

◆研修会特集◆ 1. 講演

Medical*Online 文献/書籍の電子化作業から配信に至るまで

吉岡 誠

抄録：Medical*Online は国内医学関連ジャーナルのデータベースとして2000年にリリースし、2016年8月時点で約1,250ジャーナルを配信するサイトに成長した。2015年に、電子書籍サービス（イーブックス）をリリースし、現在約2,800冊の電子書籍を配信している。今回のテーマは Medical*Online の製品に関する内容ではなく、ジャーナル（文献）配信、電子書籍配信に至るまでの電子化作業に関する記事である。電子化作業のベースとなるスキャン業務では各種課題点を解決すべくユニークな機器を導入したので、その概要を併せて紹介する。また、電子化作業に直接的な関連性はないが、Medical*Online-FAX サービスの現状についても報告する。

キーワード：Medical*Online、スキャン、SCAN RBOT 2.0MDS、テキスト化、FAX サービス

I. はじめに

Medical*Online は国内医学関連ジャーナルのデータベースとして2000年にサービスを開始し、現在約16年が経過した。その間、IT業界—特に学術分野—においては情報通信基盤は日々、目まぐるしい変化と進化を遂げてきた。このような状況下において、Medical*Online は利用者である医療従事者へ、利便性の高いサービスを提供することを目的としており、そのバックボーンとなる電子化作業は当社にとって非常に重要な部分である。

II. 配信ジャーナル数と採用実績の推移

図1に見られるとおり、配信ジャーナル数と採用実績数は毎年増加傾向にあり、配信ジャーナルでは年平均87誌が増加している。また採用実績においては、医学部大学の採用率が90%、500床以上の病院の採用率が73%である（当社調査）。

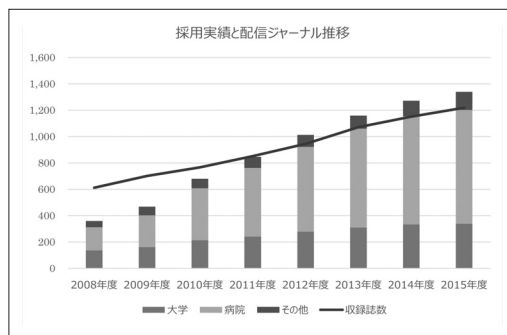


図1

YOSHIOKA Makoto
株式会社メテオ 営業部
TEL : 03-5577-5877

Ⅲ. 電子化業務について

1. 雑誌電子化業務の現状

Medical*Online は、版元である学会、研究会、出版社と電子配信許諾契約を締結し、サービスを提供している。契約締結時点で、バックナンバーは最長で過去10年分が送付され、以後新刊発行の都度、当社札幌データセンターまで送付されている。当社へ送付される雑誌の入荷形態は基本的に冊子体もしくは電子データの2種類となっており、その割合は冊子体が約75%、電子データが約25%である。そのため電子化作業の大半は冊子体からの電子化作業となる。冊子体からの電子化作業はスキャンングから始まり、各工程を経てPDFデータと書誌情報のテキストデータが完成する。

2. 雑誌電子化作業の課題点

版元との契約時にバックナンバー10年分を電子化しているが、版元側のバックナンバー保管状況は、まちまちである。電子化する際の課題点としては、在庫が少ない、1冊しかない、製本されている、紙が劣化している等々である。当社で採用している従来の高速スキャナーでは、背表紙を裁断し、コピー機と同様の工程でスキャンングが行われるため、この課題点に対応できない状況であった。このような背景を受けて、以下の課題点を解決できるスキャンング機器を探した。

- (1) 電子化作業完了後版元へ返却できるようにする。
- (2) 背表紙を裁断しないでスキャンングを可能とする。
- (3) 製本、合本雑誌のような厚い書籍（雑誌）でも画像にゆがみや影を出現させない。

- (4) ページ送り（めくり）は、手動ではなく自動とし、できる限りソフトタッチ。

3. スキャンロボット（SCAN RBOT2.0MDS）の導入

これらを解決できる機器として、各種検証した結果、SCAN RBOT 2.0MDS（写真1）を導入することになった。導入のポイントを以下に列挙する。

- (1) 背表紙の裁断不要⇒冊子を専用台にセットし、スキャンヘッドが上下してスキャンング
- (2) ページ送り（めくり）⇒空気ページを浮かせ、バキュームで吸い上げながらスキャンング、空気ページをめくることが可能
- (3) 本の厚さは最大14cm まで可能
- (4) スキャンング後の画像調整できる機能あり

