失われていくインターネット上の参考文献 図書館情報学分野の雑誌論文に参照された インターネット文献の入手可能性の分析調査

Missing references cited from sources on the Internet

An analysis of availability of Internet references in the articles of library and information science journals

藤田 節子¹

FUJITA Setsuko¹

1 川村学園女子大学教育学部社会教育学科 (〒270-1138 千葉県我孫子市下ヶ戸1133)

Tel: 04-7183-7228 Fax: 04-7183-8143 E-mail: s.fujita@kgwu.ac.jp

1 Faculty of Education, Kawamura Gakuen Woman's University (1133 Sageto Abiko-shi, Chiba 270-1138)

原稿受理 (2010-10-09)

情報管理 53(9), 492-503, doi: 10.1241/johokanri.53.492 (<http://dx.doi.org/10.1241/johokanri.53.492>)

著者抄録

近年、学術論文の参考文献に、インターネット情報の利用が増加している。Web ページは更新や削除が頻繁に行われるため、インターネットからの参考文献は読者が後になって入手できるとは限らない。本調査は、2007 年と 2005 年に刊行された図書館情報学関係学術誌 4 誌に掲載された論文の参考文献のうち、インターネット情報を参照した 858 件の入手可能性を調べた。その結果、2007 年では参照されたインターネット文献のうち 28%、2005 年では 41%が、論文の著者が記述した URL のもとでは見つからなかった。見つからなかったインターネット文献を、そのサイト内で検索したり、検索エンジンを使うと、2007 年は 89%、2005 年では 82%が入手可能になった。参照されたインターネット文献の保存、サイト運営者の適切な管理運営、著者の正確な引用と、必要十分な書誌記述が入手可能性を高める。

キーワード

インターネット引用, Web参照, 引用文献, 参考文献, インターネット文献, リンク切れ, URL消失, Webアーカイブ, 機関リポジトリ, 書誌記述

1. はじめに

現在、学術雑誌の論文の注・引用文献リストや参考文献リストには、インターネット上の情報源から入手した情報や文献（以下ネット文献という）が数多くみられる。

引用・参考文献（以下参考文献という）は、根拠となった文献の出典を記述することで、被引用文献の著者への敬意を示すとともに、論文の信頼性や新規性の評価を得る役割を持っている。また、読者にとっても有用な文献案内となる。したがって、参考文献は、論文発表後も読者が入手できなければならない。

しかし、WebサイトやWebページは、頻繁に更新や移動、削除が行われ、同じURLに同じ情報内容が存在し続けるという保証はない。学術論文の著者は、インターネット上の情報を参照する場合は、論文発行後も入手可能な、信頼性のある情報源を慎重に選択しているが、果たして後年どのくらい入手可能であろうか。筆者自身も、ネット文献を参照する機会が多いが、その後、記述したURLから入手可能かどうかを検証することはあまりない。

参照されたネット文献の入手可能性の研究は、1990年代後半から、対象とする分野や期間、方法などはさまざまであるが、いろいろな調査が行われている¹⁾。2000年以降で、本調査と同様の図書館情報学分野の学術誌を対象とした、参照されたネット文献の調査としてはCasserlyとBirdの調査^{2),3)}と、GohとNgの調査⁴⁾がある。Casserlyらは、2002年に、1999-2000年に刊行された図書館情報学雑誌の論文で参照されたネット文献500件を調査し、さらに2005年から2006年にかけて、同じネット文献を追跡調査している。Gohらは、1997年から2003年までの情報科学雑誌3誌から、2,516件のURLにアクセスし、URLの存続半減期の調査をした。

筆者も、2000年に、1998-1999年に刊行された図書館情報学分野の学術雑誌8誌に参照された電子文献333件の調査をした（以下1999年調査という⁵⁾。当時は、電子文献の正確な書誌記述がまだ普及しておらず、著者が参照した文献かどうかの判断がつかなかったため、必要な書誌要素が記述された雑誌論文と会議資料22件の入手可能性についてのみ調べた。その結果、18件は記述されたURLに文献が存在した。

それから10年が経過して、学協会の投稿規程にもネット文献の書誌記述例が掲載され、出典が正確に記載されるようになった。

そこで、本稿では、2007年と2005年発行の図書館情報学分野の学術雑誌に収録された論文を対象として、参照されたネット文献のURLにアクセスし、その入手可能性を調査し、その結果や調査プロセスの中

で浮かび上がってきた問題点を考察する。

2. 調査方法

2.1 調査対象

調査対象誌は、わが国の図書館情報学分野の学術雑誌4誌『情報管理』『情報の科学と技術』『日本図書館情報学会誌』『現代の図書館』の2007年1月から6月と2005年1月から6月に刊行された巻号とした。これらは、1999年調査をふまえて、年間の論文数が比較的多く、調査年月日から遡って約3年・5年前に刊行された、図書館情報学分野の論文誌を選択した。

