University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Faculty Publications, UNL Libraries | Libraries at University of Nebraska-Lincoln

1-1-2010

学術機関リポジトリでのオリジナルコンテンツの公開

Paul Royster
University of Nebraska-Lincoln, proyster@unl.edu

Follow this and additional works at: http://digitalcommons.unl.edu/libraryscience

Part of the Library and Information Science Commons

Royster, Paul, "学術機関リポジトリでのオリジナルコンテンツの公開" (2010). *Faculty Publications, UNL Libraries.* Paper 209.
http://digitalcommons.unl.edu/libraryscience/209

This Article is brought to you for free and open access by the Libraries at University of Nebraska-Lincoln at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, UNL Libraries by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

学術機関リポジトリでのオリジナルコンテンツの公開

Paul Royster

2008 年 1 月 30 日、オンラインで参照可能

学術機関リポジトリ（Institutional Repository、IR）は主に、学術雑誌、記念論文集、論文集など、他の場所で公開済みの学術的コンテンツを再公開する手段として開発され促進されてきた。本書では、他ではまだ公開していない資料の最初の公開場所に IR を使用する方法について考察し、学術的公開のための発展性のある「第一手段」として、また従来の商業出版社または大学出版局の代わりとなる既存の手段として IR を利用できる可能性を評価する。Serials Review 2008; 34:27-30.

©2007 Elsevier Inc. 禁無断転載。.

出版の自由は、出版社を所有する者にのみ保証されている － A. J. リープリング

まだ発展途中だが成長を続けているネブラスカ大学リンカーン校の学術機関リポジトリ（Institutional Repository、IR）[1]を管理して非常に驚いたことの 1 つは、通常の（紙の）出版物には適していない、あるいは論理的に適していないと考えられるいくつかの研究が圧倒的な人気を得ていることだ。このようなコンテンツはリポジトリ全体ではわずかな割合を占めるに過ぎないが、IRが本来アーカイブ、保存、アクセスの対象としている「正当な」つまり専門家の審査済みの論文に比べてはるかに利用回数が多い。どの月の「最もダウンロードされたドキュメント」でも、ダウンロード回数の多い論文トップ 10 の半数以上は、IRで最初に公開され、他では利用できないドキュメントだろう。

これは、IR の役割が、アーカイブ領域と利用しやすさを拡張するだけではないことを示唆している。実際、IR は、通常は除外されたり、後回しにされたり、通常の出版ルートによる予測できない変動、特異性、遅延、拘束、お役所仕事によって不利をこうむっている、広範な執筆者に発言機会を与えるオンライン出版元となるのに適している。IR の可能性を例証するため、大規模な参考資料、編集した文献シリーズ、学部プロジェクト、博士論文の 4 つの例を挙げて考察する。

Royster は米国ネブラスカ大学リンカーン校 Love Library の学術コミュニケーションのコーディネーター。
University of Nebraska-Lincoln, Lincoln, NE 68588, USA; e-mail: proyster@unl.edu

0098-7913/$-前付けを参照。©2007 Elsevier Inc. 禁無断転載。. doi:10.1016/j.serrev.2007.12.002

I.. 大規模な参考資料: 2005 年 7 月、大学のIRが稼働を始めた直後に、管理者（筆者）は、IRのコンテンツを補強するため多数のキャンパスを訪問したが、その中で国立博物館の寄生生物学研究室を訪れた。そこで、アルコール漬けの虫の瓶が何列もラック、ケース、引き出しに並んでいる中に、約 18 インチ（約 46cm）の高さのタイプ打ちされた文書を見つけた。尋ねると、研究所長が（引退した）カリフォルニア大学の同僚とその同僚の（亡くなった）妻とともに、15 年かけてまとめた無脊椎動物学の辞典であることがわかった。もちろん、これには悲しい話があった。西部の大手大学出版局に提出し、専門家の審査を受けて受領され、修正して、契約を結んだのだが、出版局に引き渡すまさにその週に、予告もなく契約がキャンセルされたのだった。出版局が（数年にわたる損失の後）動物学の出版をやめることを決めたためだった。他にもいくつかの出版社にアプローチしていたが、1 社はすでに競合する出版物を持っており、他社は原稿の長さと商業的に対象が限られることから二の足を踏んでいた。研究所長は同研究所のホームページに「オンラインで公開することを考えている」と書いていたが、そのときはまだ道のりの 4 分の 1 しか進んでいなかった。IR管理者は研究所長に機会を与えてほしいと懇願し、数日後にワープロ（WordPerfect）ファイル 98 本を受け取った。これらのファイルをMS Word形式に変換して連結し、原稿を整理して体裁を整え始めた。2 週間で 25 万クリックの作業を行った後、2005 年 9 月 6 日に『The Online Dictionary of Invertebrate Zoology』[2]がIRで電子公開された。ハーフレターサイズで約 1,000 ページの長さがあり、10 メガバイトに近い大きさになった。IRの自動PDF生成フィルタによってMS Word文書から生成されたPDFファイルだった。

