

『フフ・トグ(青旗)』のデジタルブック作成をめぐる諸問題

相原 佳之

1. はじめに

大阪大学附属図書館の貴重書コレクション「石濱文庫」に所蔵されるモンゴル語新聞『フフ・トグ(青旗)』のデータベース構築と公開を目指す本プロジェクトの中で、本稿に与えられた課題は以下の二つである。

一つ目は、筆者が関わってきた公益財団法人東洋文庫現代中国研究資料室(人間文化研究機構現代中国地域研究東洋文庫拠点)においてデジタルブックを用いた資料公開を行ってきた経験をもとに、利用可能な知識と問題点を提供することである。また二つ目は、提供された『フフ・トグ(青旗)』の紙面画像をもとに、市販ソフトを用いてデジタルブックを試作し、その中から発見された問題点を挙げ、データベース全体の中においてデジタルブックをどのように活用できるかについて、案を提示することである。

以下、これらの点につき、具体的なブック作成手順等も含めて述べていきたい。

2. 東洋文庫におけるデジタルブックを利用した資料公開

東洋文庫現代中国研究資料室では、これまでに下記のデータベースを公開してきた。(公開 URL : <http://www.tbcas.jp/ja/lib/>)¹

近代中国関係資料デジタルライブラリー (555 タイトル, 47,469 画像)

『柏原英一写真帳』(7冊, 写真数約 800 枚)

『亜東印画輯』データベース (16冊, 写真約 2600 枚)

『亜細亜大観』データベース (16冊, 写真約 2600 枚)

¹ 東洋文庫現代中国研究資料室における資料デジタル化の詳細については、下記の文献を参照。相原佳之・徳原靖浩「東洋文庫における地域研究資源の情報化・共有化の取り組み」, NIHU 研究資源共有化事業委員会編『PNC2013 NIHU 企画セッション報告集』, 85-101 頁, 2014 年 3 月, 相原佳之「東洋文庫所蔵の近代中国資料のデジタル化事業について」, 堤一昭・田中仁編『戦前期モンゴル語新聞『フフ・トグ(青旗)』のデジタル化と公開の可能性: 東洋文庫政治史資料研究班・研究セミナーの記録』(OUFC Booklet 第 7 巻), 65-74 頁, 2015 年 3 月。

このうち、写真帳を公開している、[http://www.tbias.jp/digitalimages/flipperu](#) は、市販ソフト Flipper U を用いたデジタルブックを作成し、ブックデータを web サーバーにアップロードする形で公開した。これらの資料はいずれも、写真ごとに標題(タイトル)または解説文(キャプション)が附属している。デジタルブック作成時には、この標題および解説文をいかにしてデジタルブックに反映し、利用者にわかりやすい形で提供できるかが課題となった。

においては、写真解説文を透明文字テキストとして埋め込んだ PDF ファイルからデジタルブックを作成する手順で、[http://www.tbias.jp/digitalimages/flipperu](#) においては、JPEG 画像にてデジタルブックを作成し、標題を目次用 CSV で、キャプションを検索用テキストの CSV でそれぞれ追加する手順でデジタルブックを作成し、1冊ごとではあるが、標題や解説文から該当の写真が載るページが検索できるようにした²。

3. 『フフ・トグ(青旗)』デジタルブックの試作

上記の経験を踏まえ、『フフ・トグ(青旗)』のデジタルブックを試作した。試作の素材として利用したのは、第8号～第11号(康徳8年5月10日～5月31日)の画像、計32画像(1号あたり8頁)である(画像形式は TIFF)。

本報告集の資料に見られるごとく、『フフ・トグ(青旗)』の紙面には1ページに複数の記事が載り、記事に大見出し、中見出し、小見出しが存在する。また本報告集の田中仁氏の報告に見られるように、データベースとしては記事の細目を縦書きモンゴル文字やモンゴル語のローマ字転写、日本語訳など複数文字表記・複数言語で記述し検索することが目指されている。