調査対象とした論文は、原著論文、特集論文、事例報告、解説などで、講座、コラム、オピニオン、会議報告などは除外している。

これらの論文の文末の、注、引用文献、参考文献リストに掲載された参考文献のうち、URLが記述された文献を調査対象とした。「参照URL」というような表示の下に、URLだけを羅列しているものは省いた。あくまでも本文との関連をつけている、参照されたネット文献を用いた。

2.2 入手可能性の判定

2010年3月16日～4月15日の1か月間に、参照されたネット文献に記述されているURLを入力し、その文献を入手できるかどうか調べた。自動的に新しいURLに移動した場合でも、そのURLで入手できたとした。

表示されたWebページと参照文献に記述されたタイトルや内容が一致した場合に、入手できたと判断した。参照文献リストに、著者名やタイトルが記述されていない場合でも、本文の文章内容とWebページの内容が合っていれば、入手できたと判断した。

論文の著者が参照した年月日以降に、Webページが更新されていても、タイトルや内容が同じならば同じ文献と判断した。多くのWebページの更新日時や更新箇所が明確でなく、参照年月日の記述がない場合もあるので、本調査では、参照年月日と更新日

時の検証は行わなかった。

記述されたURLで入手できなかった場合は、通常読者が行うであろう方法として、トップページにアクセスし、メニューやディレクトリをたどったり、参照文献リストあるいは本文に記述されたタイトルや著者名などの書誌要素を用いて、サイト内検索機能で検索した。

サイト内検索機能がなかったり、サイト内検索でも見つからない場合、あるいはサイトのトップページが見つからない場合などは、Googleで、タイトルや著者名、本文中の記述をキーワードに検索した。Googleは世界で最も利用されている検索エンジンなので、URLの消失に関する先行研究のほとんどは、Googleを使用している。そのため、調査結果を比較したり、今後追跡調査をする場合にも同じ検索エンジンがよいと考えた。なお、サイト内検索にGoogleを使用しているサイトもあったが、Googleであつてもサイト内にある場合はサイト内検索とした。

URLにアクセスした際に、国立国会図書館などのWebアーカイブに保存されているという指示がある場合は、それを検索した。

また、検索した結果、引用の仕方や書誌事項が不十分であったり、英語以外の外国語サイトで書誌事項が日本語に翻訳されているなどの理由で、確かに著者が参照した内容かどうか確信を持って判断ができない場合は、判断不能とした。

実際には、時間をかけてもっと丁寧な探索をすれば、入手できる可能性はあるが、今回の調査では通常の読者が初歩的に行うと思われる検索にとどめた。

3. 調査内容

調査内容は、以下のとおりである。

(1) 参照されたネット文献の概要

参照されたネット文献数と特徴を把握し、その経年変化を1999年調査と比べる。

(2) URLの持続性

記述されたURLにアクセスし、参照文献を入手できるかを調べる。

(3) Web上での入手可能性

著者が記述したURLで入手できなかった参照文献が、サイト内検索やGoogleを利用してWeb上で入手可能となるかを調べる。

(4) 文献の種類との関連性

URLの持続性およびWeb上での入手可能性と文献の種類（論文・Webページなど）に関連性があるかを調べる。

(5) トップレベルドメイン名の種類との関連性

URLの持続性およびWeb上での入手可能性とトップレベルドメイン名の種類（go, acなど）に関連性があるかを調べる。

4. 調査結果

4.1 参照されたネット文献の概要

4.1.1 ネット文献の数

表1は、1999年調査と比較して、参照されたネット文献数をまとめたものである。なお1999年は、1年間の論文を調査している。

表1 参照されたネット文献の概要

雑誌名	2007 (半年)								2005 (半年)								1999 (1年間)							
	全論文数 a	ネット文献を参照している論文数 b	b/a	全参照文献数 c	参照されたネット文献数 d	d/c	1論文あたりの参照したネット文献数d/a	1論文あたりの参照したネット文献数d/a	全論文数 a	ネット文献を参照している論文数 b	b/a	全参照文献数 c	参照されたネット文献数 d	d/c	1論文あたりの参照したネット文献数d/a	1論文あたりの参照したネット文献数d/a	全論文数 a	ネット文献を参照している論文数 b	b/a	全参照文献数 c	参照されたネット文献数 d	d/c	1論文あたりの参照したネット文献数d/a	1論文あたりの参照したネット文献数d/a
情報管理	20	15	75%	231	141	61%	11.6	7.1	21	15	71%	340	184	54%	16.2	8.8	52	13	25%	306	32	10%	5.9	0.6
情報の科学と技術	32	22	69%	337	113	34%	10.5	3.5	33	23	70%	301	168	56%	9.1	5.1	68	26	38%	581	193	33%	8.5	2.8
現代の図書館	13	8	62%	254	98	39%	19.5	7.5	13	11	85%	324	99	31%	24.9	7.6	28	6	21%	230	28	12%	8.2	1.0
日本図書館情報学会誌	5	3	60%	256	36	14%	51.2	7.2	5	4	80%	143	19	13%	28.6	3.8	7	1	14%	205	1	0%	29.3	0.1
合計	70	48	69%	1078	388	36%	15.4	5.5	72	53	74%	1108	470	42%	15.4	6.5	155	46	30%	1322	254	19%	8.5	1.6

