

XML を利用した英日全文連携検索システムの設計と実現

桶谷猪久夫*¹ Delmer Brown*² 藤本雅彦*¹ 大久保祐子*²

*¹大阪国際大学人間科学部、 *²University of California, Berkeley

本英日全文連携検索システムの目的は、特に日本神道を中心に古代日本の文化と日本人の精神生活の研究、日本古代国家の成立史や構造の研究、民俗（民族）学的研究であり、日本文化の世界への発信と国際的なコラボレーションを促進する研究である。

本英日全文連携検索システムが直接対象とする文献は、日本の記紀である「古事記」、「日本書紀」や「続日本紀」、神祇関係の法令である「延喜式」、特定の地方誌的文書である「出雲国風土記」、中世の代表的歴史書「愚管抄」である。さらに、日本古典文献 25 巻のデジタル化、Web 上で英語と日本語（または、両言語）を利用した文献内検索と文献間連携検索、閲覧と再利用を目標にしている。

文書の実体と共にその論理構造（階層構造）や属性(XML Schema)を定義可能なマークアップ言語 XML(Extensible Markup Language)が注目され、一連の規格も制定されてきており、Web 上での文書管理・流通・提供の実質的な標準になりつつある。XML は文書構造を記述するだけでなく、我々が取り扱う各文献に出現する注釈、解題、抄録、貴重な書き込み、相互参照などの情報や知識を文書本体と独立して表現することが可能である。これらを実現するため、文書の論理構造を定義可能な XML を採用し、英日全文連携検索システムを設計・構築した。

A Design and Implementation of the Full Text Coordinated Retrieval System of Japanese Historical Resources using the XML-tagged Full-texts

Ikuo Oketani*¹ Delmer Brown*² Masahiko Fujimoto*¹ Yuko Okubo*²

*¹: Osaka International University *²: University of California, Berkeley

The goals of this full text coordinated retrieval system are to study ancient Japanese culture such as Shinto and Japanese spiritual life, to study the historical development of ancient Japanese state and its structure, and to conduct ethnological/ethnographic research. It further aims to send out the information regarding Japanese culture to the world and to promote the international collaboration.

The texts we have chosen to work with are 1) Japanese ancient chronologies such as *Kojiki*, *Nihon Shoki* and *Shoku Nihongi*, 2) *Engishiki*, a collection of laws and regulations on shrines, 3) *Izumo Fudoki*, a local document of a certain geographic area, and 4) *Gukansho*, the representative history text of the Middle Ages. We intend to digitize the twenty-five volumes of Japanese classic texts and to facilitate an internet-based coordinated search, browse and reuse between the texts using both English and Japanese.

XML (Extensible Markup Language), which could define the logic structure of the documents and attributes in addition to texts, has been receiving attention, and certain standards of the language have been set, which we are using for our standard for managing, circulating, and providing documents on the Web. XML enables us to record information about each text such as notes, titles, extracts, important insertions, and interactive references independently from the text. In order to take full advantage of these features of the language we employed XML to define the logic structure of the texts, and designed the full text coordinated retrieval system. In this paper, we explain the processes and problems we encountered in developing the full text coordinated retrieval system, describing our efforts to increase efficiency and the addition of various search functions.

1. はじめに

近年のインターネットの急速な普及によって、それを利用した電子情報の公開が一般的になってきている。これは歴史学研究分野においても例外でなく、古典文献を電子化し、研究に活用しようとする動きが盛んになりつつある。つまり、研究者（ユーザ）が必要とする情報を高速・的確・効率的に見つけ出すための情報検索システムを WWW(World Wide Web, 以下、Web という)上で提供することで、歴史学研究を援用し創造的な活動の活性化を推進することが期待される。

このような状況下で、我々は日本古典史料を対象に、その文書構造や歴史的記述方法に着目し設計した英日全文連携検索システムを開発し、インターネット上に公開することにより、歴史学研究を援用し、さらに、国際的なコラボレーションを促進することを目的とする。

