

歴史文献のための電子スクラップブックシステムの設計

石川 正敏† 波多野 賢治‡ 天笠 俊之‡ 植村 俊亮‡ 勝村 哲也†

† 島根県立大学 総合政策学部

{m-ishikawa, t-katsumura}@u-shimane.ac.jp

‡ 奈良先端科学技術大学院大学 情報科学研究科

{hatano, amagasa, uemura}@is.aist-nara.ac.jp

要旨

本稿では、まず利用者による歴史文献の注釈と文献の分類を支援する電子スクラップブックシステムについて報告する。提案システムは、文献画像と注釈を管理する文献データモデルと文献の分類を管理する電子スクラップブックデータモデルに基づいて、注釈の編集、切り抜き、検索などを処理する。次に、文字や単語、地理情報などの内容に基づいた注釈の分類を行い、それぞれの注釈モデルを提案する。このような注釈モデルを用いることで、歴史文献の効率的な利用が可能になると考えられる。さらに本稿では、提案する注釈モデルの利用例として、文献画像のレイアウトに則した釈文の表示について述べる。最後に、提案モデルを用いた歴史文献に関する情報の長期的な管理について考察し、提案モデルの実用性を示す。

A Design of an Electronic Scrapbook System for Historical Documents

Masatoshi Ishikawa† Kenji Hatano‡ Toshiyuki Amagasa‡ Shunsuke Uemura‡
Tetsuya Katsumura†

† Department of Policy Studies, the University of Shimane

{m-ishikawa, t-katsumura}@u-shimane.ac.jp

‡ Graduated School of Information Science, Nara Institute of Science and Technology

{hatano, amagasa, uemura}@is.aist-nara.ac.jp

Abstract

In our research, we developed an electronic scrapbook system which aids users to annotate and classify East Asian historical document images published on the Internet such as ancient documents or ancient wooden-slips. This system uses our document data model for managing relationships between historical documents and annotations, and our electronic scrapbook data model for managing classification of document images. We designed operations of the system for editing annotations, clipping document image and searching documents. In this paper, we mentioned a variety of annotations including information of East Asian characters and meanings of words, and designed the data model for each annotation. By using our proposed annotation models, users can use document data efficiently. Display of text data according to the document image layout is shown as an example of our proposed annotation models. By our electronic scrapbook system and annotation models, users can edit and share information related to East Asian historical documents on the Internet efficiently.

1 はじめに

近年、図書館や博物館などでは、保存が困難である木簡や古文書などの東アジア圏の歴史文献の電子化を進めているところが多い。また、電子化された歴史文献をインターネットで公開する研究機関も増加している。電子化された歴史文献をそのまま展示しただけでは、一般的な利用者が文献の内容を理解できるとは限らない。従って、文献の内容理解を支援するために注釈は重要であると考えられる。しかし、歴史文献を公開している図書館などの組織が、文献に関連する情報をすべて利用者に示すことは困難であることが多い。

そこで本研究では、利用者自身がインターネット上の歴史文献に対して注釈を付けることを考える。注釈を付けられた歴史文献を利用者間で共有することによって、歴史文献自体の利用効率の向上が期待できる。そこで、本稿では、インターネット上に公開されている歴史文献の画像の利用者による注釈の添付や、文献の切り抜きなどを支援するための電子スクラップブックシステムについて述べる [1]。提案システムでは、歴史文献や注釈の管理に文献データモデルと電子スクラップブックデータモデルを用いる。文献データモデルは、歴史文献の画像と漢字や地理などの関連情報を注釈として記述するモデルである。また、このモデルでは、対応表を用いて注釈を付ける文献画像の位置を詳細に記述する。電子スクラップブックデータモデルは、文献データや電子スクラップブックデータの分類を記述するモデルである。提案モデルの記述には、インターネットを介したデータ交換に広く利用されている XML [2] を用いる。さらに、文献データモデルでは歴史文献の外字の記述に XML を用いる。本研究では、提案モデルの操作として注釈の編集、切り抜き操作、および文献の検索について述べる。切り抜き操作とは、文献画像の一部とその画像に関連した注釈を取り出し新たな文献データを生成する操作である。