求める検索をすべて満たすためには、それに適した新たなデータベースを構築する必要があるのはもちろんであるが、デジタルブックの機能を利用して、一定の検索は可能である。またデータベースとデジタルブックを併用して閲覧環境を整えることも考えられる。以下、試作した幾つかのバージョンのデジタルブックをもとに、活用の可能性を考えてみたい。

作成したデジタルブックの一覧は、【別表】に載せた。以下、適宜別表1行目の番号を用いる。なお、試作に用いたソフトとそのバージョンは、Adobe Acrobat XI Pro、FLIPPER U 4.2.4.0 バリュースタック、FlippingBook Publisher Professional 2.6.40、である。OS環境はすべて Windows7 Home Premium である。

² 同様にデジタルブックを利用した資料公開の事例として、東洋文庫イスラーム地域研究資料室が公開している Ali Haydar によるオスマン民法典(Mecelle)の注釈書 *Dürerü'l-hukkam* がある(公開 URL <http://tbias.jp/digitalimages/durerulhukkam>)

1) 事前加工

デジタルブック作成ソフトは、TIFF 画像に対応していないため、JPEG への変換 [0-1] を行った。また PDF への加工 [1-1 ~ 1-3] を行った。

PDF のうち、[1-1]、[1-2] は画像をもとに 1 枚ごと、1 冊 (8 頁) ごとに PDF を作成したもので、検索用の文字列は追加していない。[1-3] では [1-2] の PDF をもとに、記事の見出しおよび本文を PDF の該当箇所に埋め込んだ上で、テキストが表に出ないように画像の背後に隠す処理を行った。記事の一つ一つに対してテキストを埋め込む必要があるため処理に手間がかかるものの、PDF 閲覧ソフトの検索機能を用いることで、目当ての記事にたどり着く手段の一つとすることができる。

2) FLIPPER U によるデジタルブック作成

次にロゴスウェア社のデジタルブック作成ソフト FLIPPER U を用いた試作を行った。まず [2-1] は [0-1] の JPEG 画像を用い、検索用文字列を追加せずに加工をしたものである。画質を「大」で試作したものはサイズが 81.4MB となる。同じ画像を用いた PDF ファイルは 3.4MB であるため、サイズは非常に大きくなる。容量が大きいため web サーバー等にアップロードする際には時間がかかるが、PDF と異なり見開き頁ごとの読み込みになるため、閲覧はスムーズとなる。【別表】に見られるように、目次や検索用の文字列を追加したデジタルブックも、サイズは大きく異ならない。また左開き、右開きどちらのデジタルブックも作成できるため、『フフ・トグ (青旗)』のような左はじまり縦書きの言語でもページそのままの順での閲覧環境を構築できる。

続いて、目次を追加したブック [2-2] を作成した。メタデータは、タブ区切りテキストで、ページ番号に記事見出しを付加したものである。記事の階層 (大見出し、中見出し、小見出し等) は、下例のようにタブの数を増やすことでブック上に再現可能である。目次を追加することで、閲覧者はデジタルブックの左側にツリー形式で表示される目次を確認しながら、目当ての記事のあるページを探すことができる。なお、目次は同じページ番号に対して複数追加できるため、『フフ・トグ (青旗)』のごとく 1 頁に複数の記事を持つ新聞についても応用が利く。ただ、目次に追加した文字列は、このままでは検索できない。

ページ番号 [Tab] 大見出し

ページ番号 [Tab] [Tab] 中見出し

ページ番号 [Tab] [Tab] [Tab] 小見出し

さらに記事見出しの検索を可能にするため、ソフトの検索ワード追加機能を使用し、検索用テキストを追加したデジタルブック [2-3] を作成した。メタデータは目次と同じくタブ区切りテキストであるが、1 頁に対して検索文字列を 1 行で追加する。そのため、1