ネット文献を参照している論文の数は、全論文数のうち、2007年が69%、2005年が74%であった。1999年調査では30%であったので、ネット文献を参照する論文の比率は、2.3–2.5倍に増加している。この間インターネット情報源が充実したことから当然であろう。

全参考文献に対して、ネット文献の占める割合は、36%（2007）、42%（2005）で、1999年調査の19%に比べこれも大幅に増加している。

1論文あたりの参考文献数は、1999年調査が8.5件であるのに対して、2007・2005年ともに15.4件と1.8倍に増加している。1論文あたりの参照されたネット文献数も、1999年調査が1.6件であるのに対して、2007年5.5件、2005年6.5件、と増加している。

2007年の方が2005年よりも、参照されたネット文献が減少しているのは、『情報の科学と技術』のネット文献が全参考文献数に対して、56%（2005）から34%（2007）にその割合が減少したためである。これは2005年刊行の同誌では、「情報ポータル」「図書館コンソーシアム」「電子ジャーナル」というような特集が続いたのに対し、2007年では「統制語彙」「図書館への提言」「図書館情報学の研究動向」の特集が生まれ、テーマの特性からネット文献が少なかったからである。このように、ネット文献の参照の度合いは、論文のテーマや内容によって大きく変動する。

しかし、全体として、全参考文献のうちネット文献は4割程度を占めるようになってきたといえる。

1999年調査の全参考文献数1,322件および、参照されたネット文献254件を単純に半年分として1/2にして、図1のように、2005、2007年と比較すると、参考文献数は約1.6倍増加しているが、従来の紙媒体の

図1 参照されたネット文献の増加

参考文献がインターネット情報源に移行したというよりも、ネット文献の参照の増加が、全体の参考文献数を増加させていることがわかる。

4.1.2 参照されたネット文献の特徴

参照されたネット文献数は、2007年が388件、2005年が470件で、この合計858件が今回の調査対象文献である。

ネット文献の種類は、Webページ（37%）およびWebサイト全体（23%）が多く、論文（会議論文を含む）（15%）、報告書（10%）といった学術文献が続き、基準やガイドライン等（5%）やプレスリリース（5%）、答申、通達等（3%）がこれに続く（表2参照）。論文に参照された文献ではあるが、必ずしも学術的な文献ばかりでなく、WebサイトやWebページ、プレスリリースなどさまざまな種類が参照されていることがわかる。

参照されたネット文献のトップレベルドメイン名の種類は、年によって若干の違いがあるが、平均す

表2 文献の種類

	Webページ		Webサイト		論文		報告書		基準、ガイドライン等		プレスリリース		答申、通達等		統計		図書		不明		合計	
2007	156	40%	88	23%	57	15%	24	6%	32	8%	12	3%	10	3%	6	1.5%	2	0.5%	1	0.3%	388	100%
2005	162	34%	110	23%	71	15%	62	13%	13	3%	27	6%	15	3%	0	0%	2	0.4%	8	1.7%	470	100%
合計	318	37%	198	23%	128	15%	86	10%	45	5%	39	5%	25	3%	6	0.7%	4	0.5%	9	1.0%	858	100%

表3 トップレベルドメイン名

	or, org		go, gov		ac, ed, edu		co, com		ne, net		地方自治体		その他		合計	
2007	126	32%	137	35%	38	10%	49	13%	7	2%	16	4%	15	4%	388	100%
2005	136	29%	64	14%	114	24%	80	17%	8	2%	13	3%	55	12%	470	100%
合計	262	31%	201	23%	152	18%	129	15%	15	2%	29	3%	70	8%	858	100%

表4 言語

	英語		日本語		その他		合計	
2007	197	51%	170	44%	21	5%	388	100%
2005	288	61%	151	32%	31	7%	470	100%
合計	485	57%	321	37%	52	6%	858	100%

ると、団体の「or, org」が31%と最も多く、政府機関等の「go, gov」が23%で、教育機関の「ac, ed, edu」(18%), 企業の「co, com」(15%) がこれに続く。わが国の「地方自治体」は3%, ネットワークの「ne, net」が2%であった。「その他」には、国際機関やその他のccTLD (country code Top Level Domain) などが含まれている(表3参照)。

言語については、英語(57%), 日本語(37%)で両年ともに英語のネット文献が多い傾向が見られる(表4参照)。

4.2 URLの持続性

記述されたURLにアクセスし、該当するネット文献を入手できたのは、2007年では72%(279件)、入手できなかったのは28%(109件)であった。2005年では59%(278件)が入手でき、入手できなかったのは41%(192件)であった(表5参照)。つまり、3年前に参照されたURLの約30%, 5年前になると約40%が失われ、論文の著者が参照したURLでは参考文献を入手できなくなった。