これにより、今までスキャンングが困難であった雑誌に対し、電子化を可能としたことで、Medical*Online のアーカイブは、質、量ともに一層拡充することとなった。

ちなみに、この SCAN RBOT 2.0MDS は、2004年ウィーン工科大学でパイロットプロジェクトが端緒となり、2005年に特許出願し、2006年に TREVENTUS 社を設立し製造販売に至ったという経緯である。また2007年 CeBIT（セビット 国際情報通信技術見本市）ヨーロッパ ICT 賞にて「大賞」を受賞。現在オーストリア・インスブルック大学図



写真1

書館、ドイツ・バイエルン州立図書館など、世界中に200台以上納入されているとの事。日本での導入実績は、当社の1台のみと聞いている。

4. 電子化業務フローの変更

スキヤニング業務において、SCAN RBOT 2.0MDS の導入前は高速スキャンによる電子化スキヤニング業務の1ラインであったが、導入後は高速スキャンとSCAN RBOT 2.0 MDS の2ラインとなった。これにより版元から当社へ雑誌が到着後、各ラインへの振り分け作業を行い、電子化業務を実施することになった。表1に2015年4月から2016年3月までの入荷雑誌に対するデータ形式（冊子体/電子データ）と、冊子体に対しスキャンラインごとの雑誌数とその割合を集計したので以下に記載する（表1）。

年月	入荷雑誌数	冊子体入荷				電子データ入荷	
		高速スキャン		SCAN RBOT		PDF 入荷	
		雑誌数	割合	雑誌数	割合	雑誌数	割合
2015年4月	437	240	54.9%	76	17.4%	121	27.7%
2015年5月	359	219	61.0%	58	16.2%	82	22.8%
2015年6月	459	215	46.8%	111	24.2%	133	29.0%
2015年7月	541	250	46.2%	164	30.3%	127	23.5%
2015年8月	557	228	40.9%	167	30.0%	162	29.1%
2015年9月	452	241	53.3%	112	24.8%	99	21.9%
2015年10月	594	319	53.7%	96	16.2%	179	30.1%
2015年11月	558	302	54.1%	174	31.2%	82	14.7%
2015年12月	842	621	73.8%	44	5.2%	177	21.0%
2016年1月	346	237	68.5%	41	11.8%	68	19.7%
2016年2月	369	219	59.3%	25	6.8%	125	33.9%
2016年3月	608	286	47.0%	189	31.1%	133	21.9%
平均値	510	281	55.0%	105	20.4%	124	24.6%

表 1

表1の通り、入荷雑誌の約75%が冊子体で約25%が電子データ（PDF）となっている。電子データで入荷された文献は、利用者にとっても利便性が高いため、版元には電子データでの入荷依頼を定期的に行っている。