頁に複数の記事がある場合は、次のような形式で、記事の見出しをタブ以外の任意の文字で区切って追加する。この機能で追加した文字は、デジタルブックの機能を使って検索することで、検索結果一覧が表示され、結果をクリックすることで、目当ての記事の載るページにたどり着くことができる。

- 1 [Tab] 1 頁記事 1 見出し / 1 頁記事 2 見出し / 1 頁記事 3 見出し...
- 2 [Tab] 2 頁記事 1 見出し / 2 頁記事 2 見出し / 2 頁記事 3 見出し...
- 3 [Tab] 3 頁記事 1 見出し / 3 頁記事 2 見出し / 3 頁記事 3 見出し...

同じ機能を用い、メタデータとして記事見出しの代わりに記事のキーワードや本文を入力したタブ区切りテキストを用いれば、これらも検索対象になる。つまり、デジタルブックの機能を利用して、キーワードや本文の検索機能も簡易的に備えることができる。なおメタデータの文字コードは UTF-8 であり、多言語検索を目指す場合に、この範囲内の文字が使用可能である。ただ、この検索ワード追加方法では、目当ての記事の載るページが検索できるだけであり、その記事がページ内のどこに載るかは、検索結果に示されない。

上記のようにタブ区切りテキストを用いる検索ワード追加機能のほか、PDF を用いた検索ワードの追加も可能である。前述した [1-3], つまり文字列を追加した PDF を用いることで可能になる。PDF で追加された文字列は、検索された場合に文字列の背景色が反転するハイライト表示がされる。それゆえ、記事のある場所に文字列を適切に埋め込んだ PDF を利用すれば、検索された記事がページ内のどこにあるかがハイライト表示により明確になる。検索機能としてはより利便性が高まることになる。

3) Flippingbook によるデジタルブック作成

今回は、主として上記 FLIPPER U による試作を中心に検討したが、同種のソフト Flippingbook も用いたデジタルブックも試作した ([3-1] [3-2] [3-3])。FLIPPER U よりできあがりのサイズは小さくなる。文字列追加済みの PDF ファイル ([1-3]) を用いた場合にハイライト表示が実現できる点は、FLIPPER U の場合と同様である。ただ、[2-2] のようなタブ区切りテキストを用いた検索文字列追加はできない。また、目次の追加も一つ一つ手入力で行う必要がある。一方、検索結果の一覧性は FLIPPER U よりも優るという印象を持った。

4 . デジタルブックとデータベースの連携について

FLIPPER U でも Flippingbook でも、出来上がったデジタルブックは一冊が一つのフォルダに格納されたファイルの一セットとして出力できる。そして FTP クライアントソフト

等を用いてブックデータをフォルダごと web サーバーにアップロードすることで閲覧可能になる。

リンク先として、フォルダ(ディレクトリ)内の index.html ファイルを指定すると、ブックコンテンツが起動される。また各ページごとに固有の URL が生成され、その URL をリンク先として指定することで、デジタルブック内の特定のページを直接閲覧することができる。ページ固有 URL は、データベースと連携を行う際に活用できる。たとえば、データベースの記事のメタデータに、下記のように画像リンク URL としてページ固有 URL を載せることで、データベースの検索結果から容易にデジタルブックの該当ページにピンポイントで飛ぶことが可能になる。ページに飛んだ後には、前後のページにも移動できる。ただし、ページ内のハイライト表示を同時に行うことは難しい。

データの詳細な検索はデータベースにおいて行い、スムーズな閲覧環境はデジタルブックにより提供するという形で、特性を生かした併用が可能であろう。

記事番号	号	面	表題モ	表題日	表題英	作者1	作者2	年月日1	年月日2	キーワード	分類	...	画像リンク URL
1	8	1										...	http://www.*****/#page=1
2	8	1										...	http://www.*****/#page=1
3	8	2										...	http://www.*****/#page=2
4	8	3										...	http://www.*****/#page=3
...