本研究で用いる注釈は、利用者の歴史文献に対するコメントの他に文字情報や単語の意味、地理情報など様々な種類の情報を集約するために利用される。そこで本稿では、情報の種類の異なる注釈を検索などで効率的に利用するために、それぞれの情報の特徴に併せたデータモデルを提案する。また、提案する注釈モデルの利用例として、文献画像の見た目をテキストデータで再現する方法について述べる。このようなテキストデータで文献を再現することは、歴史文献の閲覧や内容理解の支援に有効であると考えられる。

最後に、提案モデルを用いた長期に渡る歴史文献と注釈などの関連情報の管理について考察し、提案モデルの有効性を示す。また、電子スクラップブックシステムと注釈モデルの利用方法の提案から、歴史文献画像を用いた情報の集約と再利用が可能になるため、意見交換などの歴史文献の利用効率の向上が期待できる。

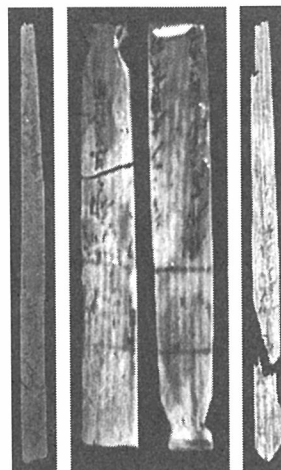


図 1: 木簡画像の例

2 歴史文献に関する考察

木簡を対象に歴史文献の特徴から、文献画像の管理と注釈の記述などに関する要求を挙げる。

古代の木簡は、荷札以外に行政文書、習字などに利用されていた木の札である (図 1)。木簡は、木の腐敗や変色が多いため、内容の読解が困難であることが多い。また、木簡には、複数の破片に分割された状態で発掘されることもある。従って、歴史文献を公開する場合、原資料の計算機上での再現に適した画像に併せて釈文が必要であると考えられる。さらに、複数の木片画像を配置して元の木簡を再現することを考える必要もある。

歴史文献の記述は、字のかすれや虫食い、草書体などのように読むことが困難な場合が多い。従って、文献の注釈を単に列挙しただけでは歴史文献のどの部分の注釈であるかを判別できないことがある。このため文献に注釈を付ける場合、注釈を関係付ける文献画像の領域を詳細に述べる必要があると考えられる。また、歴史文献の注釈は文書だけではなく、文献中の図や傷などにも付けられると考えられる。

東アジア圏の歴史文献の記述には、外字と呼ばれる Unicode などの標準的な符号化文字集合にない文字が含まれていることが多い。一般に外字は、文字列検索の対象にならないが、歴史文献では、人名などで外字が利用されるため、外字も検索対象として利用できるようにしなければならない。また、注釈や釈文として関連付けられてテキストデータの外字表示も文献を読む上で重要であると考えられる。

歴史文献の利用の一つである文献の収集は、利用者が文献や時代背景を分析するために行われる。このような文献の収集は、文献の全体の他に利用者の着目する文献の一部分だけを収集することがあると考えられる。

3 データモデル

図2は、本研究で提案しているデータモデルの構成である。図2下に示す文献データモデルは歴史文献の画像と注釈の関係性を記述するためのモデルである。図2上に示す電子スクラップブックデータモデルは、文献データなどの分類を記述するモデルである。また、文献データA'は、文献データAの文献画像の部分画像とその画像に関係付けられた注釈からなる文献データである。提案モデルの記述には、インターネットを介したデータ交換で広く利用されているXMLで記述する。