ほとんど一夜で、この著作のダウンロードは IR 全体の 3 分の 1 を占めるようになった。執筆者たちはこの分野で（社会的にも電子的にも）広いつながりを持ち、かなり以前から切望されていた著作だったことがわかった。この著作は最初の 1 か月に約 1,200 回ダウンロードされ、それ以降も毎月約 1,000 回ダウンロードされ続けている。公開後の約 6 週間で、Adobe InDesign で作成し Adobe Acrobat で PDF 化した、より見栄えのいいバージョンが用意され、ファイルサイズも 3.4 メガバイトと小さくなった。これまでに（27 か月後までに）、約 30,000 回ダウンロードされている。

II. .編集した文献シリーズ: 筆者がIR管理者になる前、（英語学で）大学院の学位を取ってから「大学教員」としての地位を得て戻るまでに 20 年あったが、そのほとんどの期間は様々な学術出版物の制作、編集、管理に従事してきた。もちろん、最初にIRに格納したコンテンツのいくつかは筆者の古い学術論文で、1980 年代に公開した、アメリカ文学初期のあまり知られていない人々を取り上げた人名辞典項目が主だった。当然だが、あまり利用されることはなく、理由の 1 つは、その人々を知らしめた作品が長い間（何百年も）出版されずまったく利用できない状態にあったためだった。そこで、こうした作品を電子的に再出版すれば、その著者の経歴が興味を引くのではないかと思われた。この計画は残念ながら失敗したが、16～20 世紀の原作の新たな権威あるエディションを再公開する、『Electronic Texts in American Studies』シリーズ（現在 65 巻）のきっかけになった。このシリーズは驚くほど好評を得ている。[3]

初期のアメリカ文学はテキストがほとんど利用できず、長いこと研究に支障が出ていた。アンソロジーには、少数のソースから選んだ 2、3 のサンプル（通常は要約）が含まれている。当時、その多くは

出版すらされていなかった（たとえば、Winthrop の『A Modell of Christian Charity』、Bradford の『Of Plymouth Plantation』、Edward Taylor の『Preparatory Meditations』、Samuel Sewall の『Diary』など）。いくつかの作品は 19 世紀に散発的に発行された様々なシリーズで再公開されていたが、ほとんどは、Early English Books Online（EEBO）、Eighteenth-Century Collections Online（ECCO）、Evans Early American Imprints などのマイクロフィルム、マイクロフィッシュ、マイクロプリント、より新しい電子デジタル化された形式で（ページイメージとして）のみ利用できる。これらは判読できないほど不完全なことがあり、アクセスは大きな研究機関に限られていて、これらのシリーズは出版日付または発行した国によって制限を受けている。商業目的で出版されたコレクションで、研究機関のサービス加入者のみが利用できる。ページイメージなので、検索や抜粋ができず、しかもその基になる作品は公開されてから長い時間がたっている。通常の出版社は、このような著作権保護のできない、読者が限られた、（そして特に）興味を引くかどうかわからない作品には興味を示さない。

一般的に利用できない初期アメリカ文学のテキストの信頼できるエディションを作成して（IR上でオンラインで）公開する作業は、筆者の「学術的活動」になっている。これまでに（最初のテキストを公開してから約 24 か月で）、36,000 回を超えるダウンロードがあった。最も人気が高く、他に大差をつけているのは、James Andersonの 1723 年の作品『The Constitutions of the Free-Masons』のBenjamin Franklinによる 1734 年エディションで、これは米国で出版された最初のフリーメーソンに関する作品だ。[4]また、同様に人気が高いのは、Cotton Matherの『The Wonders of the Invisible World』（1693）[5]と、John Smithの『A Description of New-England』（1616）[6]、Increase Matherの『A Brief History of the Warr with the Indians in New-England』（1676）[7]、David Cusickの『Sketches of Ancient History of the Six Nations』（1828）だ。[8]このシリーズの目的の一つは、この時代の文学作品の主題と取り扱い方の多様さを示すことで、よく言われるような統一性や一致した総意という印象とは異なる。シリーズに収めた作品は、フリーメーソン、魔術、魔女、インディアン戦争と捕虜、絶滅したほ乳類（Stellerの海牛）、天文学、ソドミー、性的完璧主義、服装倒錯、奴隷制と廃止、迫害、虐殺などを話題にしている。

