第20回年次大会予稿

日本の学協会誌掲載論文の機関リポジトリ収録状況

The rate of Japanese journal articles archived in institutional repositories

清水真理[1]，佐藤翔[2*]，逸村裕[3]
Mari SHIMIZU[1], Sho SATO[2*], Hiroshi ITSUMURA[3]

1 富山市立図書館
Toyama City Public Library
〒930-0085 富山県富山市丸の内1-4-50
E-mail: shimizu.mari@city.toyama.lg.jp
2 筑波大学大学院 図書館情報メディア研究科
Graduate School of Library, Information and Media Studies, University of Tsukuba
〒305-8550 茨城県つくば市春日1-2
E-mail: min2fly@slis.tsukuba.ac.jp
3 筑波大学 図書館情報メディア系
Faculty of Library, Information and Media Science, University of Tsukuba
〒305-8550 茨城県つくば市春日1-2
E-mail: hits@slis.tsukuba.ac.jp
*連絡先著者 Corresponding Author

本研究では学協会著作権ポリシーデータベース（SCPJ）とCiNiiを用い，日本の学協会誌掲載論文の機関リポジトリ収録状況を分析した．分析の結果，2000～2009年の日本の学協会誌掲載論文の機関リポジトリ収録率は約0.9%にとどまっており，いずれの分野でも収録率は低かった．収録を許可するポリシーの雑誌は許可しない雑誌より収録率が高かったが，それでも収録率は1.3%程度にとどまっていた．

We analyzed the rate of Japanese journal articles archived in institutional repositories based on the data of Society Copyright Policies in Japan (SCPJ) database and CiNii. As a result of analysis, only about 0.9% of Japanese journal articles published during 2000-2009 were archived in institutional repositories. The archive rate was low in all fields. While articles published in Green, Blue or Yellow journals were more archived than others, the archive rate of those was only about 1.3%.

キーワード: 機関リポジトリ, Open access, 著作権, 学術出版, CiNii
Institutional repository, open access, copyright, scholarly publishing, CiNii

1 はじめに

本研究の目的は日本の学協会が発行する学術雑誌に掲載された論文の，機関リポジトリへの収録状況を明らかにすることである．

機関リポジトリとは，「大学とその構成員が創造したデジタル資料の管理や発信を行うために，大学がそのコミュニティの構成員に提供する一連のサービス」であり[1]，学術情報への障壁のないアクセスを目指すオープンアクセス（OA）運動の一翼を担うものとしてあらわれた．2012年4月現在，世界全体で1,799[2]，日本国内で155の機関リポジトリが存在する[3]．日本国内での公開文献数は，本文が閲覧可能なものに限っても97万件を超えており，過去1年間で約19万件の文献が新たに収録されている[4]．それらの収録文献はCiNii等を通じ日本の学術情報流通に貢献しているのみならず[5]，非学術的な活動の中でも活用されている[6].

一方で，現在の日本の機関リポジトリ収録文献の過半数（50.5%）は大学・研究機関等が発行する紀要掲載論文であり，OA運動において重要な対象である学術雑誌掲載論文は16.2%にとどまっている[4]．その理由として，紀要論文は発行元が学内組織であるため著作権処理が容易である一方，学術雑誌論文は著作権処理に手間がかかり，著者への収録依頼を必要とすること等が指摘されている[7]．そこで著作権処理の負担を軽減し，機関リポジトリへの論文収録を促進する試みが行われており，ノッティンガム大学が運営するSHERPA/RoMEO[8]，筑波大学等が運営する学協会著作権ポリシーデータベース（SCPJ）[9]等の出版者の著作権ポリシーに関するデータベースが作成されている．

このような取り組みにより学術雑誌論文の機関リポジトリへの収録も進んでいることが考えられるが，どの程度の効果があったのかについて具体的なデータを示した形では明らかになっていない．そもそも日本で出版された論文の機関リポジトリへの収録状況や著作権ポリシーとの関係自体，これまで十分に明らかにされては来なかった．英文誌掲載論文については，2009年度に出版された日本の著者の論文の11.1%が日本の機関リポジトリに収録されていたとする見積りもあるが[10]，概算であり実際の状況を調査したものではない．また，和文誌も含めた日本の学協会誌の状況についてもわかっていない．