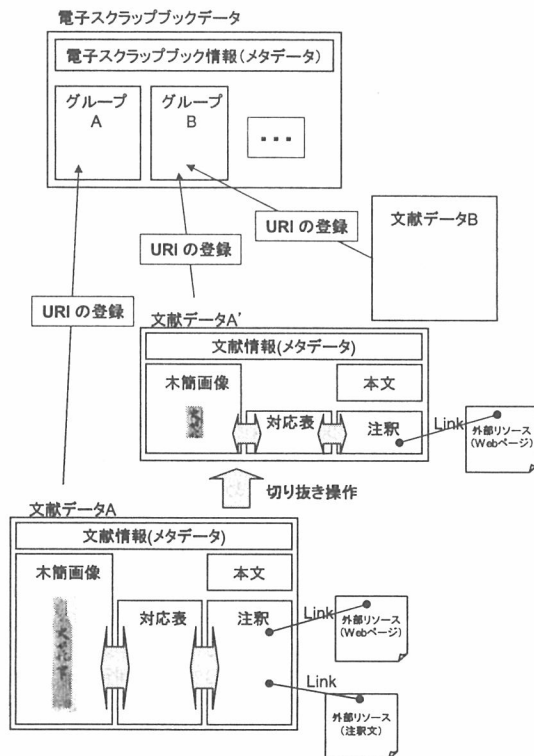


図2: データモデルの構成

3.1 文献データモデル

文献データモデルの構成は以下の通りである。また、本節では要素ごとに記述例を挙げる。

(1) 文献情報

文献画像に関する出典や著者などのメタデータを記述する。本稿では、Dublin Core Metadata Element Set Ver 1.1 [3] に従って記述する。この

ような標準に従うことで、他のメタデータシステムとの情報交換が容易になると考えられる。

(例)

```
<文献情報>
<title>官位木簡 </title>
<subject>木簡 </subject>
...
</文献情報>
```

(2) 文献画像

注釈を関係付ける対象の文献画像のURIを記述する。複数の画像を用いて一つの文献画像を表現する場合は、それぞれの画像のURIを列挙する。複数の画像を列挙する場合、列挙された画像を表示するための情報として画像のサイズと配置情報を与える。配置情報は、画像の左上の頂点を基準にした位置情報である。

(例)

```
<文献画像> mokkan1.jpg </文献画像>
```

(3) 本文

文献画像として示される歴史文献の積文を記述する。また、外字は、次に述べる注釈への参照を用いて表現する。

(例)

```
<本文>大花下 </本文>
```

(4) 注釈

歴史文献の注釈は、コメントだけではなく文字情報、単語の意味などを含む。本稿では、利用者が直接書き込むテキストデータの他に、Webページなどの外部リソースを注釈として利用する。注釈が外部リソースである場合は、そのリソースのURIを記述する。また、個々の注釈は、文献データ内で識別するための識別子を持つ。識別子は、本文での外字の参照や、対応表での文献画像の領域との関係の記述に利用する。

(例)

```
<注釈>
<注釈文><識別子>1</識別子><内容>奈良時代の官位 </内容></注釈文>
<注釈文><識別子>2</識別子><内容>... </内容></注釈文>
... </注釈>
```

(5) 対応表

インターネットで公開される歴史文献の画像に対して直接注釈を埋め込むことはできない。そこで、本研究では、対応表を用いて文献画像の領域と注釈の関係性を記述する。注釈を付ける文献画像の領域は、長方形の左上の頂点と右下の

頂点の座標の組で表現する。また、座標の原点は、文献画像の左上とする。

(例)

```
<対応表>
<要素><領域>0,0,100,200</領域><識別子>1</
識別子></要素>
<要素><領域>100,110,150,100</領域><識別
子>2</識別子></要素>
...
</対応表>
```

3.2 外字情報の記述

本研究では、外字情報の記述に注釈を利用する。注釈に記述する内容は、漢字の読みなどの属性情報や、文字画像、eKanji[4]や今昔文字鏡[5]などのインターネット上で利用可能な大規模漢字情報への参照が挙げられる。このように外字情報の記述に注釈を用いることで、文献固有の文字情報と一般的な文字情報の併記が可能になる。釈文などでの外字の記述は、注釈への参照を用いて記述する。

3.3 電子スクラップブックデータモデル

電子スクラップブックデータモデルの構成は以下の通りである。

(1) 電子スクラップブック情報

電子スクラップブック情報では、電子スクラップブックデータの所有者や概要などのメタデータを記述する。記述項目は、文献情報と同様に Dublin Core Metadata Element Set Ver 1.1 に従う。

(例)

```
<電子スクラップブック情報>
<title>木簡データ集</title>
<subject>奈良時代の木簡</subject>
...
</電子スクラップブック情報>
```