5 . 終わりに：今後に向けて

以上、『フフ・トグ(青旗)』のデジタルブックを試作した結果を述べてきた。市販のデジタルブックソフトを用いて、記事見出しや本文等の検索ができ、データベースと連携可能なブックが作成できることが示された。

今後実際の活用の際には、本稿で紹介した以外のデジタルブック作成ソフトの機能や、作成したデジタルブックの見映え等の問題も同時に検討し、より容易な閲覧環境の整備に向けた取り組みが必要になるであろう。

また、『フフ・トグ(青旗)』の1冊ごとにブックを作成するのか、発行年月ごとに月単位でブックを作成するのか、検索用テキストとしてどのような情報を載せるか、記事のハイライトを実現するために比較的手順のかかる PDF へのテキスト埋め込みを行うか否か、デジタルブックの作成を業者委託するか否かなどの課題については、完成期日や人員など

の条件に基づいて作業量を見積もり、実現可能な最適なデジタルブックを作成することが
のぞまれる。

本稿での初歩的な試作結果が、『フフ・トグ(青旗)』のデータベース構築・公開の過程
において快適な閲覧環境を実現する際に活用され、『フフ・トグ(青旗)』がより多くの研
究者・学生等に利用されることに少しでも資することができれば幸いである。

別表：作成したデータの一覧表

	0-1	1-1	1-2	1-3	2-1	2-2	2-3	2-4	3-1	3-2	3-3
種類	JPEG	PDF1枚 ごと	PDF一冊 ごと	PDF文字埋め込 み	FLIPPER検索なし	FLIPPER目次 つき	FLIPPER目次検 索つき	FLIPPER PDF文 字埋め込み	FlippingBook カタログ グ	FlippingBook 文書	FlippingBook PDFより
素材	tiff	tiff	tiff	1-2のPDF (あり)	0-1のjpeg	0-1のjpeg	0-1のjpeg	1-3のPDF	0-1のjpeg	0-1のjpeg	1-3のPDF
目次	なし	なし	なし	(あり)	なし	あり	あり	あり	なし	なし	あり
検索用	なし	なし	なし	(あり)	なし	なし	あり	あり	なし	なし	あり
OS	Windows 7	Windows 7	Windows7	Windows7	Windows7	Windows7	Windows7	Windows7	Windows7	Windows7	Windows7
ソフト	フリー ソフト	Adobe Acrobat	Adobe Acrobat	Adobe Acrobat	FLIPPER U	FLIPPER U	FLIPPER U	FLIPPER U	FlippingBook	FlippingBook	FlippingBook
作業	拡張子の 変更	ファイル からPDFを 作成	ファイル をPDFに 結合	目次または検索 用テキストを記 事箇所に埋め込 み(手作業)	デジタルブック加工	デジタルブック 加工、目次は タブ区切りテ キストフア イルでインポ ート	デジタルブック加 工、目次と検索用 テキストはタブ区 切りテキストフア イルでインポ ート	デジタルブック加 工、テキストは PDFの埋め込みテ キストから自動抽 出、目次はタブ区 切りテキストフア イルでインポ ート	デジタル ブック加 工	デジタル ブック加 工	デジタルブック加 工、テキストは PDFの埋め込みテ キストから自動抽 出
設定									カタログ	文書	カタログ
見開き / 片面	片面	片面	片面	片面	見開き	見開き	見開き	見開き	見開き	片面	見開き
画質					大 / 標準 / 小	大	大	大	最高品質 / ライト	最高品質	最高品質
備考				画像のみ				検索結果文字のハ イライトが可能。			検索結果文字のハ イライトが可能。
サイズ		0.5MB	3.4MB		81.4MB / 31.8MB / 15.4MB	81.5MB	81.6MB	83.0MB	60.5MB / 29.6MB	60.4MB	44.2MB