我々は、本システムの設計と開発をカリフォルニア大学バークレー校のブラウン教授が運営する JHTI (Japanese Historical Text Initiative)プロジェクトと共同で開発した。このプロジェクトの目的は、日本神道を中心に古代日本の文化と日本人の精神生活の研究、その当時の事物や社会の様相を研究する資料を提供することにより、日本文化の世界への発信と国際的なコラボレーションを促進する研究である。また、外国人研究者の古典入門や研究支援だけでなく日本に関する教育にも役立つと思われる。最終的には、日本古典文献 25 巻（後述）のデジタル化とデータベース化である。

日本古典史料の日本語、英訳文とページイメージ画像ファイルに対して、連携して検索を可能にするため、それら各文書に対して文献間連携用と属性のデータ記述の定義が必要である。現在、文書の実体と共にその論理構造（階層構造）や属性(XML Schema)を定義可能なマークアップ言語 XML(Extensible Markup Language)が注目され、一連の規格も制定されてきており、Web 上での文書管理・流通・提供の実質的な標準になりつつある。XML は文書構造を記述するだけでなく、我々が取り扱う各文献に出現する注釈、解題、抄録、貴重な書き込み、相互参照などの情報や知識を文書本体と独立して表現することが可能である。

本稿では、文書の論理構造を定義可能な XML を採用し、各文献の XML 化した文書をデータベース管理システム(OpenText)に格納し、高速検索エンジン(PAT70)を利用し、UNIX システム固有のバッファであるパイプ(PIPE)を利用して検索要求とその検索結果がやり取りできるように設計した。そのため、利用者と高速検索エンジン間のユーザーインターフェースをアプリケーションプログラムで作成した。我々が設計し実現した英日全文連携検索システムの目的と概要、各種機能の特徴と問題点、今後の検討課題について述べる。

2. 英日全文連携検索システムの目的と概要

2-1. 対象文献の概要

本英日全文連携検索システムが直接対象とする文献は、日本の記紀である「古事記」、「日本書紀」や「続日本紀」、神祇関係の法令である「延喜式」、特定の地方誌的文書である「出雲国風土記」、歌集「万葉集」、天台座主の慈円の作で承久 2 年(1220 年)ごろ成立した中世の代表的歴史書（全 7 巻から構成）「愚管抄」である。さらに、後述の日本古典文献 25 巻のデジタル化、Web 上で英語と日本語（または、両言語）を利用した文献内検索と文献間連携検索、閲覧と再利用を目標にしている。

以下に、本英日全文連携検索システムが直接対象にした文献のいくつかの概要を簡単に説明する。

(1) 日本書紀⁴⁾

「日本書紀」は、官撰の国史という性格を持っており、漢文で編年体の体裁で書かれている。巻 1 と巻 2 が神代の物語、古事記上巻とほぼ同じく、天地開闢、男神の伊弉諾尊（いざなぎのみこと）と女神の伊弉冉尊（いざなみのみこと）の国生み・神生み、天照大神・素戔鳴尊の誕生などが続く。各段に「一書」と称する別伝を種々掲げているという特徴がある。巻 3 から初代神武天皇の東征から第 40 代持統天皇までの天皇の代が記され、養老 4 年(720 年)、舎人親王が合計 30 巻と系図 1 巻を撰進奏上した（「続日本紀」）。

本英日全文連携検索システムが対象とした「日本書紀」の文献は、江戸時代の儒学者で尾張藩士、河村秀根、益根父子が 60 年の月日を費やし刊行した「日本書紀」の注釈書である「書紀集解」30 巻（1756 年第一巻序）である。

(2) 続日本紀^{[6][7]}

「続日本紀」は、文武天皇元年(697)から桓武天皇の延暦十年(791)まで、九代 95 年間の歴史を記述した漢文の史書である。少なくとも 8 世紀中葉以後 2 度の編さん作業を経過した後、延暦 16 年 (797 年) に桓武天皇の命により、藤原朝臣継縄 (ふじわらのあそんつぐただ) ・菅野朝臣真道 (すがののあそんまみち) らが、ほぼ現存する形に近い 40 巻を完成させ桓武天皇に進上した。「日本書紀」に次ぐ第 2 の正史として編纂され、古代の歴史・国文学を研究する上で重要な文献であり、奈良時代の資料として信頼できる史実性を備えており、また万葉集の歌の背景を理解するのに重要な資料ともなっている。