そこで本研究ではSCPJ収録誌を対象に，その掲載論文の機関リポジトリへの収録状況について，全体の傾向および発行元である学協会のポリシーや学術分野との関係を調査した．これは日本の機関リポジトリの現状を理解する上での，基礎となるデータを提供するものと考えられる．

2 方法・対象

2.1 分析対象

調査対象はSCPJ収録雑誌中，ISSNが記載されていた2,583誌に2000-2009年に掲載された論文とし，2011年9月時点での機関リポジトリへの収録状況を調査した．2,583誌の対象期間中の掲載論文データについては，国立情報学研究所の論文データベースCiNiiのAPIにより取得した．CiNiiのAPIでは1回に取得できる論文データの上限が200件であったため，年間200本以上論文を掲載している雑誌については新しいものから200件までを取得した．

データ収集の結果，2,583誌中463誌については対象期間中，CiNii収録論文が1本もなかった．この中にはCiNii未収録誌のほか，休刊・廃刊もしくは未創刊の雑誌が含まれると考えられる．これらの雑誌については分析から除外し，1本以上CiNiiに論文が収録されていた2,120誌の掲載論文のみ対象とすることとした．取得した分析対象論文数の合計は1,010,822本である．

2.2 分析方法

分析対象論文の機関リポジトリへの収録状況について，全体および著作権ポリシー別

に集計した．また，学術分野別の状況についても分析した．これは著作権ポリシーと別に，学術分野の特性が機関リポジトリ収録状況に影響する可能性を考慮したためである．

機関リポジトリ収録状況については，CiNiiの書誌画面から機関リポジトリへのリンクが付与されていたものを「収録あり」とみなし分析した．CiNiiから機関リポジトリへのリンクの有無については論文データ収集時にあわせて取得した．

著作権ポリシーについてはSCPJのデータを利用した．SCPJでは学協会の著作権ポリシーを，Green（査読前・後いずれの版も著者によるアーカイブを認める），Blue（査読後の版のみ認める），Yellow（査読前原稿のみ認める），White（アーカイブを認めない），Gray（検討中・非公開・無回答・その他）の5つに分けている．本研究ではこれをいずれかの版の機関リポジトリへの収録が許可されている場合（Green，Blue，Yellow），許可されていない場合（White，Gray）の2つに分けて分析した．

また，学術分野については発行元の学協会に基づき，学会名鑑の学術分野を参照して表1の3領域，30分野に分類した．学会名鑑に分野の記載がない学協会や未収録の学協会については内容を参照し，著者が独自にいずれかの分野に分類した．分野が多岐に渡る雑誌や30分野のいずれにも含まれないと判断した雑誌は複合領域として区分した．

表1 学術分野区分

領域	分野	雑誌数	論文数
人文・社会	言語・文学	101	18,249
	哲学	66	13,436
	心理学・教育学	161	37,469
	社会学	55	11,305
	史学	97	23,186
	地域研究	47	8,699
	法学	36	9,659
	政治学	9	2,704
	経済学	50	10,969
	経営学	46	11,545
生命科学	基礎生物学	83	29,032
	統合生物学	15	7,166
	農学	136	68,706
	食料科学	14	8,883
	基礎医学	106	58,146
	臨床医学	346	217,906
	健康・生活科学	107	41,607
	歯学	63	29,443
	薬学	32	23,691
理学・工学	環境学	35	16,759
	数理科学	24	7,155
	物理学	23	23,034
	地球惑星科学	70	35,273
	情報学	45	17,614
	化学	50	42,875
	総合工学	83	73,966
	機械工学	40	30,179
	電気電子工学	37	38,667
	土木工学・建築学	29	24,683
	材料工学	32	35,331
複合領域		59	18,102