Casserlyらの調査^{6), 7)}では、URLの持続性は、2-3年前で56.4%, 6年後の追跡調査で39%に減少したと報告されている。また、Gohら⁸⁾は、1年前では70%, 2-3年前で約65%, 4年前で60%がアクセス可能で、5年で半減期を迎えると報告している。したがって、本調査のURLの持続性の結果は、Gohらの調査にはやや近いが、いずれの調査よりもアクセス可能性が高い結果となった。

この要因のひとつとしては、本調査ではCasserlyらとGohらの調査よりも、サイトの参照が多く、なおかつその持続性が高かったためと考えられる。

CasserlyらとGohらは、URLのディレクトリの深さと持続性の関係を調査している。これは、URLのディレクトリの深さの0レベル(例えばhttp://aaa.bbb.ccc/)から7レベルまでの持続性を調査したもので、この0レベルが、本調査の文献の種類におけるサイトにほぼ該当する。Casserlyらの調査では、0レベルが全体の15.2%で、その持続性は68.4%⁹⁾、追跡調査では62.2%¹⁰⁾であった。Gohらの調査¹¹⁾では、0レベルは全体の23%を占め、持続性は74.4%であった。

しかし、本調査では4.4に述べるように、サイトの参照は全体の23%前後を占め、URLで入手できたのは、91%(2007年)、82%(2005年)と高かった。これが、全体のURLの持続性を押し上げているひとつの要因と考えられる。

表5 URLの持続性

	2007				2005			
	入手可能		入手不可		入手可能		入手不可	
情報管理	105	74%	36	26%	126	68%	58	32%
情報の科学と技術	74	65%	39	35%	90	54%	78	46%
現代の図書館	76	78%	22	22%	52	53%	47	47%
日本図書館情報学会誌	24	67%	12	33%	10	53%	9	47%
合計	279	72%	109	28%	278	59%	192	41%

図2 Web上の入手法と入手率 (2007)

図3 Web上の入手法と入手率 (2005)

そのほか、Casserlyらの調査では、著者の書誌記述が不十分なものが多く、それに伴う入手の判断の違いなどもあるのではないかと考えられる¹²⁾。

なお、ネット文献の言語によるURLの持続性の違いは、ほとんどなかった。したがって、本調査が日本語のネット文献も対象としていることは、特別な影響を与えるものではないと判断した。

4.3 Web上での入手可能性

URLで見つからなかったネット文献全301件（2007年109件、2005年192件）を、トップページに戻ってサイト内検索をしたり、トップページのメニューから探したりして、サイト内で入手できた場合と、Googleで検索したり、案内されたWebアーカイブで検索してWeb上に存在した文献の割合を調べた。

その結果、図2、3に示したように、2007年は、入手できなかった109件のうちサイト内検索で35%、トップページからの探索で4%、Google検索で20%、Webアーカイブで2%見つけることができた。全体では、61%が入手でき、32%が入手できなかった。残り7%は著者が参照した文献なのかどうかの判断がつかなかった。

2005年では、URLで入手できなかった192件のうちサイト内検索で26%、トップページからの探索で7%、Google検索で21%、Webアーカイブで2%が入手で

表6 Web上での入手可能性

	2007		2005	
URLで入手可 (a)	279	72%	278	59%
Web上で入手可 (b)	66	17%	108	23%
入手できた全件数 (a+b)	345	89%	386	82%
入手できなかった件数 (c)	43	11%	84	18%
合計 (a+b+c)	388	100%	470	100%

き、全部で56%が入手可能であった。残り34%が入手不可能で、10%は判断ができなかった。

入手不可能な参考文献の中には、著者名やタイトルが書誌事項に記述されておらず、本文との関係も明確でないため、サイト内検索や検索エンジンによる検索が的確にできなかった文献も含まれている。

この結果、表6のように、サイト内検索や検索エンジンを利用して、2007年では最終的に入手できたネット文献は89%、2005年では82%になり、2年間で7%減少するものの、かなりの参照されたネット文献がWeb上のどこかにはあり、入手可能になることがわかった。しかし、それでも入手できないネット文献(判断不能の文献を含む)は、3年前では11%、5年前では18%あることになる。

Casserlyらの調査では、Web上で入手できたネット文献は、2-3年前で81.4%、6年後の追跡調査では73.2%で、8.2%減少している。

4.4 文献の種類との関連性

表7は、ネット文献の種類と、URLの持続性および

Web上の入手可能性の関係を示している。

2007年では、Webサイト（91%）と論文（84%）が、著者が記述したURLで入手できた割合が高い。

一方2005年では、全体的に入手の割合が低くなるが、Webサイトが82%と変わらず最もURLの持続性が高い。論文のURLの持続性が他に比べて低くなっているのは、複数の同じURLの記述ミスが生じていたり、URLで入手不可能な同じ文献が何回か参照されていることが、割合を引き下げている原因だと考えられる。