2-2. 英日全文連携検索システムの目的

本英日全文連携検索システムの目的は、特に日本神道を中心に古代日本の文化と日本人の精神生活の研究、日本古代国家の成立史や構造の研究、民俗 (民族) 学的研究であり、日本文化の世界への発信と国際的なコラボレーションを促進する研究である。つまり、英語圏の研究者や学生の日本史・国文学の研究に貢献することであり、また、日本の研究者との共同研究を促進することで研究の相乗的な効果を追求することである。さらに、外国人研究者の古典入門や研究支援だけでなく日本に関する教育にも役立つと思われる。本研究は、Web 上で日本古典史料の文書構造を解析し、また歴史的記述方法に着目した検索手法を開発し、英語と日本語・漢文 (または、両言語) を利用した文献内検索と文献間連携検索、閲覧、再利用を目標にしている。

3. XML 文書の設計概念

3-1. XML による文書標準化の特徴

格納文献数が少量であり多彩な情報検索サービスを提供する場合、我々が既にインターネット上で公開している英日全文連携検索システムはかなり有効に作用すると思われる。しかし、格納文献数が増大し、デジタル化された情報が膨大になると単純な文字列パターンマッチング技法のみでは、その検索効率が低下するという問題がある。また、歴史史料のようなデータに一定の形式をもっていない、文書構造 (論理構造) を持つ文書から特定の項目だけを抽出することは困難であるなど、大きな制約と弱点が生じてくる。これらの問題を解決するため、文書の論理構造を定義可能なマークアップ言語 (XML: Extensible Markup Language) を採用し全文検索システムを実現した。

XML 形式はデータ交換規約として国際的に標準化された規格であり、その利用方法に関しても規格が整備されており、ツールや開発環境等の多くが公開されている。本テキストデータを XML 形式とすることにより、厳密な構造チェックがツールにて容易に可能であり、またデータ利用においても、多くのツール、開発環境が利用でき、効率よくデータ整備が行える。

XML の特徴として、(1) 任意の新しいタグや属性名の定義が可能である。(2) 章立てや段落など文書構造の情報を定義可能である。(3) 文書構造を定義する文書型定義 DTD (Document Type Definition) を設定することによって情報の操作、記述が容易である。(4) 通常使用する文書 (キャラクターベースの文書データ) により、機器及びアプリケーションから独立している。

本設計では以下のデータに関する構造定義を行う。

- (1) 日本書紀テキストデータ (本文和文、本文英文、ノート英文)
- (2) 続日本書紀テキストデータ (本文和文、本文英文、ノート英文)

データ形式は XML をベースとし、文字コードは UTF-8 コードとする。また、UTF-8 コードにて定義されていない外字に関しては、%ufxxx;形式にて定義し、外字を画像ファイルとして貼り付ける。

3-2. 文書型定義 DTD の設計

日本古典史料のような文献は文献一次情報のままであり、内容的に複雑な意味を持つ文書構造を持っている。日本書紀や続日本紀は、巻番号、序文、要約、所見、本文などから構成され、その本文はパラグラフ、年月日構造、勅撰、割注、注釈 (Note)、各項目などの論理構造から構成される。また、複数テキスト情報、テキスト画像ファイルと分割画像ファイルの連携検索を実現するために重要な物理構造としての

ページ番号やパラグラフ番号などで構成されている。

4. 英日全文連携検索システムの概要と各種機能

我々は、本英日全文連携検索システムを近年の有力な研究基盤となっている Web 上で設計し構築した。つまり、Web 上で日本語と英語（または、両言語）を利用した文献内検索と文献間連携検索、閲覧、再利用を目標に設計し実現した。英語圏と日本語圏の研究者が、歴史学研究に有効な史料検索システムを利用し、研究を進めるには、日本語文書と英訳文書が連携して、検索可能にならなければならない。そのため、4 種類の文献、つまり日本語文書、英訳文書、ローマ字読み文書、原文に近い底本の画像ファイルに対して、文書構造が定義可能な XML でタグ付けを行った。その XML タグ付けされた 4 種類の文書ファイルが連携して検索可能になる。