3 分析結果

3.1 機関リポジトリ収録状況

分析対象論文1,010,822本中，機関リポジトリに収録されていたのは18,604本で，1.8%にとどまった．さらに機関リポジトリ収録論文の多くは特定の大学等所属者を中心に運営されている学協会等の発行誌（いわゆる学内学会誌）掲載論文が，当該機関のリポジトリに収録されている場合であった（他大学のリポジトリにはほとんど収録されていなかった）．特に機関リポジトリ収録論文数の多い学内学会誌上位23誌のみで10,000本以上の論文が収録されており，これを除くと総論文数995,619本に対し機関リポジトリ収録論文は8,518本（0.9%）で，収録率は1%未満まで下がる．学内学会誌は紀要に近いものと考えれば（実際に各機関リポジトリ上で「雑誌論文」ではなく「紀要論文」として収録されている場合もあった），実質的に日本の学協会誌掲載論文で機関リポジトリに収録されているものの割合は1%に満たないと言えよう．

なお，出版年別に見ると2007年までは新しい論文ほど機関リポジトリ収録率が高い一方，2008・2009年については収録率が下がっていた．出版後機関リポジトリ等への収録を制限する猶予期間（エンバーゴ）が影響している可能性が考えられる．

3.2 著作権ポリシーとの関係

表2は著作権ポリシーと機関リポジトリ収録状況の関係を示したものである．タイトル単位で見ると機関リポジトリへの収録が許可されている割合は約33%にとどまる．しかし論文単位で見ると収録が許可されているものの割合は約45%にのぼっており，半数には至らないまでも多くの論文が，機関リポジトリへの収録が発行元によって許可された状態にあると言える．

表 2 著作権ポリシーとリポジトリ収録状況

著作権ポリシー	雑誌数	論文数	リポジトリ収録数	リポジトリ収録率
学内学会誌含む				
収録可	690	451,375	11,674	2.6%
収録不可	1,430	559,447	6,930	1.2%
学内学会誌除く				
収録可	679	442,166	5,683	1.3%
収録不可	1,418	553,453	2,835	0.5%

表 3 領域別の機関リポジトリ収録状況

領域	総論文数	収録許可論文数	収録許可率	リポジトリ収録数	リポジトリ収録率
人文・社会	147,221	49,168	33.4%	1,415	1.0%
生命科学	484,580	182,299	37.6%	3,570	0.7%
理学・工学	345,536	204,624	59.2%	3,439	1.0%
複合領域	18,102	5,895	32.6%	94	0.5%

表 4 分野別の機関リポジトリ収録状況

領域	総論文数	収録許可論文数	収録許可率	リポジトリ収録数	リポジトリ収録率
機械工学	30,179	22,097	73.2%	842	2.8%
数理科学	7,155	3,808	53.2%	156	2.2%
電気電子工学	38,667	30,865	79.8%	637	1.6%
基礎生物学	29,032	18,762	64.6%	469	1.6%
情報学	17,614	11,808	67.0%	271	1.5%
哲学	13,436	5,124	38.1%	203	1.5%
薬学	23,691	8,550	36.1%	355	1.5%
心理学・教育学	37,469	13,247	35.4%	545	1.5%
史学	23,186	6,570	28.3%	273	1.2%
地球惑星科学	35,273	15,880	45.0%	395	1.1%
農学	68,706	36,121	52.6%	706	1.0%
食料科学	8,883	3,814	42.9%	90	1.0%
環境学	16,759	8,624	51.5%	167	1.0%
言語・文学	18,249	7,591	41.6%	156	0.9%
基礎医学	58,146	24,889	42.8%	424	0.7%
地域研究	8,699	3,238	37.2%	60	0.7%
健康・生活科学	41,607	17,428	41.9%	273	0.7%
化学	42,875	13,414	31.3%	277	0.6%
統合生物学	7,166	5,683	79.3%	45	0.6%
歯学	29,443	5,852	19.9%	176	0.6%
物理学	23,034	20,901	90.7%	130	0.6%
経済学	10,969	3,843	35.0%	58	0.5%
複合領域	18,102	5,895	32.6%	94	0.5%
経営学	11,545	4,222	36.6%	56	0.5%
材料工学	35,331	17,088	48.4%	168	0.5%
臨床医学	217,906	61,200	28.1%	1,032	0.5%
総合工学	73,966	48,711	65.9%	326	0.4%
土木工学・建築学	24,683	11,428	46.3%	70	0.3%
社会学	11,305	1,501	13.3%	32	0.3%
政治学	2,704	1,106	40.9%	7	0.3%
法学	9,659	2,726	28.2%	25	0.3%

