

# トランスクリプトマッピングのための書式とその活用

服部 雄輝<sup>†1</sup> 寺沢 憲吾<sup>†1</sup>

近年、文書の書面画像をデータとして扱う機会が増加している。特にデジタルアーカイブにおいて扱われる史料は、その内容のみならず、文書そのものの画像データも貴重な資料とされることから、トランスクリプトと文字画像の各文字位置を対応付ける、トランスクリプトマッピングの研究が盛んに行われている。しかし、現状の研究では、マッピングされたデータが活用される場面は、研究目的に限られるため、より広い分野でマッピングデータを利用するのは困難である。

そこで本研究では、トランスクリプトマッピングのための共通データ書式を提案する。様々な環境のアプリケーションにおけるデータの相互利用を容易にすることで、有益な文献のマッピングデータをより広く活用できるようにすることを目指す。

## Data format for transcript mapping and its application

YUKI HATTORI<sup>†1</sup> and KENGO TERASAWA<sup>†1</sup>

Recently, we often have been treating document images as digital data. In the historical documents, the data have values not only for its contents but also for letters themselves, thus researches for transcript mapping are becoming active. However, the mapping results are tend to be used for the specific research purpose. It is difficult to use the result of the transcript mapping for diverse fields besides the research even if they are useful documents. In this research, we will propose a common data format with XML for transcript mapping. It will make it possible to use the data for mutual collaboration, in various environments.

### 1. はじめに

本研究では、トランスクリプトマッピングを行うための共通データ書式を提案し、データ

<sup>†1</sup> 公立はこだて未来大学 システム情報科学部 情報アーキテクチャ学科  
Department of Media Architecture, School of Systems Information Science, Future University  
Hakodate



図1 トランスクリプトマッピング  
Fig. 1 Transcript mapping.

の相互利用を容易にすることで、有益な文献のマッピングデータをより広い分野で活用できるようにすることを目的とする。

トランスクリプトとは、文書の内容を活字のテキストに書き起こしたデータのことを指す。トランスクリプトマッピングとは、一つの文書に対するトランスクリプト(テキストデータ)と、同文書を画像として取り込んだ画像データ上にある、同じ文字の位置情報を対応付けることで、両データ間の同一文字の位置を関連付ける手法である。図1の例で言うならば、「四」という字は、トランスクリプト上では一文字目にあり、画像データ上では位置座標点からなる矩形範囲で示されている。この二つの文字位置情報同士を一対一で対応付けることが、トランスクリプトマッピングである。つまり、文書画像中の「四」という座標範囲内が、トランスクリプト上における一文字目の「四」に対応している、ということが分かるようになり、逆もまた然りである。この対応付けを文書中の各文字に適用することで、文書画像中の文字の位置からトランスクリプト上のテキストを参照する、あるいはトランスクリプトのテキストから画像データ上の該当する文字位置を参照することができるようになる。

#### 1.1 研究背景

近年において、文書をデジタル画像データとして取り込んで扱う機会は増加の一途をたどっている。特に、スマートフォンやタブレット端末が牽引役となった電子書籍の普及の影響が大きい。ごく最近では、文書の紙面画像をそのまま取り込み、携帯端末で閲覧するという使い方も一般的になるなど、文書画像データの活用が広く浸透しつつあると言える。無論、研究分野でもその有用性は高い。例として、デジタルアーカイブの動きがある。これは様々な文化財をデジタル化して、劣化のない形で保存・公開を行う取り組みである。文書に関しては、歴史的価値のある古文書や、著名作家の原稿など、文化的に有益な文書の電子化が行われている。このような文書は、その文書内容だけではなく、文字の字体や筆跡などの

情報も、時代背景や執筆者の特徴など、様々な歴史的価値のある情報を映し出す、貴重な資料である。よって、文書のテキストデータを電子化して保存することに加え、文書そのものを画像データとして電子化し、アーカイブとして扱う重要性が高い。そのような中で、それぞれのデータ上の文字の位置情報を関連付けるために用いられるのが、トランスクリプトマッピングである。