また、XML 化した文書に対しては、適切な文字列や語（文字）から利用者は所望する目的に合致した検索が、高速に効率的に実行できる機能が要求される。そのため、高速検索を実現するデータベース管理システム OpenText（カナダ、OpenText 社）を採用し、そこに XML 化された文書を格納した。さらに、近年のインターネットの急激な普及により、歴史学研究者も間便に利用できる Web からの検索を可能にするため、OpenText に実装されている検索エンジン PAT70 を使用し、Web ブラウザからの融通性のある検索機能を実現した。

4-1. データベース管理システム OpenText について

本英日全文連携検索システムは、XML 文書を格納するデータベース管理システムとして、研究機関等で実績のある全文検索システム OpenText（カナダ OpenText 社）を採用した。OpenText は、検索アルゴリズムとしてパトリシア (Patricia: Practical Algorithm To Retrieval Information Code In Alphanumeric) ツリー方式を採用している。パトリシアツリー方式とは、検索キーの文字列がどの半無限部分文字列と一致するかを判定して検索する方式である。具体的には、文書の全ての文字から始まる半無限部分文字列 (Semi-infinite strings) のインデックスを作成し、これを検索することで該当部分を探し出す。つまりインデックス作成時に与えられた文書に対してポインターを一文字ずつずらして置き、それぞれのポインターから文末までを半無限部分文字列とし、それをリスト化する。検索時は、検索パターンの文字列がどの半無限部分文字列と一致するかを調べる。実際の文字列はビットストリングで扱われ、0 か 1 で分岐するツリー構造で表現されるため、単語の間をスペースで区切られる英語はもちろん、2 バイトコードで膠着言語である日本語の文書に対しても高速に全文検索が可能である。

4-2. PAT70 を利用した英日全文連携検索システムの概要

格納された文献の文字列や語彙に対する検索は、全文検索システム OpenText が用意している検索機能（キーワード検索、範囲検索、複合検索）を利用して可能である。しかし、簡便な Web ブラウザを利用し、XML タグを有効に利用した、効率的で融通性のある検索を実行することが不可能である。これを実現するには、OpenText に備わっている高速テキスト検索に有効なツールであり、コマンド・モードで実行可能な高速検索エンジン PAT70 を利用する必要がある。

PAT70 はテキスト検索に有効な高速検索エンジンであり、PAT70 を使って prefixes (プレフィックス)、words (単語)、phrases (フレーズ) を検索でき、長いフレーズに対しても単語と同じくらい高速に検索可能である。また、特定したテキストの近くに出現する文字列の検索 (Proximity : 近接に基づく検索)、頻出するデータ (テキスト) の検索、頻出頻度 (回数) などの取得も可能である。

さらに、PAT70 はテキスト全体からの検索は当然として、検索エリアの限定が可能であり、テキストの一部を検索することも可能である。たとえば、ある文献中の所見、または要約だけに限定して検索を行うことができる。検索エリアの限定、つまり検索対象とするテキストのコンポーネントの選択には、各コンポーネントが何らかの始点と終点を示すタグが存在すればよい。

本英日全文連携検索システムは、インターネットの Web 上で、効率的で融通性のある検索機能を実現することを目標にした。そのためには、前述の高速検索エンジン PAT70 を Web サーバの CGI (Common

Gateway Interface)で動作させるアプリケーションプログラムを開発することが必要であった。この CGI プログラムは、Web クライアントからの動的な検索要求を受け付け、この要求を解析・処理し、高速検索エンジン PAT70 にコマンド・モードで発信する。さらに、その検索結果を受信し、利用者が見易いようフォームを作成しブラウザに表示する必要がある。これらを実現するため、CGI プログラムと PAT70 を別個のプロセスとして、動作させる必要があり、CGI プログラムは parent プロセスとして、PAT70 は child プロセスとして作動し、UNIX システム固有のバッファであるパイプ(pipe)を利用して検索要求とその検索結果の相互やり取りが可能になるように設計・開発した。