(2) グループ

利用者が収集したデータの分類を記述するための要素であり、電子スクラップブックデータ内に1つ以上含まれる。分類されたデータの記述には分類されるデータのURIを用いる。分類の対象のデータは、文献データと電子スクラップブックデータである。さらに、グループは、グループの分類目的を表すメタデータとデータごとの配置の記述を許す。

(例)

```
<グループ>
```

```
<キーワード>奈良時代の木簡</キーワード>
<収集データ>mdata01.xml</収集データ>
<収集データ>mdata02.xml</収集データ>
...
</グループ>
```

3.4 操作

本節では、提案モデルの操作として注釈の編集、切り抜きおよび検索について述べる。

(1) 注釈の編集

注釈の編集は、文献に関連する情報を集約するための操作である。ここでは、注釈の追加と削除について述べる。注釈の追加は、注釈文の作成と文献画像の領域との関係付けからなる。注釈文は、テキストデータの他に既存のWWWページなどの外部リソースへのリンクを含む。さらに注釈文の作成では、注釈文に文献データ内で識別するための識別子を与える。次に、文献画像の領域と注釈の関係付けは、対応表に文献画像の領域と参照する注釈文の識別子を登録する操作である。注釈の削除操作は、利用者が与える注釈の識別子に従って、注釈文と対応表に登録された領域情報を削除する操作である。

(2) 切り抜き操作

切り抜き操作は、文献画像から利用者の着目する部分画像の抽出と、抽出された画像に関係付けられている注釈を取り出す操作である。また、抽出されたデータは文献データモデルに従うので、そのデータに対してさらに注釈編集などの操作が可能である。

(3) 検索

文献画像を用いた文献検索は、検索結果がOCRなどの画像処理の性能に依存するため効率的であるとは言えない。特に、木簡などの文字の判別が困難な画像の場合、画像を用いた文献検索は不可能であると考えられる。そこで、本研究の検索は、文献画像に付けられた注釈を利用する。注釈を用いた検索は、注釈を付けた利用者と同じような視点を持つ他の利用者に対して有効であると考えられる。また、このような検索は、文献の記述だけではなく、文献の意味に基づいた検索が実現できると考えられる。

本稿の検索処理は、文献データの文献情報や注釈などのテキストデータに対する文字列一致を用いる。検索条件は注釈などの検索対象の要素名と文字列の組の集合であり、検索は次の2段階で処理する。(1) 検索条件の各組に従って文字列一致を行い指定した文字列を含む文献データのURIを一時的な結果として保存し、すべての条件の組に対して同様の文字列一致処理を行う。(2) 処理(1)で得られるURIを集計し、URIの重複の多いものから順にn個を検索結果として利用者に示す。

4 電子スクラップブックシステム

本節では、提案モデルの実用性などを検証するために実装した電子スクラップブックシステムのプロトタイプについて述べる。

4.1 システム構成

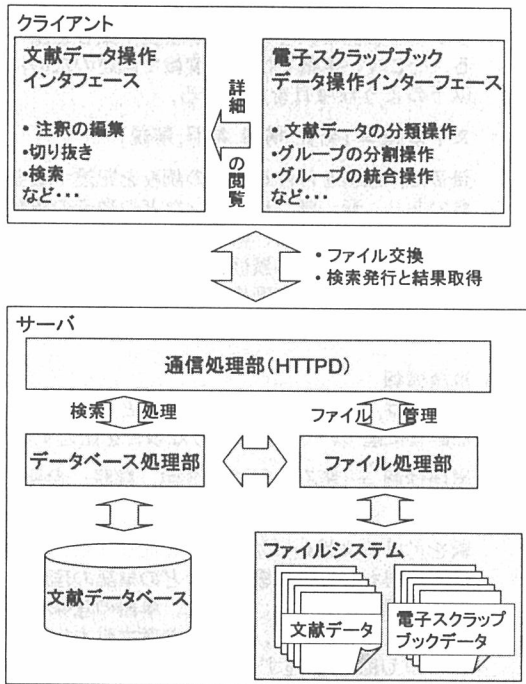


図3: プロトタイプシステムの構成

本研究で実装しているプロトタイプシステムの構成を図3に示す。本システムは、利用者による文献画像の注釈編集などを支援するクライアントと、文献データの管理と検索などを処理するサーバから構成され、クライアントとサーバ間の通信の protokolには HTTP を利用している。