## 1.2 研究目的

現状のトランスクリプトマッピングを題材とした研究は、文書のトランスクリプトと画像データから、トランスクリプトマッピングデータを自動的に生成して利用する研究が一般的である。しかし、生成されたデータは、それぞれの研究で定められた目的の範囲内での活用にとどまっているのが現状である。これは、研究の中でマッピングデータを生成する目的が、文字認識のための学習用データのために利用することを前提としたデータであるなど、用途が限られたものであるため、他の用途に利用することを前提に作られていないことに起因している。そのため、研究者以外の人物がマッピングデータを利用するのは困難である。さらに、マッピングデータを扱うための規格が存在しないため、各々の研究ごとにデータの扱いが異なる。これらのことから、仮にマッピングデータを生成する研究の対象となる文献が有益なものであったとしても、そのデータを一般利用などの用途で使うことは難しく、幅広い分野におけるマッピングデータの活用は実現されていない。

そこで本研究では、研究以外の用途においても、共通して用いることのできるデータ書式を提案することで、有益な文献のマッピングデータをより広い分野で活用できるようにすることを目的とする。データ書式を共通化することで、携帯端末や Web を通じたオンライン利用など、様々な環境のアプリケーションの下での相互利用性を確保し、幅広い場面でデータを連携利用できるようになることが期待できる。

## 1.3 関連研究

トランスクリプトマッピングを題材とした研究は以前から行われており、Tomai ら<sup>1)</sup> は英文の古文書を対象として、文書画像から各単語の位置を切り出し、トランスクリプトの単語とのマッピング処理を行っている。英文などの西洋圏の言語の場合、単語間にスペースがあるため、単語の範囲座標の切り出しや、単語単位でのトランスクリプトマッピングを比較的容易に行うことができ、マッピングされたデータは、文書画像上におけるワードスポットティング（文書中から特定の単語の位置を検索する）に用いられている。

日本語や中国語など、単語間でスペースによる区切りがされない言語では、文字単位でのマッピングが中心となる。Zhou ら<sup>2)</sup> は、中国語および日本語のオンライン手書き文字（文

字を書いている筆跡の情報が判明している手書き文字）を対象に、文字単位でのトランスクリプトマッピングを行っている。マッピングの目的は、手書き文字の認識精度を向上するためのサンプルを収集するためであるため、生成されたマッピングデータを、対象となった文の閲覧補助など、他の目的に利用することは特に想定されていないようである。

マッピングデータの一般利用に向けた研究として、小西ら<sup>3)</sup> は、日本語による毛筆手書きの古文書画像を対象としたトランスクリプトマッピングを行っている。歴史的資料などを扱うデジタルアーカイブの取り組みの中では、古文書などを扱う機会が多く、草書体のような崩された字が多い。そのため、文書を楷書の文字で一般に読めるように、翻刻作業を行った上で内容を電子化することが一般的である。そのような文書画像と、トランスクリプトの間でトランスクリプトマッピングを行うことで、難読文字の多い文献を対応する活字を参照しながら閲覧できるようになり、一般利用が容易になるというメリットがある。また、相原、林<sup>4)</sup> は、難読文字の存在する資料の翻刻を支援するツールを開発している。この中では、ネットワークでデータを共有して翻刻作業を行うための、協調型のデータ活用例を提示している。

実際にトランスクリプトマッピングが行われたデータを一般向けに活用している例としては、イスラエル考古学庁と Google が共同でインターネット上で公開している、死海文書を閲覧するためのビューワーである“Digital Dead Sea Scrolls”<sup>5)</sup> がある。死海文書は、聖書の原本として確認されている中でもっとも古い文書で、歴史的価値が非常に高い。ビューワーを使用すると、文書画像をスクロールして閲覧することが可能である。死海文書には、文単位でトランスクリプトマッピングが行われており、画像上の文を選択すると、該当する文のトランスクリプトを英訳した文を参照することが可能である。

## 2. 提案データ書式

本研究では、記述形式として XML<sup>\*1</sup> を用いたトランスクリプトマッピングデータを提案する。XML は、データを階層構造で記述するマークアップ言語の一つである。

マッピングデータは、最小限のデータを考えると、文字の位置と座標、という極めて単純なデータであるため、XML 形式を使用しなくても、独自形式や、CSV<sup>\*2</sup> 形式（パラメータをカンマ区切りで記述する形式）でも十分対応可能である、しかし、共通のデータ書式と