これらの機関リポジトリへの収録が許可されている論文は，許可されていない論文に比べて有意に機関リポジトリへの収録率が高かった（カイ二乗検定，有意確率<0.01）．ただし高いと言ってもそのリポジトリ収録率は2.6%とわずかである（収録不可の場合は1.2%）．さらにこの中には前述の学内学会誌掲載論文が含まれており，それを除いた場合の収録率は1.3%に下がる（収録不可の場合も0.5%に下がるので，収録が許可されている雑誌の方が収録率が高い傾向は変わらない）．

機関リポジトリへの収録を許可されたものの方が実際に収録される割合が高いことは確かであるが，大部分の論文は学協会によって機関リポジトリへの収録が許可されていても，収録されないままになっていることがわかる．

3.3 分野との関係

表3は領域別，表4は分野別の，機関リポジトリへの収録が許可されている論文の割合と，実際の機関リポジトリ収録状況を示したものである（いずれも学内学会誌は除いている）．

領域別に見ると，論文数の少ない複合領域を除けば，生命科学領域で他の領域に比べリポジトリ収録率が低い．カイ二乗検定よりこの差は有意であった（有意確率<0.01）．領域によって機関リポジトリへの収録率には差があると言える．機関リポジトリへの収録が許可されている論文の割合（収録許可率）については人文・社会領域の方が生命科学領域よりも低いことから，この差は著作権ポリシーの影響によるものではないと考えられる．

ただし表4から，分野別により詳細な状況を見ると，同一領域に属する分野であっても機関リポジトリ収録状況は大きく異なることがわかる．機関リポジトリ収録率が最も高いのは機械工学分野（2.8%）であるが，同じ理学・工学領域，中でも機械工学により近いと考えられる工学領域の中でも，材料工学（0.5%），総合工学（0.4%），土木工学・建築学（0.3%）等では機関リポジトリ収録率はごく低い値にとどまっている．また，基礎生物学（1.6%）と総合生物学（0.6%）も隣接分野間で機関リポジトリ収録率に大きく差が開いている．これらの分野による差についても領域別の結果と同様に，カイ二乗検定より有意確率<0.01で有意であった．

一方，領域別の結果と異なり，分野間の機関リポジトリ収録率の差については著作権ポリシーとの関係も見受けられた．収録許可率と機関リポジトリ収録率の間には中程度の有意な相関関係があり（スピアマンの順位相関，ρ=0.406，有意確率<0.05），機関リポジトリへの収録が許可されている論文の割合が多い分野ほど，実際の収録率も高い傾向があった．ただし例外もあり，収録許可率が最も高い物理学（収録許可率90.7%），3番目に高い総合生物学（79.3%）ではいずれもリポジトリ収録率は0.6%にとどまっていた．逆に収録許可率が低いのにリポジトリ収録率が高い分野は哲学（収録許可率38.1%，収録率1.5%），心理学・教育学（収録許可率35.4%，収録率1.5%），史学（収録許可率28.3%，収録率1.2%）など，人文・社会領域に多かった．

もっとも，リポジトリ収録率が高い／低いと言っても，最も収録率の高い分野でも2.8%と3%に至っておらず，いずれの分野においても，収録が許可されているものも含めてほとんどの論文は機関リポジトリに収録されていないと言える．なお，学内学会誌を含めた場合，非常に掲載論文数の多い学内学会誌2誌を有する複合領域で機関リポジトリ収録率が11.3%，学内学会誌23誌中13誌が集中する基礎医学分野で収録率が9.7%と高くなっていたが，いずれも学内学会誌を除けば収録率は全体の値（0.9%）にも至っていなかった．