WebサイトのURLの持続性が他に比べ高いことがわかったが、本調査を見るかぎりでは、それ以外の文献の種類による差は明確ではなかった。

また、Web上で入手可能性を比較すると、2007年では文献の種類にかかわらず、かなり見つけることができ、最終的には80%以上入手できているが、比較すると、Webページの割合が低い。2005年になると、Webページやプレスリリースの最終的な入手可能性が低くなる。論文や報告書、基準などは、書誌事項が明確なこともあり、Web上で見つけられる割合が高い。

4.5 トップレベルドメイン名の種類との関連性

トップレベルドメイン名の種類によって、URLの持続性とWeb上の入手可能性が異なるかを調べた（表8参照）。

2007年では、「co, com」が82%でURLの持続性が高く、「go, gov」（74%）,「or, org」（72%）,「ac, ed, edu」（71%）と続き、地方自治体が、全体の件数が少ないとはいえ、31%と他と比較して最も低かった。

2005年では、それぞれ持続性は下がるが、「co, com」,「go, gov」,「or, org」,「ac, ed, edu」は54%から64%で大体同じくらいとなり、トップレベルドメイン名の種類による差は少なくなったが、地方自治体はやはり全体の件数が少ないものの、23%と最も低かった。

地方自治体のURLの持続性は、他の機関のサイトに比べて低いことがわかった。例えば、地方自治体の調査の中間報告が、本報告が掲載されると削除されているケースが見られた。

URLで入手できなかったネット文献が、Web上で入

表7 文献の種類との関連性

		Webページ		Webサイト		論文		報告書		基準、ガイドライン等		プレスリリース		答申、通達等		統計		図書		不明		合計	
2007	URLで入手可	101	65%	80	91%	48	84%	14	58%	19	59%	7	58%	4	40%	4	67%	2	100%	0	0%	279	72%
	Web上で入手可	24	15%	5	6%	8	14%	7	29%	12	38%	4	33%	4	40%	2	33%	0	0%	0	0%	66	17%
	入手できた全件数	125	80%	85	97%	56	98%	21	88%	31	97%	11	92%	8	80%	6	100%	2	100%	0	0%	345	89%
	入手できなかった件数	31	20%	3	3%	1	2%	3	13%	1	3%	1	8%	2	20%	0	0%	0	0%	1	100%	43	11%
	合計	156	100%	88	100%	57	100%	24	100%	32	100%	12	100%	10	100%	6	100%	2	100%	1	100%	388	100%
2005	URLで入手可	87	54%	90	82%	32	45%	41	66%	5	38%	15	56%	8	53%	0	0%	0	0%	0	0%	278	59%
	Web上で入手可	32	20%	12	11%	32	45%	17	27%	6	46%	5	19%	3	20%	0	0%	1	50%	0	0%	108	23%
	入手できた全件数	119	73%	102	93%	64	90%	58	94%	11	85%	20	74%	11	73%	0	0%	1	50%	0	0%	386	82%
	入手できなかった件数	43	27%	8	7%	7	10%	4	6%	2	15%	7	26%	4	27%	0	0%	1	50%	8	100%	84	18%
	合計	162	100%	110	100%	71	100%	62	100%	13	100%	27	100%	15	100%	0	100%	2	100%	8	100%	470	100%

表8 トップレベルドメイン名の種類との関連性

		or, org		go, gov		ac, ed, edu		co, com		ne, net		地方自治体		その他		合計	
2007	URLで入手可	91	72%	101	74%	27	71%	40	82%	6	86%	5	31%	9	60%	279	72%
	Web上で入手可	23	18%	22	16%	8	21%	3	6%	1	14%	6	38%	3	20%	66	17%
	入手できた全件数	114	90%	123	90%	35	92%	43	88%	7	100%	11	69%	12	80%	345	89%
	入手できなかった件数	12	10%	14	10%	3	8%	6	12%	0	0%	5	31%	3	20%	43	11%
	合計	126	100%	137	100%	38	100%	49	100%	7	100%	16	100%	15	100%	388	100%
2005	URLで入手可	83	61%	41	64%	61	54%	48	60%	5	63%	3	23%	37	67%	278	59%
	Web上で入手可	34	25%	18	28%	34	30%	14	18%	3	38%	0	0%	5	9%	108	23%
	入手できた全件数	117	86%	59	92%	95	83%	62	78%	8	100%	3	23%	42	76%	386	82%
	入手できなかった件数	19	14%	5	8%	19	17%	18	23%	0	0%	10	77%	13	24%	84	18%
	合計	136	100%	64	100%	114	100%	80	100%	8	100%	13	100%	55	100%	470	100%

手可能かどうかについては、トップレベルドメイン名による大きな差はない。しかし、サンプル数が少ないものの、地方自治体から発信された文献の入手可能性は低いことがわかった。