\*1 eXtensible Markup Language

\*2 Comma Separated Values

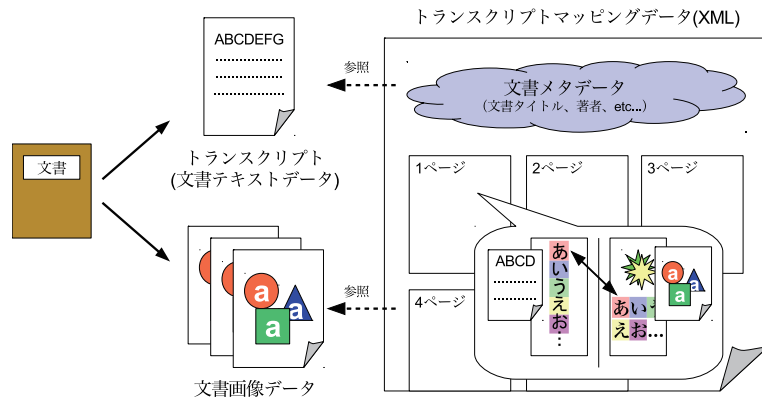


図 2 提案データ書式のデータ構造図  
Fig. 2 Data structure diagram of proposing format.

して運用する場合、研究者から一般ユーザーまで、様々な立場からデータは扱われることを考慮しなければならない。それ故に、XML 形式そのものに対する一般的な認知度をはじめとして、データ内容の視認性、アーキテクチャを問わず利用できる共通性、柔軟な記述に対応する拡張性などの利点を踏まえた上で、XML を採用した。特に研究目的でのマッピングデータにおいては拡張性は重要であり、研究で利用するためのパラメータが標準のデータ書式で定義されていない場合でも、独自の属性を付加することでパラメータを拡張できるメリットがある。

### 2.1 基本構造

提案するトランスクリプトマッピングデータの基本構造を示した略図を、図 2 に示す。一般的なマッピングデータは、トランスクリプトを記述したテキストファイル、文書画像をページごとに取り込んだ画像データファイル群、対応付けの情報を記述したトランスクリプトマッピングデータファイルの三種類のファイルで構成されている。

トランスクリプトマッピングデータは、文書画像を基準として、位置情報などのマッピングのためのデータを文字単位で格納する。文字の位置情報の対応付けは、文書画像上の範囲座標と、トランスクリプト上のテキストの文字位置をペアで指定することで行う。この対応付け情報は、文書を構成する各ページごとに分割して管理される。これは、一般的に画像ファイルがページ単位で管理されることに基づいたもので、各ページで異なる画像ファイルを参照し、マッピング情報をページ別に格納する。また、マッピングデータの他に、単語や

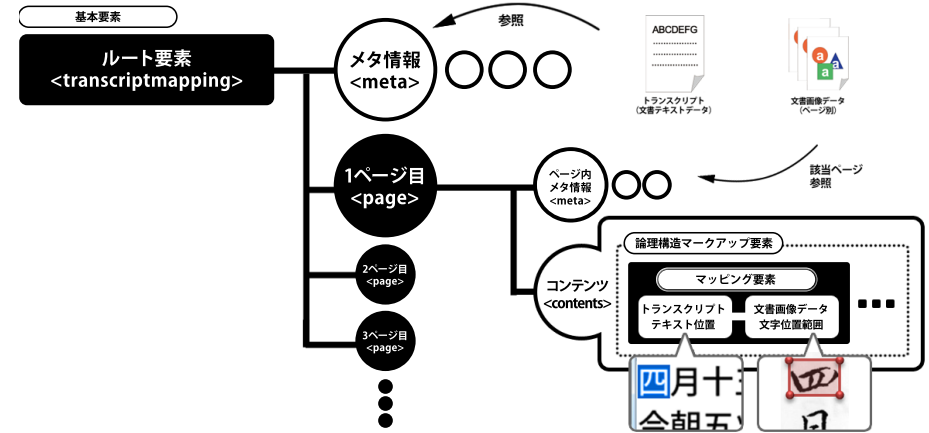


図 3 マッピングデータの XML 要素構造図  
Fig. 3 XML elements structure diagram of mapping data.