4 考察と今後の課題

「はじめに」で述べたように日本の機関リポジトリは設置数・収録文献数ともに順調に増加してきている．しかし本研究の結果から，日本の学協会誌掲載論文の中で機関リポジ

トリに掲載されているものはごくわずかにとどまっていることが明らかになった．2000-2009年に出版された論文の中で機関リポジトリに収録されているものは1.8%で，いわゆる学内学会誌掲載論文を除けば収録率は0.9%と1%にも満たなかった．また，学内学会誌の影響を除けば，いずれの分野においても機関リポジトリへの論文収録率は低い値にとどまっていた．

機関リポジトリへの収録を許可する学協会誌掲載論文については機関リポジトリに収録される割合が有意に高く，分野別の著作権ポリシーと機関リポジトリ収録状況の間にも有意な相関関係があった．著作権ポリシーが機関リポジトリへの論文収録に影響すること自体は確かであり，SCPJのような取り組みは有効であると言えよう．調査対象論文の過半数は収録を許可しないか，方針を定めていない（不許可であると明言はしていない一方，許可もされていない）学協会誌に掲載された論文であった．特に方針を定めていない学協会への働きかけは引き続き重要と考えられる．

一方で，収録を許可するポリシーを有する学協会誌掲載論文であっても実際に収録される論文の割合は1.3%とごくわずかな値にとどまっていた．学協会が許可した場合でもほとんどの論文は機関リポジトリに収録されておらず，著作権ポリシーの明示は重要ではあるものの，それだけでは収録率は上がらないことも本研究の結果から明らかである．すでに膨大に存在する，学協会が機関リポジトリへの収録を許可した文献について，如何に著者の許諾を得て収録していくかは著作権ポリシーの確認以上に今後，重要な課題となると考えられる．

研究上の今後の課題として，本研究では機関リポジトリへの収録の有無のみ分析したが，今後はNII-ELSや他の電子ジャーナルサイト等での入手可能性についてもあわせて分析したいと考えている．これにより，日本の学協会誌掲載論文の電子的利用可能性と，その中での機関リポジトリの貢献状況を明らかにすることができるであろう．

参考文献

[1] Lynch, Clifford A: “Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age”, ARL Bimonthly Report, No.226, 2003, http://www.arl.org/resources/pubs/br/br226/br226ir.shtml （2012年4月3日参照）

[2] “Directory of Open Access Repositories”. http://www.opendoar.org/（2012 年4 月3 日参照）

[3] “機関リポジトリ一覧”. http://www.nii.ac.jp/irp/list/（2012年4月3日参照）

[4] “IRDB コンテンツ分析システム”. http://irdb.nii.ac.jp/（2012年4月3日参照）

[5] 佐藤翔；大向一輝；関戸麻衣；逸村裕：「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年度日本図書館情報学会春季研究集会，三重，2012-05-12，発表予定.

[6] 佐藤翔；逸村裕：「非学術的活動におけるオープンアクセス文献の活用：機関リポジトリ収録文献のリンク分析」，図書館情報メディア研究，Vol.9，No.1，pp.51-64，2011.

[7] 国立大学図書館協会 学術情報委員会 学術機関リポジトリワーキンググループ：「「学術機関リポジトリに関する調査」報告書」，52p，2010，http://wwwsoc.nii.ac.jp/anul/j/publications/reports/repository1.pdf（2012年4月3日参照）

[8] “SHERPA/RoMEO”. http://www.sherpa.ac.uk/romeo/（2012年4月3日参照）

[9] “学協会著作権ポリシーデータベース”. http://scpj.tulips.tsukuba.ac.jp/（2012年4月3日参照）

[10] Harnad, Stevan: “Estimating Japan's Annual Rate of Journal Article Self-Archiving”, Open Access Archivangelism, 2010, http://openaccess.eprints.org/index.php?/archives/763-Estimating-Japans-Annual-Rate-of-Journal-Article-Self-Archiving.html（2012年4月3日参照）