これらの著作は、研究者、教師、学生、その他の関心のある人が使用できるように、無償でオープンアクセスの電子形式で利用できる。教員には以前より格段に幅広いテキストを選択し利用する機会を与え、出版業界によって長く支えられてきた分野を広範に再評価する機会を提供している。

III. .学部プロジェクト: 1898 年の夏、ネブラスカ州オマハではThe Trans-Mississippi and International Expositionが開催され、開催地内の様々な会場で毎日音楽が演奏された。この期間のプログラムと新聞を調査したUNL音楽学部のある学部生は、博覧会が開催された 5 か月（6 月 1 日～10 月 31 日）に毎日演奏された音楽プログラムのほとんどを再構成することができた。この女子学生は、毎日誰がどこで何を演奏したかを示すインタラクティブなWebサイトを作成した。サイトを永続的にホストする場所がないため、彼女は教授を通してIRに問い合わせて、自身のHTMLファイルを図入りの検索可能なテキストベースのPDF文書に組み合わせ、学生研究用のシリーズとしてIRに投稿した。[9]「情報を平板に並べた」（106 ページの）PDF版をIRに格納するとともに、本来のインタラクティブなWebサイトを再構築したダウンロード可能な「zip圧縮した」HTML版も添付した。（Digital Commonsシステムを使用すると、

任意の数と形式の「関連」ファイルまたは補足ファイルを添付できる。）公開してから 15 か月で、この研究は 953 回ダウンロードされ、これは 1 日平均 2 回のペースである。この著作は別の形式では公開できなかったと思われるが、CDに格納して箱や棚におかれていれば世間一般から利用されることはなかっただろう。

IV. 博士論文: IRで博士論文を公開することは珍しくない。実際に、米国内のIRの最近の調査結果では、単一のドキュメントタイプとして最多数を占めている。[10]しかし、特に注目すべき大きな違いは、アクセスを制限するか会員に限定するのではなく、オープンにアクセスできることにある。我々のIRでの経験では、学位論文へのアクセスを制限した場合に対して、オープンにした場合にはダウンロード回数が（そして影響力が）約 60 倍増加するという差が出ている。

UNL の IR では、サービスの開始時に 9,500 件を超える電子学位論文（Electronic Theses and Dissertations、ETD）を導入した。これらは、マイクロフィルムのコレクションから ProQuest/UMI によってデジタル化されたものだ。大学はデジタル化した ETD にキャンパス内で永続的に電子アクセスするための料金を支払い、ProQuest/UMI はこれらを他のあらゆる形式、あらゆる場所で配布する権限を保有した。2007 年の夏に ProQuest が Digital Commons の管理を bepress（Berkeley Electronic Press）に返すまで、我々の利用レポートは、この会員加入に基づく（キャンパスでは無償、キャンパス外では有償の）論文のダウンロード数を示している。最近の ETD のダウンロードは、1 年に平均 7.20 コピーである（27 か月に 1,527 件の論文を 24,770 回ダウンロード）。カバーページが「ヒット」した論文を実際にダウンロードする割合は 16.0%である（383,473 件ヒットした中から 61,507 件をダウンロード、2005 年 4 月には 9,391 ETD、2007 年 6 月には 10,446 ETD）。言い換えれば、タイトルと要約を見た 6 人の内 5 人は、その ETD をダウンロードしていない（キャンパス外の利用者の料金は約 30 ドル）。（IR のオープンアクセス部では、ヒットに対するダウンロード率は 70～75%である。）

2006 年、大学では論文の電子保存を要求するようになった（これによって、図書館が製本と格納に費やすコストを軽減した）が、大学院ではすべての対象 PhD が適切な PDF ファイルを生成できるとは限らないことを懸念した。IR 管理者（筆者）は PDF 化について学生を補助したが、新しい PhD の約 10%にあたる約 200 人は、ファイルを PDF に変換するために助力を必要とした（特に、音楽、現代語、英語の年配の学生[35 歳以上]にその傾向が見られた）。ProQuest/UMI の論文保存契約は非独占的なものである（つまり、対象者は、別の場所に任意の形式で論文を公開または投稿する権利を保持する）。そのため、IR 管理者は、新しい PhD の PDF ファイル生成を支援するたびに、IR のオープンアクセス側で電子公開することの利点を候補者に説明した（ただちに公開できる、Google 検索ができる、永続的にリンクできる URL、全世界から 24 時間 365 日アクセスできるなど）。ほとんどの対象者は論文をオープンアクセス IR に投稿することに同意した（おそらく、何よりも感謝から同意してくれたのだろう）。こうして、66 件の ETD が完全にオープンアクセスで IR で公開された。