段落など、文書内における各文字の論理構造を定義するためのマークアップや、文書のタイトルや著者などの、文書そのものに関する情報を記述するためのメタデータ情報を格納することができる。

### 3. マッピングデータの構成要素

本項では、マッピングデータファイルを構成する XML の要素について解説する。要素の役割は三つに大別され、データを構成するための「基本要素」と、実際にトランスクリプトマッピングを行い文字の位置情報を対応付ける「マッピング要素」、文書の論理構造を定義する「論理構造マークアップ要素」がそれぞれ存在する。マッピングデータの各要素の木構造図を図 3 に示す。

#### 3.1 基本要素

基本要素は、マッピングデータが成立するために必要な要素群であり、データの核となる。ルート要素 (transcriptmapping) XML がマッピングデータであることを宣言するために、ルート要素として、(transcriptmapping) 要素を取る。このルート要素の子要素には、トランスクリプトが書かれたテキストファイルの場所などの、文書全体のメタ情報を定義するメタデータ要素と、文書のページを構成するページ要素を置くことができる。ページ要素 (page) ページ要素は、文書の各ページを管理するための要素であり、(page)

で表記される。ページ要素でも同様に、メタデータ要素を子要素として配置することが可能で、ページ別に分かれた文書画像データファイルを指定するなど、各ページごとに異なるメタ情報を定義するために使用する。ページ内における実際のマッピング情報は、子要素としてコンテンツ要素 (contents) を定義し、その中でマッピング要素や論理構造マークアップ要素を記述する。

メタデータ要素 (meta) 文書全体および、各ページに必要なメタ情報は、メタデータ要素 (meta) を使用して定義する。ここでは、参照するトランスクリプトや画像データファイルの指定を行う他、文書に関する任意の情報を格納する。XML 要素の属性として、定義するメタ情報の種類を指定する name 属性と、その値を指定する value 属性を取る。属性の指定例を表 1 に挙げる。研究で独自に使うパラメータなど、標準で定義されていない情報に関しては、name 属性に予約されていない任意の名前を指定することで、独自にメタ情報を定義することができる。

### 3.2 マッピング要素

実際に文字の位置を対応付けるトランスクリプトマッピングは、マッピング要素で主に行われる。マッピング要素は、コンテンツ要素 (contents) 内で定義され、文書中の文字一文字ずつをマッピングの単位として、トランスクリプトのテキスト位置と文書画像上の文字位置の範囲の関連付けを行う。

文字マッピング要素 (char) 文字ごとにトランスクリプトマッピングを行っているのが、文字マッピング要素 (char) である。この要素の属性に、該当する文字のトランスクリプト上のテキスト位置、および文字列の長さ、そして文書画像上の文字位置の範囲を指定することで、二つの文字が対応付けられ、マッピングが成立する。

トランスクリプトの文字位置は、テキスト位置を示す index 属性と、対象となるテキストの長さを示す length 属性で指定する。length 属性は文字単位の対応付けには不要にも

表 1 (meta) 要素におけるメタ情報指定例  
Table 1 The example of defining metadata in (meta) elements.

name 属性の値	value 属性の値	ルート要素での定義	ページ要素での定義
transcript	トランスクリプトファイル	TRUE	TRUE
image	文書画像ファイル	TRUE	TRUE
title	文書のタイトル	TRUE	FALSE
author	文書の著者 (複数指定可能)	TRUE	FALSE
text-direction	文字の方向	TRUE	TRUE
page-number	ページ番号	FALSE	TRUE



図 4 複数文字のトランスクリプトへの文字マッピング例  
(左: トランスクリプト 2 文字分, 右: 文書画像 1 文字扱い)

Fig. 4 The example of character mapping to multi-character transcript.  
(Left: Transcript, Two characters. Right: Document image, treating one character.)

思えるが、文書画像上では 1 文字に見えても、トランスクリプトで表記すると 2 文字以上になる文字の存在があるために、属性を設けている。例えば、日本語の草書体で書かれた「より」という言葉は、トランスクリプト上の表記は 2 文字だが、文書画像で見ると 1 文字のように表記されることがある (図 4)。

また、text 属性に文字列を定義することで、マッピングデータ内に直接、対応する文字を埋め込むことができる。これにより、トランスクリプトが存在しない場合でも、該当する範囲に任意の文字を対応づけることが可能になる。