5. 考察

5.1 参照されたネット文献の保存

本調査で、参照されたネット文献のURLは、3年前では約3割、5年前では約4割が見つからなくなるといことがわかった。URLの持続性は、年とともにますます減少すると考えられる。URLは、あくまでも著者が参照した時点での、その文献の位置するホストのドメイン名やファイル名であって、永久にその場所に存在するとは限らない。

また、著者が参照した情報内容そのものも、更新、消去される可能性がある。本調査では、論文の著者が記述したURLと合致し、記述されたタイトルと合った文献はアクセスできたと判断したが、確かに著者が参照した文献であるという同定の確証は、実際は得ることができない。ネット文献の更新年月日は明確に書かれていない場合が多く、参照年月日と比べることができない。たとえ更新年月日が書かれていて、著者の参照年月日と現在の更新日が違うことがわかったとしても、どこが違うのかは更新履歴がないので明確にわからない。そのため、著者が参照年月日を書いても、現在は有効に機能していない。

紙媒体ならば、改訂や絶版になっても、出版者が廃業しても、どこかに存在し、図書館に保存されている。つまり現在のところ、ネット文献は、参照文献の役割である、読者によるアクセスと学問的検証を完全に保証しえない媒体と言える。論文の著者としては、ネット上の文献を参照する場合には、このことを十分考慮に入れて参照することが求められる。

この問題の解決には、URLに左右されないDOI (Digital Object Identifier) やISBNのような固有の識別記号を付与することが、ひとつの方法であることは

確かである。論文の著者も、このような固有識別記号のある文献を参照し、参照文献の書誌要素として記述するようにすれば読者の入手が容易になるだろう。

しかし、調査に見るように、参照されたネット文献の種類は学術論文だけでなく、多岐にわたっている。それらすべてに、統一した固有識別記号を付与することはできない。また、識別記号とともに、Web上で後年入手が可能でなければならない。

つまり、参照した時点のネット文献の保存が急務なのである。

まず、論文の著者が、自分で参照したネット文献の複製と保存をしておかなければならないだろう。もし、読者からの問い合わせがあったら、著者としては参考文献を提供できなければならない。

参照されたネット文献を保存する母体としては、学術雑誌を刊行している学協会や出版者も考えられる。原稿の提出や査読の時に、参照したネット文献を提供してもらう必要が出てくるだろう。法的な問題は残されているが、学協会や出版者が参照されたネット文献を保存することも考えられる。

次項に述べるように、ネット文献の発信者やサイト運営者もWebページの保存を考える時期にきている。公的あるいは教育研究機関による、Webアーカイブや機関リポジトリは、世界各国で行われている。わが国では、国立国会図書館のWebアーカイブが、ようやく緒についたところである。著作権や収集方法などさまざまな問題があるが、特に学術分野においては、安定したネット文献の蓄積と保存が求められている。

5.2 サイト運営者の課題

本調査では、参照文献に記述されたURLで見つからなかったとしても、サイト内検索や検索エンジンによる検索が有効で、2007年では89%、2005年でも82%が入手可能となった。

ネット文献の入手可能性には、サイト運営者の管

理や運営の仕方が影響する。

例えば、同一サイト内におけるリンクの変更やサーバーの移動をした際は、自動的に移動した該当のURLに案内してあれば、探す時間も手間も省かれ、適切なアドレスを入手できる。また、サイト内検索があると、ディレクトリが変わっていたとしても、検索しやすく入手可能性が高くなる。

本調査では、サイト内にアーカイブを作成したり、あるいは外部のWebアーカイブに保存されていて、そのことを案内しているサイトもあった。例えば、総務省のページにアクセスすると、当該ページは、掲載期間が終了しているというメッセージとともに、国立国会図書館のWARPにアクセスして検索をするように案内され、入手することができた。

一方、Webページに、更新年月日が記載されていないものがかなり多かった。著者が参照年月日を記述していたとしても、被参照文献に更新年月日、あるいは更新履歴がなければ、著者が参照した文献であることの判断がつかない。

サイト運営者は、これまでは、情報を発信するツールとしてWebサイトを利用し、その掲載情報の更新や消去、保存に対して、特に方針を持っていなかったと思われる。今後は、特に教育、行政、学術研究の分野の公的なサイト運営者は、発信する情報の履歴を記録し、消去する情報と保存する情報を区別し、保存の基準を作り、それを公開して、保存する情報はサイト内アーカイブを構築するか、あるいは外部アーカイブのどこに保存されているかの情報を公表して、保存されたネット文献の検索を可能にしておく必要がある。

5.3 入手可能性を高める書誌記述

サイト内検索や検索エンジンで検索する場合は、通常参照文献リストに記述された著者名やタイトル等で検索する。参照文献リストに、必要十分な書誌要素が書かれていることが、URLを記述することよりも、実は重要なことが今回の調査で明らかになった。