オープンアクセスにした論文のダウンロード数は飛躍的に伸びた。毎月のトップ「ダウンロード数」リストの半数を占め、その人気によって、教授が執筆した、審査済みで学術雑誌公開済みの論文はすっかり影が薄くなるほどだった。オープンアクセス論文の 1 年間の平均ダウンロード数は 424 だった

（利用できた 726 か月で 25,658 回のダウンロード）。これを非オープンアクセスの論文の 1 年間の平均 7.2 回と比較すると、5889%の増加で、約 60 倍にもなる。いささか驚いたのは、最も人気のある（19 か月で 3,759 回ダウンロードされた）論文はすべてスペイン語で書かれた、Julio César Sánchez Rondónの『Poética de lo Soez: Luis Rafael Sánchez: Identidad y Cultura en América Latina y en el Caribe』だったことだ。[11]（そして、文芸批評は死んだ、と言った）。
他のオープンアクセスの論文と比べても、この論文の人気は高い。リポジトリのオープンアクセス部では 1 か月の 1 件ごとのダウンロード数は平均して 3～5 回だが、このオープンアクセス論文の平均は 35 回を超えている。

より広く認知されるために

誰もいない森の中で木が倒れたら、音はするのだろうか － 哲学的な問い（作者不詳）

紙でも電子的形態でも、著作を単に発表するより公開する方が手間がかかるのは明らかだ。潜在的な利用者が著作を利用できることに気付かなければ、木が倒れても音はしない。
リソースの存在を潜在的な利用者に知らせるため、主にオンラインリンクという形態で、いくつかの努力が行われてきた。アクセス増加に最も効果を発揮しているのはThe Online Books Page[12]で、ここにリストされると、ほとんどすぐに 1 日のダウンロード数が 3 倍になるのが通常だ。Wikipedia[13]も役立つ人気サイトだが、利用者は外部リンクについてのガイドラインを調べてこれに従う必要がある。[14]有効性が実証されている別のリソースにMLA International Bibliography[15]とIntute[16]がある。専門分野を特化したWebページも有効で、我々のアメリカ文学研究はSociety of Early Americanistsリソースページの支援によって認識され宣伝されている。[17]ただし、各オンラインリソースサイトには様々なポリシーと更新方法がある。多くのサイトはホストするテキストを掲載するだけなのに、しばしばHTMLファイル形式を要求される。サイトによっては、最新情報だと宣伝しているのに 2001 年以降更新されていないことがある。
オリジナルの著作を目録化して（我々の図書館のテクニカルサービス部門が実施）、我々が所有するオンラインカタログ用のMARCレコードを作成し[18]、これを取り入れてWorldCatに組み込むと「オンラインリソース」として表示できる。[19]
検索エンジン向けの最適化（Search Engine Optimization、SEO）は、広く認識されるためのもう一つの重要な要素である。主題をタイトルに盛り込んで、要約のメタデータを最大限活用することが重要だ。『Online Dictionary of Invertebrate Zoology』の場合、個々の文字に追加のファイルを作成して、定義されている用語を（定義そのものは除いて）要約にロードした。こうすると、Google で「androsynhesmia」（後尾時期に集まってくる雄のグループ）を検索したとき、結果の最上部に ODIG が表示されるようになる。

社会的な地位

オンライン電子公開の悩みは、社会的地位が比較的低いことである。著作の真価について正しい評価を得るのも困難なことがある。我々の大学院の研究事務局でさえ、1 冊丸ごと電子的に公開された著作を、大学奨学金の年次調査および報告に含めることを拒否している。一般的に教授からは、オンライン公開はブログから半歩だけ進んだものと見なされているようだ。

これに対する答えは時間しかない。オンライン学術論文が増え、世代が変われば、最終的に流れが変わるだろう。Grisha Perelmanが、最近、ポアンカレ予測の解決でフィールズ賞（数学での最高の栄誉）を受賞したことは意味深い。その論文はオンラインのオープンアクセスリポジトリarXivでのみ公開されているためだ。[20]