文書画像データ上の文字の範囲座標は、coords 属性を用いてピクセル単位で指定する。この属性は、Web ページを記述するための言語である HTML<sup>\*3</sup> で用いられる、クリックブルマップの記述方法を採用することで、習得や理解を容易にしている。ただし HTML のそれとは異なり、デフォルトでは、多角形ポリゴンによる座標をカンマ区切りで指定することで、文字の範囲を指定する。画像データを取り込む際に生じたずれなどによって、文字の傾きなどが発生しうるため、矩形範囲のみでは、複雑な形状の文字をカバーすることが難しいためである。shape 属性によって、coords 属性の範囲指定を、矩形に変更することも可能である (polygon, rectangle に対応)。

セグメント範囲指定要素 (segment) 基本的に (char) 要素は子要素が定義されない空要素だが、もし文字より小さな単位の範囲データを作成する場合、子要素として、セグメント範囲指定要素 (segment) を置くことができる。セグメントとは、文字をさらに細かく分割した、文字を構成する最小単位を指す。漢字を部首に分割したり、あるいは図 5 のような

\*3 HyperText Markup Language

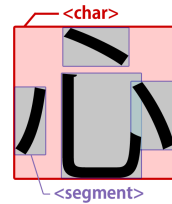


図 5 (char) 要素と (segment) 要素の文字範囲指定  
Fig. 5 Specify character areas by (char) element and (segment) element.

連結成分単位に分割する必要がある場合の利用が想定されている。マッピングデータを生成する研究の中には、まず文書画像上の文字をセグメント単位まで細かく分割してから、それぞれを文字になるように統合する処理が行われるものがあるため、そのような研究においてもデータ書式を活用できるように、セグメント単位の範囲指定が実装されている。使用できる属性は coords 属性と shape 属性のみで、あくまで文書画像上の範囲指定に特化しており、セグメントとトランスクリプトの文字をマッピングさせることはできない。

### 3.3 論理構造マークアップ要素

マッピングデータに文書の論理構造を反映する、論理構造マークアップ要素は、研究システムや一般利用アプリケーションでの活用が期待できる。

このデータにおけるマッピングの最小単位は文字としているが、実際の研究の場におけるトランスクリプトマッピングではそうとは限らない。単語単位・文字行単位など、ある程度大きな単位でマッピングを行っている研究は数多い。よって、一文字一文字のマッピングだけではなく、単語や行などの大きな部分でマッピングができる構造でなければ、より広い分野で活用できるデータ書式とは言えないだろう。研究分野だけでなく、一般利用の点でも、単語・文・段落などの文書の論理構造をマッピングデータに記録することには利点がある。文書データが意味のあるまとまりで細分化されることで、特に Web のような通信媒体を用いたデータ利用の場面では、「特定の部分の文が見たい」などのユーザーのリクエストに合わせてデータを転送することで、一度のデータ転送量の削減と、高速なレスポンスを実現することができ、データ内容にアクセスする際の効率が向上する。よって、論理構造によってデータをマークアップすることは、広い分野で活用できるデータ書式の実現という意味では極めて重要であると言える。

論理構造マークアップ要素では共通して、マッピング要素と同じようにトランスクリプトマッピングのための情報を定義することが可能で、(char) 要素で使用できる属性が同じよ

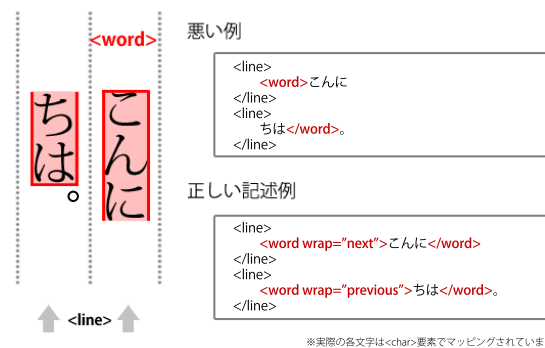


図 6 wrap 属性を利用したオーバーラップ  
Fig. 6 Overwrapping with wrap attribute.

