今回の調査対象とした「とりたて」文献目録は、日本語研究論文の情報流通においてCiNiiが果たしている役割を十分に活用するとともに、ウェブ上に分散している論文情報に効率よくアクセスするための「つなぐ」機能を一定レベルで果たしていると言える。

5. まとめ

本稿では、電子化された日本語研究論文の流通の実態の一端として、「とりたて」研究論文の電子化の状況を見てきた。

ここから見てきたのは、論文の本文情報の流通という「経路」の問題だけでなく、日本語研究者がいかに学術情報を発信していくべきかという「始点」の問題であり、それらの情報をどのように工夫して入手するかという「終点」の問題である。最後に、調査のまとめとともに、このような課題についても述べていく。

「とりたて」研究論文の掲載媒体に関しては、紀要論文が最も多く、直近3年間に発行された紀要論文では約6割の本文が入手可能であること、またそのほとんどが機関リポジトリで公開されていることが分かった。

ただし、紀要論文でも、学内学会や研究室など比較的小規模な組織が発行元となっているものは、全般的に電子化が低調な傾向にある。このタイプの紀要の公開には著作権処理が問題となっている可能性があり、組織的な対応が望まれる。

一方、学会誌の「とりたて」研究論文に関しては、それほど電子化率が高くなかった。また、CiNiiやJ-STAGEといった大規模プラットフォームではなく、学会や研究会のウェブページ上で公開しているものが約半数あった。これらの独自の取り組みは評価すべきものであるが、論

文の流通経路の複雑化が危惧されることから、ユーザーへの配慮という点では、何らかの形でCiNiiとの連携を考える必要があると思われる。

科研費の研究成果報告書の収録論文は、今回の調査では全文公開されていなかった。科学研究費助成事業データベースKAKEN（国立情報学研究所）では、採択課題の概要や成果のリストは見られるものの、論文の全文公開は行われていない。また、研究実績報告書の「発表文献」欄の論文情報からCiNiiへのリンクが設けられている場合があるが、日本語学分野の課題を見るかぎりでは、それほど多くない。現状では、科研費の成果としての論文の入手コストは比較的高い状況にあると言える。

また、今回の調査範囲では、論文著者によるセルフアーカイブはほとんど行われていなかった。機関リポジトリを利用した学会誌論文の公開は行われておらず、それ以外の場所での公開も、5人の研究者による6編の論文（全文が公開されている「とりたて」研究論文全体の3.6%）にとどまっている。これらの論文の入手行動としては、1編を除いてGoogle Scholarの検索で情報が得られるため、CiNiiの利用に加え、Google Scholarの併用が有効であるように思われる。

以上のように、電子化された日本語研究論文の本文情報の公開は、現状では、CiNiiと機関リポジトリを中心として行われており、同時に、学会・研究会や著者が個別に独自の取り組みを行っているという状況にある。このような情報の分散傾向はいずれ解消されるのかもしれないが、過渡期の問題解決の一つの方法として、「つなぐ」文献目録は有効に働くのではないかとと思われる。

注

※本研究は、日本学術振興会科学研究費若手研究(B)「電子化された日本語研究論文情報の流通・活用に関する基礎的研究」(研究課題番号:23720230)の助成を受けたものである。

1) 機関リポジトリにおける「深層ウェブ」問題については、宮田ほか(2012)に分析が見られる。

2) 図中の「CiNii Articles」(旧CiNii)、「J-STAGE」(旧Journal@rchive)に関しては、サービスの名称に変更が見られたため、前稿の図から修正を行った。これらのサービスに関しては、それぞれ大向(2012)、佐藤・久保田ほか(2012)参照。

3) 「「とりたて」関連研究文献目録」は、以下で公開している。

http://www.naruto-u.ac.jp/facultystaff/tmogi/fp_biblio/

4) ただし、清水ほか(2012)によれば、機関リポジトリへの学会誌論文の登録が低調であるのは人文・社会系だけの傾向ではないとされる。なお、「とりたて」研究論文が掲載されている学会誌・研究会誌のうち14誌については、「学協会著作権ポリシーデータベース」(SCPJデータベース)において、著者が論文本文の公開を希望する際の条件が確認できる。これによれば、「出版社版」の公開を認めない学会誌が3誌あり、公開までの猶予期間は最大で4年であった。

5) 佐藤・大向ほか(2012)では、このようなアクセス経路の需要が人文・社会系の論文で高いことが示されている。

参考文献

- 大向一輝(2012)「電子書籍化する学術論文—CiNii Articlesの展開を中心に—」『情報処理』53(12), pp. 1282-1286, 情報処理学会。
- 倉田敬子(2010)「オープンアクセスとは何か」『情報の科学と技術』60(4), pp. 132-137, 情報科学技術協会。
- 国立情報学研究所(2011)『変容する学術情報流通、進展する機関リポジトリ—学術機関リポジトリ構築連携支援事業 第2期報告書—』, 国立情報学研究所。
<<http://www.nii.ac.jp/irp/archive/report/index.html>> [2013年1月18日最終確認]
- 佐藤 翔・大向一輝・関戸麻衣・逸村 裕(2012)「アクセスログに基づくCiNiiによる本文提供とその利用状況の分析」『2012年日本図書館情報学会春季研究集会発表要綱』, pp. 73-76, 日本図書館情報学会。
- 佐藤竜一・久保田壮一・青山幸太・土屋江里・宮川謹至(2012)「J-STAGE 新システムが加速する国内学術論文誌の電子化と流通」『情報管理』55(2), pp. 106-114, 科学技術振興機構。
- 清水真理・佐藤 翔・逸村 裕(2012)「日本の学協会誌掲載論文の機関リポジトリ収録状況」『情報知識学会誌』22(2), pp. 77-82, 情報知識学会。
- 日詰梨恵・逸村 裕(2010)「CiNii収録率から見たわが国の学術情報電子化の現状—人文学4領域を対象に—」『中部図書館情報学会誌』50, pp. 19-35, 中部図書館情報学会。
- 宮田洋輔・安形 輝・池内 淳・石田栄美・上田修一(2012)「深層ウェブの実態とその要因:機関リポジトリに登録された文献を用いた調査」『日本図書館情報学会誌』58(2), pp. 97-109, 日本図書館情報学会。
- 茂木俊伸(2010)「日本語研究論文情報の電子化の実態と論文探索スキル」『鳴門教育大学情報教育ジャーナル』7, pp. 9-14, 鳴門教育大学。

茂木俊伸 (2012) 「電子化された日本語研究論文の流通実態と問題点」『鳴門教育大学情報教育ジャーナル』9, pp. 23-29, 鳴門教育大学.

若松昭子 (2011) 「日本における大学の学術情報発信システムの発展と課題—紀要創刊から機関リポジトリ構築へ—」『聖学院大学論叢』24(1), pp. 161-170, 聖学院大学.

参照ページ

CiNii Articles

<http://ci.nii.ac.jp/ja>

Google Scholar

<http://scholar.google.co.jp/>

J-STAGE

<https://www.jstage.jst.go.jp/browse/-char/ja>

KAKEN (科学研究費助成事業データベース)

<http://kaken.nii.ac.jp/>

Read&Researchmap

<http://researchmap.jp/>

学協会著作権ポリシーデータベース (SCPJデータベース)

<http://scpj.tulips.tsukuba.ac.jp/>

[いずれも 2013 年 1 月 18 日最終確認]