● クライアント

クライアントは、文献画像に対する注釈の編集や切り抜き操作などを支援する文献データ操作インタフェースと、文献データや電子スクラップブックデータの分類を支援する電子スクラップブックデータ操作インタフェースからなる。

● サーバ

サーバは、クライアントとの通信を処理する通信処理部、クライアントからの検索要求などを処理するデータベース処理部とファイル処理部、

検索を効率的に処理するための文献データベースからなる。通信には HTTP を利用するので通信処理部は、一般的な WWW サーバを利用する。また、文献データベースは、一般に広く利用されている関係データベースを利用する。

4.2 実行例

プロトタイプシステムの実行例として、文献画像に対する注釈の編集例を図4に示す。本プロトタイプシステムでは、日本古代木簡選に収録されている木簡画像と解説を利用している [6]。

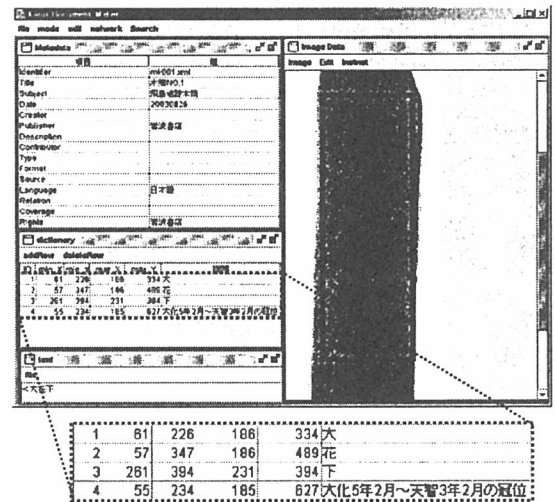


図4: 文献画像の注釈の追加

図4は、文献データ操作インタフェースの実行例であり、このインタフェースは、文献画像、文献情報、本文、注釈と対応表を併せたデータを表示するウィンドウ群からなる。例では、文献画像にある長方形で示す領域に含まれる語に対して注釈を付けており、この領域と注釈文の関係は、図4下部に拡大表示している表によって管理する。注釈編集の他に本インタフェースは、切り抜き操作、検索操作が可能である。

図5は、利用者が収集した文献データの分類例である。このインタフェースでは電子スクラップブックデータのグループごとに文献画像を表示している。図5では、官位を表す記述のある木簡とそれ以外の木簡に分類した例である。各木簡に付けられた注釈などの閲覧には、文献データ操作インタフェースを用いる。

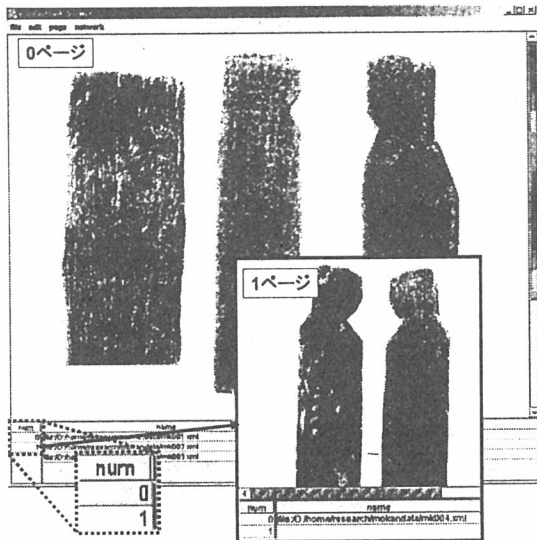


図 5: 文献データの分類

5 注釈の内容に基づくモデル化

前節まで述べてきた文献データモデルで用いる注釈は、記述に制約がなく利用者による自由な形式での記述を許していた。注釈の記述に形式や制約がない場合、文献の注釈から人名辞典を生成するなどの計算機による注釈の分類や処理の効率が落ちる。そこで、本節では、注釈を文字情報や、単語の意味、地理情報、リンクなどに分類し、それぞれの内容に基づいて注釈を記述するためのデータモデルを定義する。さらに、本節では外部リソースとして作成する注釈について述べる。