4-3. 英日全文連携検索システムの各種検索機能

我々が設計・開発した英日全文連携検索システムの各種機能について、最初に開発した江戸時代の儒学者で尾張藩士、河村秀根、益根父子が 60 年の月日を費やし刊行した「日本書紀」の注釈書である「書紀集解」を例に各種機能を説明する。

(1) キーワード検索機能

「日本書紀」の巻第三(神日本磐余彦天皇 神武天皇)から具体的な検索例で説明する。巻第三は、神武天皇の東征から書き始まり、文章の要約は、以下のとおりである。

“THE Emperor Kami Yamato Ihare-biko's personal name was Hiko-hoho-demi.,

From the date when our Heavenly ancestor descended until now it is over 1,792,470 years. But the remote regions do not yet enjoy the blessings of Imperial rule. Every town has always been allowed to have its lord, and every village its chief, who, each one for himself, makes division of territory and practises mutual aggression and conflict.

I have heard from the Ancient of the Sea, that in the East there is a fair land encircled on all sides by blue mountains. Moreover, there is there one who flew down riding in a Heavenly Rock-boat. I think that this land will undoubtedly be suitable for the extension of the Heavenly task, so that its glory should fill the universe. It is, doubtless, the center of the world. The person who flew down was, I believe, Nigi-haya-hi. Why should we not proceed thither, and make it the capital?"

All the Imperial Princess answered, and said: "The truth of this is manifest. This thought is constantly present to our minds also. Let us go thither quickly." This was the year Kinoye Tora (51st) of the Great Year.

In that year, in winter, on the Kanoto Tori day (the 5th) of the 10th month, the new moon of which was on the day Hinoto Mi, the Emperor in person led the Imperial Princes and a naval force on an expedition against the East. When he arrived at the Haya-suhi gate, there was there a fisherman who came riding in a boat. The Emperor summoned him, and then inquired of him, saying: "Who art thou?" [5]

本文は漢文であり、巻第三以降は編年体で記述されている。それを以下に示す。

自天祖降跡。以逮于今。一百七十九萬二千四百七十餘歲。而遼%u9088;之地。猶未霑於王澤。遂使邑有君。村有長。各自分疆。用相凌轢。抑又聞於塩土老翁曰。東有美地。青山四周。其中亦有乘天磐船而飛降者。余謂彼地。必當足以恢弘天業。光宅天下。蓋六合之中心乎。厥飛降者。謂是饒速日歟。何不就而都之乎。諸皇子對曰。理實灼然。我亦恆以爲念。宜早行之。是年也。太歲甲寅。其年冬十月丁巳朔辛酉。天皇。親帥諸皇子。舟師東征。至速吸之門。時有一漁人。乘艇而至。天皇招之。因問曰。汝誰也。 [4] (注) : 「東征」を下線で示す。

具体的な検索例として、上記本文(漢文)の5行目の「東征」を検索文字列(キーワード)で検索した結果である。まず、カリフォルニア大学バークレー校の JHTI (Japanese Historical Text Initiative) プロジェクトのホームページのプルダウンメニューから検索する文献を選択し、本英日全文連携検索システムを実行する。図1に示す入力フォーム画面(Interactive Searching of Nihon Shoki)の Search words: ボックスに、神武天皇が東方に美しい国ある。四方を青山が囲んでいる。天下に君臨する地であり、我が国の

中心の地であると言われ、諸皇子と舟軍を率いて東征の途に就かれたと書かれている箇所から、検索文字列（キーワード）として「東征」を入力し、検索した例を示す。もし、外字を含んだ検索文字列が必要ならば、図1の下部分の Gaiji Select ボックスをクリックすると日本書紀の文献に出現するすべての外字が別ウィンドウに表示される。この画面から該当する外字を視覚的に指定可能である。

図2は検索結果表示画面であり、検索文字列「東征」でヒットしたパラグラフ一覧が表示される。画面の上部から、検索文字列(東征)とヒットしたパラグラフ数(:3)が表示され、それ以降に該当するパラグラフ一覧が巻番号、ページ数とパラグラフ番号と共に表示される。また、指定したキーワードは、見やすくするため赤色太字で表示される。詳細表示（英日対応文書）を希望するとき、該当パラグラフの Detail ボックスをクリックすると図3に示す日本語と英語の対応パラグラフが表示される。また、表示画面の右側に、「書紀集解」の画像ファイルから切り出された該当文字列が画像で表示可能である。さらに、画面上部の[Show page image]をクリックすると、Original Image of Document の該当ページ画像が表示される。