うに使用できる。そのため、(char) 要素などのマッピング要素を使用せず、論理構造マークアップ要素を用いて単語単位・行単位でトランスクリプトマッピングをすることも可能な設計になっている。

論理構造を定義するにあたって問題となるのは、XML 文書の仕様上、要素同士がオーバーラップできないことにある。例えば、行と単語の両方に論理構造マークアップを行っていて、単語が行をまたいでいる場合は、素直に書くと XML 文書として正しくないものになってしまう。そこで、全ての論理構造マークアップ要素において、オーバーラップに対処するための wrap 属性を指定することができる(図 6)。これは該当する要素のマークアップが、他の要素の終了をまたいで連続していることを定義するもので、値に “next” を指定して、次に記述される同じ論理属性の要素まで、内容が連続していることを示す。受ける側の要素には “previous” を指定する必要がある。

汎用ブロック要素 (block) (char) 要素でマッピングされた文字からなる、単語や文などの意味のある構造のことを“ブロック”と呼ぶ。汎用ブロック要素 (block) は、単語・文・行・段落などに属さない、汎用的なブロックを作成するために用いる。これは仮に文法的には意味を成さないブロックでも、研究利用などにおいて指定する範囲が特別な意味を成す際に活用できる要素である。ブロックがどのような意味を持つのかを定義するための type 属性が存在し、任意の文字列を定義し、ブロックの用途を任意に定めることを可能にしている。

基本論理要素群 単語や文など、一般的に利用される論理構造は、論理構造マークアップ要

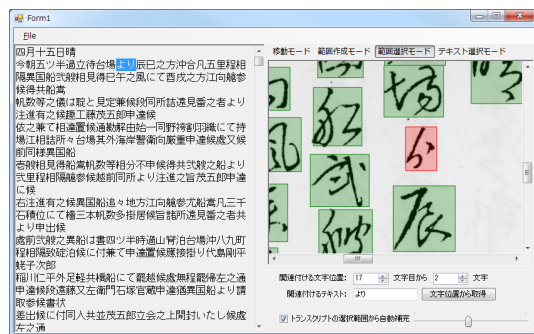


図 7 マッピングデータを作成するプロトタイプツール  
Fig. 7 Prototype tool for making transcript mapping data.

素の基本論理要素群として定義されている。現状の策定段階では、要素として

- 〈word〉(単語)
- 〈line〉(行)
- 〈sentence〉(文)
- 〈paragraph〉(段落, HTML 互換記法 <p> で略記可)

が定義されており、それぞれ対応する文書の論理構造を表現するために使用することができる。

#### 4. 実験

以上のようなデータ書式の提案に至るまで、我々はマッピングデータを作成するプロトタイプツールの製作と、実際にマッピングデータの作成実験を通して、データに必要な書式の策定を進めた。

プロトタイプツールは、トランスクリプトと画像データの文字の位置を手動で対応づけ、提案するデータ書式を用いてマッピング情報の読み書きを行うことのできるツールである(図7)。言語には C# を利用して開発した。マウスで画像上に範囲を作成し、トランスクリプト上の文字の位置を指定して対応付けを行う。現時点では、単体ページ(〈page〉要素が1つのみ)に対するマッピングに対応している。論理構造マークアップ要素や、多角形範囲指定、トランスクリプト・文書画像データファイル以外のメタデータは、プログラム内のデータ構造としては実装されているが、現時点ではインターフェースに反映されていない。

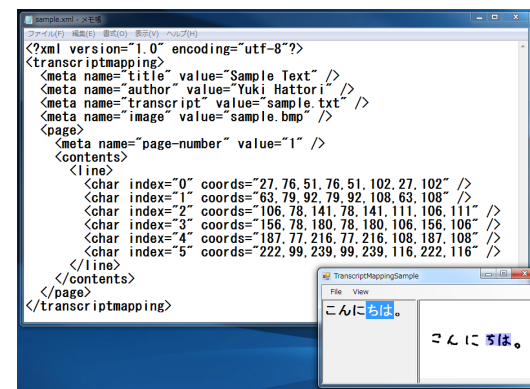


図 8 実際のデータ記述例と、インターフェース実験ツール(画像データ上の文字選択)  
Fig. 8 The example of data and interface experimental tool. (Selecting characters on image data.)

また、このツールとは別に、提案するマッピングデータを活用するためのアプリケーションインターフェースの実験ツールとして、画像データ上で文字選択を行うツールを開発した(図8)。これはマッピングデータを用いることで、画像データをテキストのようにドラッグで選択し、該当部分のトランスクリプトのデータを抽出できるものである。これにより、文書画像に書かれた文字を、テキストデータのように直感的に扱うことができるようになる。その結果、画像とテキストの垣根が無くなり、文書画像を扱う上での利便性の向上に繋がるものと考えられる。