注釈文の内容に基づくデータモデルを定義する場合、利用者による記述の容易さを優先させたモデルと計算機による情報収集の効率を優先させたモデルの定義が考えられる。一般に、前者と後者は制約に関してトレードオフの関係にあり、後者を考慮して設計した場合、注釈の記述に関する制約が強くなるため利用者による注釈の記述が減少すると考えられる [7]。そこで本稿では、一般的な辞書などの形式に基づいて直感的なデータモデルを定義することによって、利用者による記述の容易さと計算機による分析効率の両立を目指す。また、一般にデータの記述に用いる用語の統一の方が計算機による集計の効率が良くなる。しかし、遺物の名称や古文書の解釈は研究者によって記述が異なることが多いと考えられるので、本研究では用語の統一を考慮しない。ただし、ピンインなどの発音の表記や地図上の位置の記述は一般的な表記に基づく。

(1) 文字情報

文字情報では、文献中の文字に関する属性情報や文字画像への参照を記述する。文字に関する

情報の記述形式には、国文学研究資料館の外字辞書 [8] や、XML による画像参照交換方式 [9] などの提案がある。前者は読みなどの要素ごとに著者登録の記述があり、注釈で用いるには記述内容の重複が多い。後者は、文書中に埋め込む外字画像の指定方法を定義した規格であり、文字属性の記述に関する制約は弱いと考えられる。そこで本研究では、国文学研究資料館の外字辞書を簡略化した形式を考える。文字画像への参照は、ekanji や今昔文字鏡などのインターネットで公開されている大規模漢字集合を用いる。そこで、本稿では文字情報で記述のために以下のような項目を定義する。

文字情報 = (発音, 構造, 参照, 解説)

発音は、注釈を付ける文字の読みを記述する要素であり、音、訓、ピンインなどの発音の種類と合わせて記述する。構造は、文字に関する部首名、総画数などの属性情報を記述する。参照は、出典情報や文字画像への URL を記述する。解説では、文献固有の文字情報を記述する。

(2) 単語情報

本稿では、単語の意味の記述形式として一般的な辞書に基づいて下記のような項目を記述する。

単語情報 = (読み, 記述, 分類, 解説, 参照)

読みは、単語の読み方を記述する。記述は、注釈を付ける文献の単語の積文を記述する。分類は、品詞もしくは仏教用語などの単語の理解を促す情報を記述する。解説は、単語の意味やコメントなどを記述し、参照は参考文献や出典などへの URI を記述する。

(3) 地理情報

地理情報は、注釈を付ける文献画像自体の出土場所や、文献の記述にある地名に関する情報を記述する要素である。従って、地理情報の記述では、緯度や経度、標高などの位置情報の他に、住所や地名などを併せて記述することがあると考えられる。

地理情報 = (位置情報, 住所, 参照)

位置情報は、緯度、経度、標高などの数値情報を記述する。住所は、地名などのテキストデータを記述する。また、参照は、参考文献や外部リソースへの URI を記述する。

(4) 外部リソースへのリンク

注釈文として既存の Web ページを参照する場合に用いる要素であり、URI と併せて、外部リソースの分類を記述する。

リンク = (分類, 参照)

分類では、外部リソースの種類を記述する。参照は、参照する外部リソースの URI を記述する。

(5) 文字位置

この注釈は、文献データの本文要素の積文と文献画像中の文字ごとの対応を表現する注釈である。この注釈によって、切り抜き操作で、切り抜かれた文献画像に関連する積文を併せて取り出すことができる。

文字位置 = (開始位置, 終了位置)

草書体の文書には、一つの文字で二文字以上の積文を表現することがある。そこで、注釈の文字位置では、上記の通り積文の文字の開始位置と終了位置を記述する。文献画像中の1文字が積文の1文字に対応する場合は、開始位置、終了位置ともに同じ値を入れる。

その他に、外部リソースとして利用者が注釈文を生成することを考える。利用者が外部リソースとして注釈を作成した場合、歴史文献間で注釈を共有することが可能になる。このような注釈は、注釈文の他に作成者など記録するためにメタデータを併せて記述する。メタデータの記述する項目は、Dublin Core Metadata Element Set Ver. 1.1 に従う。