(2) 閲覧(ブラウジング) 機能

閲覧(ブラウジング) 機能は、言語（英語、日本語、両言語対応）を選択し、文献の巻番号を指定して先頭から、またページ指定やパラグラフ番号指定で、連続して閲覧することが可能である。さらに、閲覧するパラグラフ数を指定可能である（5,10,20,30 デフォルト値：10）。閲覧(ブラウジング) 機能は、日本史・国文学を学習する

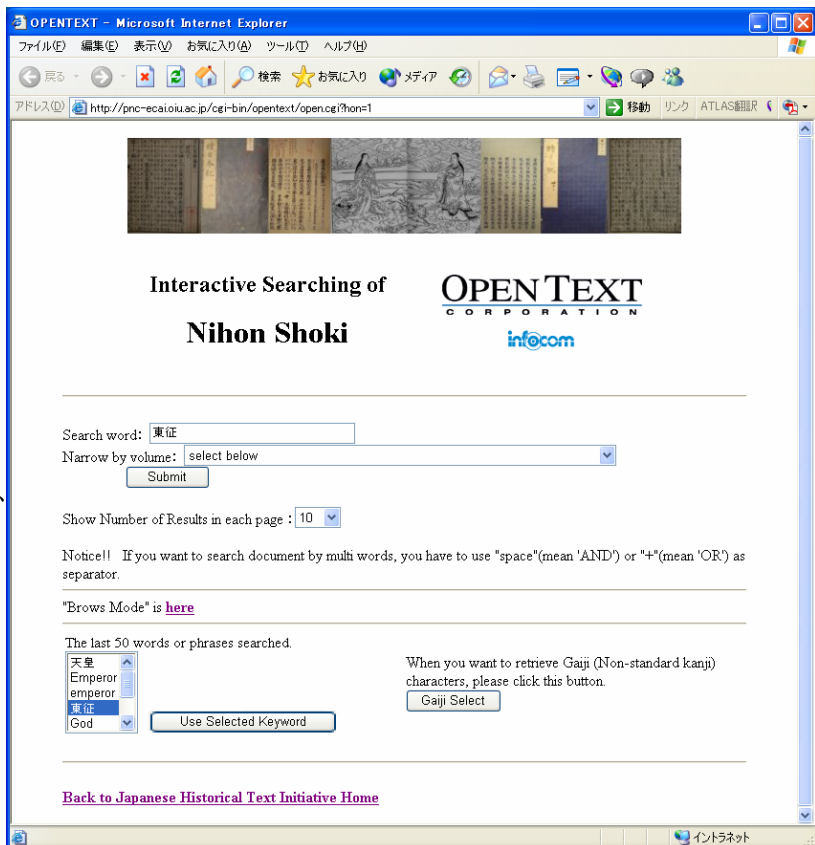


図1. 検索画面で検索文字列（キーワード）「東征」を指定した例

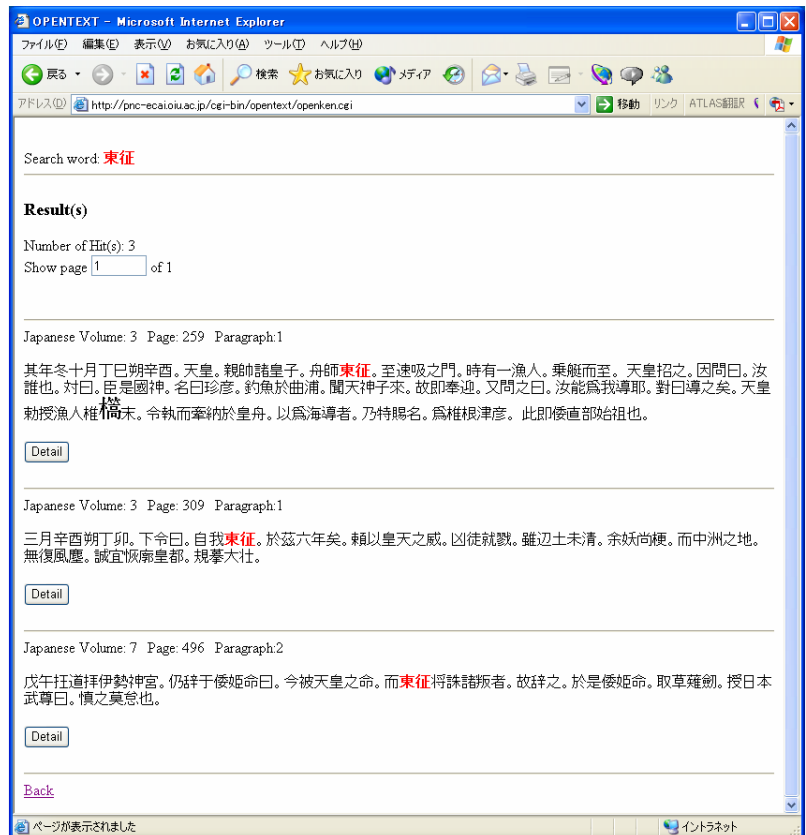


図2. 英日全文連携検索システムの検索結果画面

初心者にとって、また、それらの教育用に有効であると思われる。

(3) 項目検索機能

項目検索機能は、神の名前、神社名、神社の場所名、儀式などから文献を効率的に検索することが可能である。そのため各項目のテーブルを作成し、タグ付けをプログラムで自動的に行うと共に、関連する画像（人物像や神社など）や絵画を格納している。

その他の機能として、史料の定量的解析の実験として、巻単位、文献単位や複数文献単位に語彙検索や文字頻度分析ツールなども実装している。

本英日全文連携検索システムは、研究・教育に試験的に公開され使用されている。開発における問題点として、