すなわち、URLだけ書いていても、それは消失する可能性があるが、著者名やタイトル等が記述されていれば、なんらかの方法で探すことができる。

調査対象とした4誌は、いずれも投稿規程にWeb情報の記述の仕方が例示されている。しかしながら、本調査では参照文献リストに、URLだけしか記述されていなかったり、機関名やサービス名とURLだけしか表示されていないものが依然としてあった。これではサイト内検索でも検索エンジンでも探すことができない。それは次項の引用方法にも関係している。

また、Webページには通常ページ付けがないため、参照した部分がどこかを探すのに時間がかかったり、同定することが難しい。本調査では、「第2段落参照」のように段落を示し、同定を迅速にできた例があった。Webページの位置情報をどうするのかも、今後の課題であろう。

ページに関して、調査で気がついたことは、参照文献がPDFで、通常のページがあるにもかかわらず、何故かネット文献になると、参照ページを書かなくなる傾向がある。ネット文献であっても、基本的な書誌記述は紙媒体と同様に記述しなければならない。

インターネット上の情報源には、紙媒体の奥付に当たるような統一的な書誌要素の記述がまだない。したがって、著者は、ネット文献を参照する場合には、読者が後に入手できるように必要十分な書誌要素を記述することが重要である。

学術雑誌の編集者も、投稿された原稿を検証・査読する際に、参照されたネット文献の書誌的記述にも注意を払う必要がある。

5.4 ネット文献の参照方法

ネット文献の参照の増加に伴って、参照の仕方に問題がでてきている。

ひとつは、「JAPIC¹⁾は…」 「MEDLINE²⁰⁾, Dialog²¹⁾」というように機関名やサービス名等に注番号をつけ、参照文献リストには、そのトップページのURLだけを記述している参照方法である。これらは、そのネッ

ト文献を参照して本文を記述しているのではなく、単にその機関やサービスのWebサイトがあることを示しているにすぎない。これは、本調査で、Webサイトの参照が23%を占めている理由の一因になっている。Huberは、こうした参照を「余計なURL」であるとして、「参照としては使用すべきでない」と指摘している¹³⁾が、同感である。意味のないトップページの参照とURLのみの表示は、避けるべきであろう。

もうひとつは、「自治体でも…する動きが出始めた⁵⁾。」のように、本文には参考文献の著者名も引用文もなく、また要約引用でもなく、事実や意見を記述するにあたって参照した文献を注番号で示しているような場合である。従来の紙媒体でもこのような参照の仕方は見られたが、これまでは、出版年と該当ページなど必要な書誌要素が書かれており、著者が参照した内容と個所を特定することができた。

しかし、これがネット文献になると、著者が記述したURLにそのタイトルの文献があったとしても、ページ付けがないと、著者の参照した文章がどこにあるか探すのに時間がかかり、引用文がないために、それが同じかどうかを識別することが難しくなる。内容が更新されている場合には、さらに論文の著者がどこを参照したのかを特定することは難しい。まして、書誌事項が不十分で、かつURLが消滅してしまうと、検索のためのキーワードの選定ができないため、探すこともできなくなってしまう。

著者は、ネット文献を参照する場合には、読者が同じ情報内容を迅速に見つけられるように、本文中に明確に参照した文章や情報を記述する方法を取り、できるだけあいまいな参照を避けなければならない。

6. おわりに

現在のところ、失われていくネット文献を確実に

つなぎ止めておく方法はない。しかし、われわれはインターネット上の情報源を参照しないわけにはいなくなっている。

論文の著者は、ネット文献は入手ができなくなるものであることを認識して、安易な参照を避け、慎重に文献を選択し、読者が参照した文献だと同定できるような適切な方法で参照し、後年検索エンジンを含むさまざまな方法で見つけられるように、必要十分な書誌事項を記述し、少なくとも自分自身はその参考文献を複製保存することを心がけることが必要だろう。

学術論文を刊行する学協会や出版者の編集者も、参照されたネット文献が適切で、書誌事項が必要十分であるかをチェックし、今後その蓄積保存も考慮する必要がでてくるだろう。

教育機関、官公庁、学術団体のサイト管理者も、発信した情報の年月日を付与し、保存する必要がある情報については、サイト内での蓄積が求められるだろう。図書館は自らもアーカイブを構築するとともに、アーカイブ構築を支援し、その所在情報や内容情報を蓄積し、検索できるようにする役割を担うことになる。

本調査は、2007年と2005年の、それぞれの論文に参照されたネット文献の入手可能性を調査したが、参照されるネット文献は、分野やテーマによって大きく異なる。したがって、2007年の参照されたネット文献と同じURLを数年後に調査したり、別の分野や期間、方法で調査をすることにより、より客観的なデータが得られるであろう。