6 文献画像のレイアウトに基づいた積文の表示

前節で定義した注釈データモデルの応用として、文献画像のレイアウトに基づいた積文の表示について述べる。

文書には、記述位置が意味を持つことがある。例えば、荷札を考えた場合、木簡上の記述位置で荷物の送り先と送り元を表現していると考えられる。また、仏典には、仏教の思想を表すために文書を渦巻き状に記述しているものがある。従って、歴史文献の文書の2次元的な配置が文献を理解するための重要な情報になることがある。テキストデータとして歴史文献を表示する場合であっても原資料のレイアウトに併せる必要があると考えられる。また、木簡などのように墨のにじみなどで文字の判別が困難な歴史文献も多いため、文献画像の表示に併せて文献画像のレイアウトに基づいた積文の表示は、歴史文献の閲覧支援にとって有効であると考えられる。

文献画像のレイアウトに従ってテキストデータを表示する手法として、SVGを用いる方法が提案されている [10]。しかし、SVGを用いた場合、表示領域に収まりきれない画像を表示することができないため、巻物などの大きな文献画像に併せたテキストデータの表示が困難であると考えられる。

そこで本研究では、文献画像のレイアウトに従った積文の表示方法として以下の二つの方法を考える。

(1) 文献画像の文字とフォント画像の置き換え。

文献データの対応表に記述された領域と対応付けられた積文のフォントに基づいて文字画像を作成し、文献画像上に文字画像を表示する方法である。この

方法では、積文を画像として表現するため、テキストとして利用できないが文献データとして再利用が可能である。

(2) HTML 文書としての表現。

この方法では、文献データの対応表の領域情報に従って文字ごとの位置を HTML タグや CSS を用いて原資料のレイアウトを表現する。この方法では積文をテキストデータのまま表示できるので、文字列検索などの処理が可能である。また、SVG のような表示領域の制限することなしに文献の見た目の再現が可能になると考えられる。

図 6 に文献画像の見た目を再現した HTML 文書の表示例を示す。この例では、文字ごとに DIV タグを追加し表示位置と文字の大きさを定義し、元の文献の見た目を再現している。



図 6: 文献画像の見た目を再現した積文の表示例

7 考察

歴史文献に関する情報は、文字や単語、地理情報など多種多様であるため、一つのデータモデルですべての歴史文献の情報を記述することは困難である。そこで本研究で提案した文献データモデルは、まず、関連情報を集約する方法としてどのような歴史文献でも同じである歴史文献の画像と注釈の関係を記述する基本的な構造を定義した。さらに、本稿では、提案モデルの拡張として歴史文献の注釈に関するデータモデルを定義した。従って、文献データモデルは、基本モデルと注釈モデルの二つの構造を持つデータモデルである。このような構造を用いることで、歴史文献に対して柔軟な知識の集約が可能であると考えられる。また、XML の名前空間を利用すれば、注釈の編集者の手によって注釈モデルを拡張が可能となるので、長期に渡る歴史文献への知識の集約に提案モデルが利用できると考えられる。

電子スクラップブックデータモデルは利用者が収集した情報を整理するためのモデルである。また、

電子スクラップブックデータは、収集データの目録として利用できると考えられる。このような目録を電子スクラップブックシステムに依存せずに、利用者間で共有することができれば、作成されたデータの長期に渡る利用が可能であると考えられる。このような利用を実現するために、システムに依存しないデータの記述を可能にするXMLは有効なデータ記述言語であると考えられる。

8 関連研究

歴史文献の検索システムとして、奈良文化財研究所の木簡データベース [11] がある。木簡データベースでは、木簡画像に併せて木簡研究の成果に基づいたメタデータと釈文をデータベース化している。このデータベースでは、メタデータや釈文を用いて木簡画像を検索している。しかし、このようなメタデータは研究者を対象としたものであり、専門的な知識を持たない利用者は必ずしも容易に扱えるものではない。一方、本研究の提案モデルでは、利用者による注釈の編集を許すことで専門的な知識以外の知識の追加が可能になるため、考古学、歴史学分野以外の利用者による検索の効率の向上が期待できる。