①単語の表記上の違いがあった。たとえば「古事記」の英訳本では、男神伊弉諾尊は、“IZANAGI”となり、「日本書紀」の英訳本では、“Izanagi”となっている。これらの問題は、検索用プログラムで対処(解決)した。②各文献で研究者に有効な検索方法の採用が必要である。例えば、「続日本紀」は、編年体の史書の特徴として、巻番号の後に年・季・月・日が記述されている。そのため検索機能として文字列（キーワード）のみでなく年月日や指定期間内の検索も可能にした。

4. 今後の課題

本稿では、日本古典史料を題材に、インターネット上の Web を利用した英日全文連携検索システムを実現し、複数文献ファイル、つまり日本語ファイル、英訳ファイルと画像ファイルの連携と各種検索機能について述べた。本英日全文連携検索システムはインターネット環境下で、複数文献のテキスト連携表示機能、ページ画像連携表示機能、外字の混在した文字列の検索機能と外字表示機能／転送機能に対して、有効に作用し、URL : sunsite.berkeley.edu/JHTI¹⁰⁾で試験運用されている。また、今回の XML 化と同時に進めた文字コードの UTF-8 化により、「日本書紀」で、文献内に出現する外字総数：969 字数、外字種類数：305 字が、それぞれ外字総数：131 字数、外字種類数 75 字に、また、「続日本紀」で、外字総数：145 字数、外字種類数：54 字が、それぞれ外字総数：11 字数、外字種類数 7 字に大きく減少した。このことにより利用者から本英日全文連携検索システムの使用が簡便になることを期待している。