最後に、本稿をまとめるにあたって、仲本秀四郎氏ならびに古谷実氏に適切なご助言をいただいた。ここに心から感謝の意を表し、お礼申し上げます。

参考文献

- 1) Wagner, Cassie et al. Disappearing act: Decay of uniform resource locators in health care management journals. *Journal of the Medical Library Association*. 2009, vol. 97, no. 2, p. 122-123. doi:10.3163/1536-5050.97.2.009.
- 2) Casserly, Mary F.; Bird, James E. Web citation availability: Analysis and implications for scholarship. *College & Research Libraries*. 2003, vol. 64, no. 4, p. 300-317. <http://crl.acrl.org/content/64/4/300.full.pdf>, (accessed 2010-10-01).
- 3) Casserly, Mary F.; Bird, James E. Web citation availability: A follow-up study. *Library Resources & Technical Services*. 2008, vol. 52, no. 1, p. 42-51.
- 4) Goh, Dion H.; Ng, Peng K. Link decay in leading information science journals. *Journal of the American Society for Information Science and Technology*. 2007, vol. 58, no. 1, p. 15-24.
- 5) 藤田節子. 図書館情報学分野における電子文献の参照調査. *実践女子短大評論*. 2000, no. 21, p. 102-123.
- 6) Casserly, Mary F.; Bird, James E. Web citation availability: Analysis and implications for scholarship. *College & Research Libraries*. 2003, vol. 64, no. 4, p.307.
- 7) Casserly, Mary F.; Bird, James E. Web citation availability: A follow-up study. *Library Resources & Technical Services*. 2008, vol. 52, no. 1, p. 44-45.
- 8) Goh, Dion H.; Ng, Peng K. Link decay in leading information science journals. *Journal of the American Society for Information Science and Technology*. 2007, vol. 58, no. 1, p. 17-18.
- 9) Casserly, Mary F.; Bird, James E. Web citation availability: Analysis and implications for scholarship. *College & Research Libraries*. 2003, vol. 64, no. 4, p. 313.
- 10) Casserly, Mary F.; Bird, James E. Web citation availability: A follow-up study. *Library Resources & Technical Services*. 2008, vol. 52, no. 1, p. 49.
- 11) Goh, Dion H.; Ng, Peng K. Link decay in leading information science journals. *Journal of the American Society for Information Science and Technology*. 2007, vol. 58, no. 1, p. 37.
- 12) Casserly, Mary F.; Bird, James E. Web citation availability: Analysis and implications for scholarship. *College & Research Libraries*. 2003, vol. 64, no. 4, p. 305-307.
- 13) Huber, Andy. “余計なURL”. ホームページは文献として引用できるか? . Vetter, Ron編. 宍戸博訳. *情報処理*. 1998, vol. 39, no. 7, p. 663-664. 原論文: URLs in print: To cite or not to cite. *IEEE Computer*. 1998, vol. 31, no. 2, p. 114-117.

参考資料

- a) Dellavalle, Robert P. et al. Going, going, gone: Lost Internet references. *Science*. 2003, vol. 302, no. 5646, p. 787-788.
- b) Falagas, Matthew E.; Karveli, Efthymia A.; Tritsaroli, Vassiliki. The risk of using the Internet as reference resource: A comparative study. *International Journal of Medical Informatics*. 2008, vol. 77, no. 4, p. 280-286. doi:10.1016/j.ijmedinf.2007.07.001.
- c) 指宿信. “学術文献としてのオンライン情報の引用”. *法情報学の世界*. 第一法規, 2010, p. 258-270.
- d) Koehler, Wallance. A longitudinal study of Web pages continued: A consideration of document

persistence. Information Research. 2004, vol. 9, no. 2, paper 174. <http://InformationR.net/ir/9-2/paper174.html>, (accessed 2010-10-10).

- e) 二通信子. 論文の引用に関する基礎的調査と引用モデルの試案. アカデミック・ジャパニーズ・ジャーナル. 2009, vol. 1, p. 65-74.
- f) Sellitto, Carmine. A study of missing Web-cites in scholarly articles: Towards an evaluation framework. Journal of Information Science. 2004, vol. 30, no. 6, p. 484-495. doi:10.1177/0165551504047822.
- g) Zhang, Yin. The impact of Internet-based electronic resources on formal scholarly communication in the area of library and information science: A citation analysis. Journal of Information Science. 1998, vol. 24, no. 4, p. 241-254.

Author Abstract

Recently, the use of Internet resources has increased in the references of articles in scholarly journals. Readers may not be able to access to the Internet references later, because the Web pages are frequently revised or removed. This study examines the availability of 858 Internet references cited or referred to in the articles published in four library and information science journals in 2005 and 2007. 28% of the Internet references in 2007 could not be located at the URLs that authors of articles provided, and 41% could not be located in 2005. The author tried to access the Internet references that were not found at the URLs by searching in the Web sites and using Google. The author found that 89% of the Internet references were available on the Web in 2007, and 82% were available in 2005. Archiving of Internet references, proper management of Web sites, correct citations and complete bibliographic description will improve the accessibility to the references from sources on the Internet.

Key words

Internet references, Web references, cited resources, Internet resources, link decay, decay of URL, Web archive, institutional repositories, description of bibliographic references