木簡と注釈などの関連情報を半構造データとしてモデル化しオブジェクト指向データベースで管理する提案がある [12]。この提案システムでは、木簡の公開後に増加する注釈などを対応しており、本研究と類似点が多い。これに加えて、本研究では、データの記述にXMLを利用してインターネットを介した歴史文献と関連情報の共有と再利用を可能にしている。

9 まとめ

本稿では、木簡や古文書などの歴史文献とインターネット上に公開される関連情報の関係を利用者が編集するためのデータモデルとプロトタイプシステムについて述べた。提案モデルは、文献画像と関連する知識を注釈として集約し管理する文献データモデルと、文献間の関連を管理する電子スクラップブックデータモデルをからなる。また、提案モデルを用いた文献利用のための操作として注釈の編集、文献の切り抜きおよび検索について述べた。プロトタイプシステムでは、利用者が提案モデルの構造を知らなくとも容易に注釈の編集や文献の分類ができることを示した。さらに、本稿では、利用者が編集した注釈を効率的に利用するために、注釈で扱う主な内容の分類を示し、それぞれの特徴に併せた注釈モデルを提案した。提案する注釈モデルの利用例として、本稿では、文献画像の見た目をテキストデータで再現する方法について述べた。さらに、本稿では提案モデルを用いた長期に渡る歴史文献と注釈などの関連情報の管理について考察した。本稿で述べ

たプロトタイプシステムと注釈モデルから本稿では利用者による歴史文献に関する知識の集約と再利用に関して提案モデルの有効性を示した。

今後の課題は、提案モデルに従ったデータと地理情報などを利用した歴史文献の分類や分析の支援について考察する。

本研究の共同発表者であります勝村哲也先生には、かねてから熱心にご指導いただいておりますが、去る平成15年9月10日に急逝されました。謹んで哀悼の意を表します。

参考文献

- [1] 石川正敏, 波多野賢治, 天笠俊之, 植村俊亮, 勝村哲也: “歴史的文献画像のための電子スクラップブックシステム”, 情報処理学会論文誌: データベース, Vol.44, SIG12(TOD19), pp.110-122, 2003年9月.
- [2] Tim Bray, Jean Paoli, C. M. Sperberg-McQueen et al.: “Extensible Markup Language (XML) 1.0 (Second Edition)”, <http://www.w3.org/TR/2000/REC-xml-20001006>, 6 October 2000.
- [3] Dublin Core Metadata Initiative: “Dublin Core Metadata Element Set, Version 1.1: Reference Description”, <http://dublincore.org/documents/1999/07/02/dces/7/2>, 1999.
- [4] 勝村哲也, 丹羽正之: “eKanji”, <http://nohara.u-shimane.ac.jp/ekanji/>, 2000.
- [5] 文字鏡研究会: “今昔文字鏡”, <http://www.mojikyo.org/html/index.html>
- [6] 木簡学会: “日本古代木簡選”, 岩波書店, 1990
- [7] Marc Nanard, Jocelyne Nanard: “Cumulating and sharing end users knowledge to improve video indexing in a video digital library”, *Proceedings of the first ACM/IEEE-CS joint conference on Digital libraries*, pp. 282-289, Virginia USA, 2001
- [8] 原正一郎: “人文科学研究支援コラボレーション機能に関する実証的研究～国文学研究資料館の資料・史料を主な対象として～”, 平成12年度～平成14年度科学研究費補助金基盤研究(B)(2)展開研究成果報告書, 平成15年3月
- [9] 川俣 晶: “XMLによる画像参照交換方式”, 日本工業規格協会, JIS-TRX004 5, 2001
- [10] 師 茂樹: “SVGを用いた『華嚴一乗法界圖』の表現実験”, <http://www.l.u-tokyo.ac.jp/sat/japan/tech/experimental/svg.html>, Nov. 2000.
- [11] 奈良文化財研究所: “木簡データベース”, <http://acd.nabunken.jp/Open/mokkan/mokkan1.html>
- [12] Shinichi Ueshima, Kazuhiro Ohtsuki, Jun-ya Morishita et. al.: “Incremental Data Organization for Ancient Document Database”, *Proceeding of the Fourth International Conference on Database Systems for Advanced Applications (DASFAA '95)*, pp.457 - 466, Singapore, April, 10 - 13, 1995