まとめ

ここで紹介した（筆者の意見としての）活動が、電子公開またはデジタル学術研究の最先端を示していると思われてはいけない。実際のところ、筆者は、ウィンカーを点滅させながら「情報ハイウェイ」を時速 35 マイル（56 キロ）で運転する老人のような気がしている。公開している著作は「平板な」PDF ファイルで、所々ハイパーリンクを設定しているのがせいぜいで、インタラクティブでもハイパーテキストでもない。助成金を受けることも、新しいデジタルツールを開発することも、従来型または電子的な学術論文に関する技術や手段を分析し、覆し、改革することもしていない。

革命をもたらす可能性があるのは、これによって、すべての学部と新しい生産関係を切実に求める労働者階級が、公開資料（つまり学術的な生産手段へのアクセス）を広く利用できるようになることだ。他の専門家が、メンバーに労働生産物（著作権）の無償提供を求めるのは、仕事を維持するためか。しかし、彼らは印税を受け取っているではないか？　いや、一部はそうかもしれないが、多くはそうではない。時折広く採用される入門教科書を除けば、形ばかりの収入以上のものを受け取ることはほとんどない。

出版社はこれまで、専門家の審査を実施して重要な「門番」としての役割を果たしていると自己弁護してきたが、それはどうだろう。審査は出版社ではなく、教員が行うものだ。門番と言うが、閉め出されているのは（または、長い列を作って入場を待たされているのは）、従来の出版社にとって必要な収入が得られない著作だ。1 冊の研究論文を作成する直接費は（ほとんどはフリーランサーや外部業者が行うため）5、000 ドルに収まるが、出版社の 20,000 ドル以上の間接費をまかなわなくてはならず、収支を合わせるだけで 50,000 ドル以上の売り上げが必要になる。このような環境では、出版社には多くの分野の出版ができるほどの予算はなく、学術的価値より「出版社の判断」が優先されて、あらゆるものが、学部ではなく出版社の自己防衛に関連付けられてしまう。出版とは、学術出版といえども本を売るビジネスで、研究は市場動向によって定義される売却可能なユニットに組み込まれてしまう。

しかし、著者が自分自身で公開できるとしたらどうだろう。学部のすべてのメンバーが自分自身の出版局を持ち、自分が適当だと思う方法で自由に利用できたらどうだろう。その結果、伝統主義者が予想するような無秩序状態になるだろうか。たとえば、出版社の John Wiley & Sons がどれを出版するか

決定しなかったと言って、大学構内で暴動が起こることがあっただろうか。図書館が（無償の）リソースを収集して保存してはどうだろう。終身在籍権（テニュア）委員会は再任すべき人をどのように知るのか。問題はあまりにも深く幅広いため、フットボールプログラムで有名な中西部の大学の IR 管理者では答えることができない。それでもなお、筆者は、実現しはじめたばかりの学術論文のオープンアクセスのオンライン公開には可能性があり、2015 年または 2020 年の学術界は今とはずいぶん異なっているだろうと感じている。

注記

1. http://digitalcommons.unl.edu/
2. http://digitalcommons.unl.edu/onlinedictinvertzoology/
3. http://digitalcommons.unl.edu/etas/
4. http://digitalcommons.unl.edu/libraryscience/25/
5. http://digitalcommons.unl.edu/etas/19/
6. http://digitalcommons.unl.edu/etas/4/
7. http://digitalcommons.unl.edu/libraryscience/31/
8. http://digitalcommons.unl.edu/libraryscience/24/
9. http://digitalcommons.unl.edu/musicstudent/4/
10. C. S.McDowell、"Evaluating Institutional Repository Deployment in American Academe Since Early 2005"、D-Lib Magazine 13:9-10（2007 年 9 月/10 月）、doi:10.1045/september2007-mcdowell、http://www.dlib.org/dlib/september07/mcdowell/09mcdowell.html
11. http://digitalcommons.unl.edu/modlangdiss/1/
12. http://onlinebooks.library.upenn.edu/
13. http://en.wikipedia.org
14. http://en.wikipedia.org/wiki/Wikipedia%3AExternal_links
15. http://www.mla.org/bibliography
16. http://www.intute.ac.uk/
17. http://www.societyofearlyamericanists.org/links.html
18. http://iris.unl.edu/
19. http://www.worldcat.org/
20. http://arXiv.org。論文はhttp://arxiv.org/abs/math/0307245、http://arxiv.org/abs/math/0303109、http://arxiv.org/abs/math/0211159で参照できる。