#### 5. 評価と考察

プロトタイプツールを製作しながら、実際の史料を用いたトランスクリプトマッピングを行い、データの評価・策定を進めた。マッピングを行った結果、当初のデータ書式に問題点を発見次第、改善・フィードバックを積み重ねて、現在のデータ書式に至る。

初期のデータ書式では、「亜国来使記」\*4のトランスクリプトと画像のマッピングを試行した際、草書体に1文字の表記でトランスクリプト2文字分の字があることを発見した結果、〈char〉要素の length 属性として反映されている。また、英文のハイフネーションのような、テキストでは無視すべきだが、画像上では無視できない文字の存在があり、初期のデータ書

\*4 安政元年(1854年)、ペリー艦隊が函館を訪問した際の、松前藩による応接記録。

式でそのようなものが含まれる文書に対し、画像データ上で文字選択を行うツールを用いて実験を行ったところ、正しくテキストを抽出できなかった。当初のデータ書式の〈char〉属性は、対応付けが必須だったことが原因であり、現在は位置座標のみを記述し、対応するトランスクリプトの位置情報を省略可能にすることで解消している。

現段階では、策定段階の初期にあたるため、データ書式の提案に留まっているが、今後はこのデータ書式を具体的に広く活用させることを構想している。まだデータ書式を使った具体的なアプリケーションというものは存在しないので、実際にデータ書式を活用できるアプリケーションを開発しなければ、その有用性は証明されない。近年の躍進が目覚ましい携帯端末や、マッピングデータを活用した実績のまだ少ないWebなど、データを活用し得るフィールドは多岐に渡っており、データの連携利用という利点を生かしたアプリケーションの開発を視野に入れている。また、より幅広い分野での利用を促すなら、データ書式を扱うためのAPIを整備する必要もある。このような動きの中で、データ書式の有用性を評価し、現状の書式で本当に幅広い分野での活用に耐えうるかを見極めなければならない。

勿論、現時点のデータ書式そのものの改善も必要である。現状の書式は、冗長的なメタ情報の記述方法など、使い勝手の面で改善の余地がある。マッピングデータに必要とされる情報、あるいは不必要な情報を判断し、現在のデータ書式をより洗練しなければならない。そのためには、トランスクリプトマッピングのシーズを貪欲に追うだけではなく、現状求められるニーズをきちんと把握した上で、書式を思慮するべきであろう。

## 6. おわりに

本研究では、トランスクリプトマッピングを扱うためのデータ書式を提案した。データの記述形式としてXMLを採用し、現状のトランスクリプトマッピングの研究・活用に沿ったデータ書式を提案した。実際に提案するデータ書式をマッピングに用いたツールを実験材料としてデータ書式を評価し、発見した問題点を逐次フィードバックした上で、どのような情報がデータ書式に必要であるか考慮した。研究の最終的な目標は、幅広い場面・幅広い環境における、マッピングデータの相互連携利用であり、実際にデータ書式を活用した場面で、有用性を評価する必要がある。

今後は、トランスクリプトマッピングに本当に必要な情報は何かを取捨選択しなければならないという点に留意しながら、実際にデータ書式を活用したアプリケーションを製作し、利用する中でデータ書式の有用性を評価することで、よりデータ書式の完成度を高める作業を進めていく予定である。

## 参 考 文 献

- 1) C. Tomai, B. Zhang, and V. Govindaraju, "Transcript mapping for historic hand-written document images", Int. Workshop on Frontiers in Handwriting Recognition (IWFHR-8), pp. 413-418, 2002.
- 2) Xiang-Dong Zhou, Fei Yin, Da-Han Wang, Qiu-Feng Wang, Masaki Nakagawa, and Cheng-Lin Liu, "Transcript Mapping for Handwritten Text Lines Using Conditional Random Fields", International Conference on Document Analysis and Recognition, pp. 58-62, 2011.
- 3) 小西隆弘, 寺沢憲吾, 川嶋稔夫, "毛筆手書き文字画像へのトランスクリプトマッピング", 信学技報, PRMU2005-215, pp. 79-84, Feb. 2006.
- 4) 相原健朗, 林晋, "画像化主義に基づく文献資料研究用ツール SMART-GS とその発展", 情報処理学会研究報告, 2011-DD-79, No.5, pp. 1-5, 2011.
- 5) The Israel Museum, "Digital Dead Sea Scrolls", available from (<http://dss.collections.imj.org.il/>) (accessed 2011-12-14)