IV. SCAN RBOT 2.0MDS の操作について

この機器の作業フローの非常にユニークな動作に関し、写真を交え簡単に紹介する。

1. 本をクレイドル（筐体中央部）にセットする。



2. スキャンヘッドが下降する。



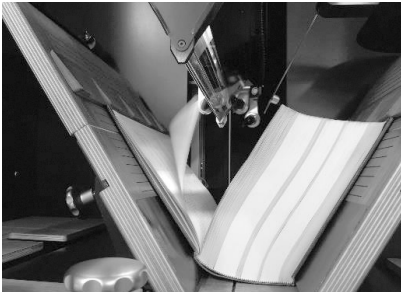
3. 左右2ページを吸い上げる。



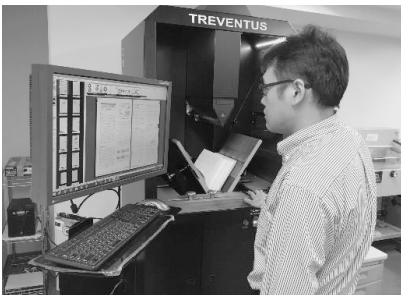
4. 上昇しながらスキヤニング



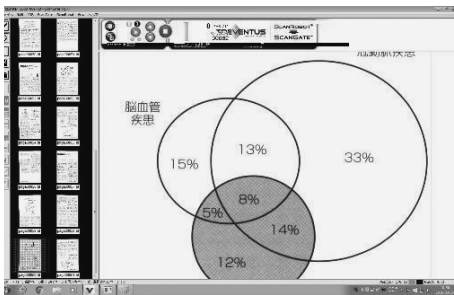
5. ヘッドがトップ位置に戻りページめくり



6. 2-5の繰り返し



7. 取り込み画像の確認



8. 本の取り出し



9. SCAN RBOT 2.0MDS の電子化業務に携わる作業者の総括

SCAN RBOT 2.0MDS は全自動機器ではなく、スペック通りの生産性を生み出すことが非常に難しい機器であり、作業者にとってノウハウと熟練を要する機器である。このSCAN RBOT 2.0MDS は通常のフラット型スキャンに比べ、完成品は圧倒的に高いクオリティを確保できるが、そのコストは、通常の電子化業務よりはるかに高額になる。ただこの投資は、貴重な知的財産である書物を傷つけることなく電子化し、保存、配信するという社会貢献的な視点からも意義のある業務であると認識しており、今後もクオリティの向上に向け日々努力し続けたいと考えている。

V. 書誌情報のテキスト化とリンク設定業務について

Medical*Online の電子化業務は、PDF 作成だけではない。書誌情報のテキスト化業務や、引用文献のリンク (Medical*Online PDF へのリンクや PubMed へのリンク) 設定業務、各リンクリゾルバ/ディスカバリーサービスとの連携業務等を行っている。このうちテキスト化業務とリンク設定業務について基本的な部分を紹介する。

1. テキスト化業務について

Medical*Online の「文献」で提供する情報は、大きく分類すると4つの項目に分かれている。(1) 検索結果に表示する文献の書誌情報、(2) 文献のアブストラクト、(3) 文献PDF、(4) 文献FAXである。このうち書誌情報とアブストラクトがテキスト化の対象になる。書誌情報のテキスト化する項目は基本的に文献のタイトル、著者名、著者所属、雑

誌名、巻数、号数、ページ、発行年である。アブストラクトは、文献 PDF から要旨部分を抜粋しテキスト化を行う。アブストラクトは日本語のみの文献と、英語のみの文献、日本語と英語の併記の文献があるが、基本的に原文に併せてテキスト化を行っている。

2. 引用文献のテキスト化業務とリンク設定業務

テキスト化業務の一部に引用文献のテキスト化がある。これは Medical*Online「電子書籍」サービス (EBooks) の引用文献情報を抜粋し、テキスト化業務およびその引用文献へのリンク設定を行っている。この引用文献のテキスト化業務とリンク設定業務とその実装は、電子書籍サービス (EBooks) だけではなく、Medical*Online「文献」へも実施する予定である。

VI. Medical*Online-FAX サービス

1. Medical*Online-FAX サービスの対象雑誌について

Medical*Online で配信している雑誌は、大きく分けて学会誌と商業誌に分かれており、商業誌については、そのほとんどが新刊発行後、即時の全文 PDF 配信ができない配信制限期間が設けられている。この配信制限は、雑誌によってその期間や、制限内容が異なっている。基本的に書誌情報とアブストラクトには配信制限は設けられていない。全文 PDF に対して配信制限が設けられている。具体的な配信制限の内容としては、(1) 一定期間の全文 PDF 配信が不可、(2) 一定期間の全文 FAX 送信の 2 通りとなっている。FAX 送信可能な雑誌と PDF 配信不可能な雑誌の割合は86 (FAX 送信可) : 14 (PDF 配信不可)

となり、ほとんどが FAX 送信可能となっている。

2. 直近 1 年間分の FAX 送信状況 (2015/07 ~2016/06)

年月	送信数
2016年6月	2,266
2016年5月	1,963
2016年4月	2,020
2016年3月	1,987
2016年2月	1,951
2016年1月	1,816
2015年12月	1,762
2015年11月	2,221
2015年10月	2,247
2015年9月	1,958
2015年8月	1,817
2015年7月	2,288

表 2

直近 1 年間の送信数の平均は2,025送信/月となっている。

3. Medical*Online-FAX サービスの今後

FAX サービスは全文 PDF 配信に比べ、その利便性とクオリティは低い媒体であり、当社もこの配信制限期間をできる限り短縮してもらえよう版元との継続的な交渉が必要であると認識している。

VII. おわりに

Medical*Online は文献のデータベースという性質上、配信雑誌数の増大化による利便性の向上を至上命題として取り組んでいる。しかしながら、配信雑誌数の増大化は、制作コストという側面では増加傾向となってしまう。当社は経営努力の一環として、作業効率の向上と製品クオリティの向上に多くのリソースを割いている。日々進化している学術情報の流通環境に対応できるよう鋭意努力していきたい。