今後、表 1 に示す 25 文献の格納を計画している。我々は、これまでに英日全文連携検索システムで開

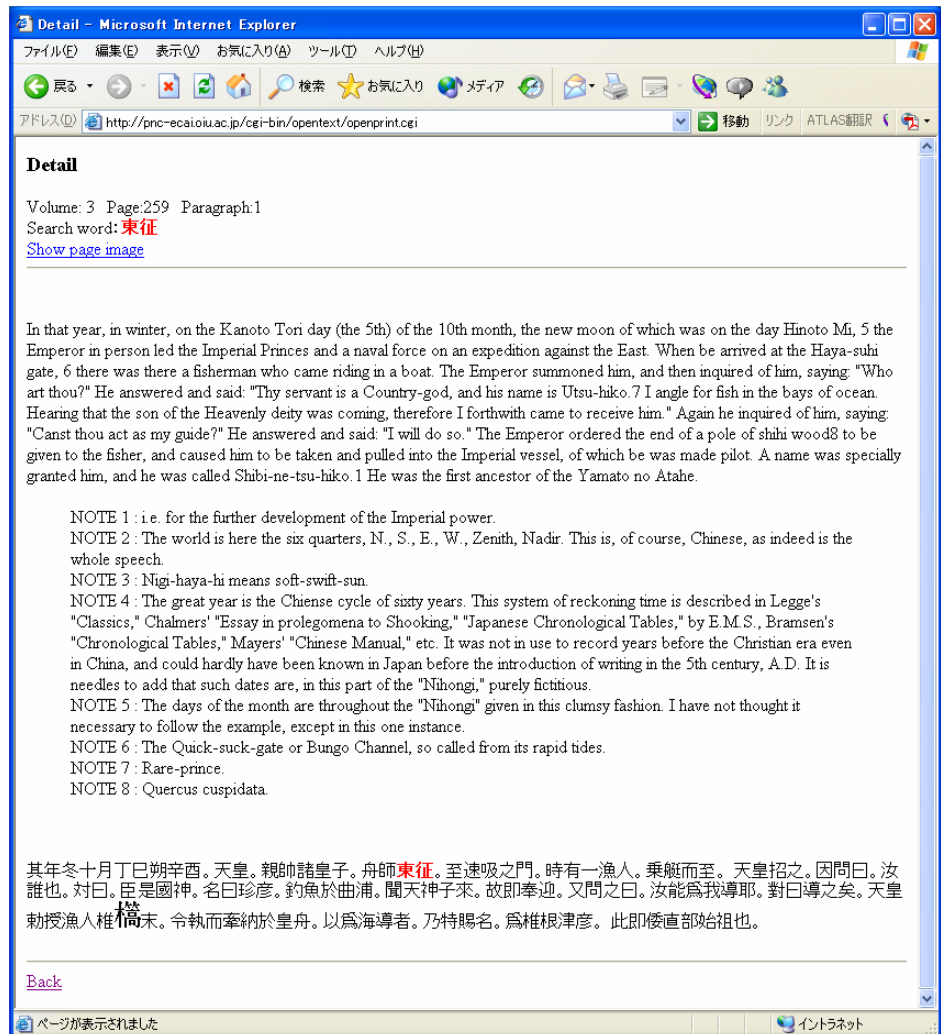


図 3. 英日対応パラグラフの表示例

発・構築した文献は、(1) 古事記、(2) 日本書紀、(3) 続日本紀、(4) 出雲風土記、(5) 延喜式、(6) 愚管抄、(7) 神皇正統記、(8) 太平記、(9) 大鏡、(10) 明治以降神社関係法令史料、である。

本英日全文連携検索システムが日本の史料の新たな解釈・解析など歴史学研究を促進し、研究支援へのコンピュータの有効性を示し、新しい視点を与え、新しい研究課題と研究方法を生み出す契機になっていくことを期待したい。

なお、本研究は科学研究費基盤研究(B)(2)「Web を利用した歴史史料英日全文連携検索システムの設計と開発に関する研究」(平成 15~17 年度、研究代表者 桶谷猪久夫) の下で行った。

表 1. デジタル化対象文献 (* : 検索可能文献、** : 現在構築中文献)

Text 1: Kojiki (古事記) *	Text 14: Dokushi Yoron (読史余論) **
Text 2: Nihon Shoki (日本書紀) *	Text 15: Meiji igo Shukyo kankei Horei (明治以降神社関係法令史料) *
Text 3: Shoku Nihongi (続日本紀) *	Text 16: Kokutai no Hongi (国体の本義) **
Text 4: Izumo Fudoki (出雲風土記) *	Text 17: Tenri-kyo (天理教)
Text 5: Kogoshui (古語拾遺)	Text 18: Kurozumi-kyo (黒住教)
Text 6: Engi Shiki (延喜式) *	Text 19: Konko-kyo (金光教)
Text 7: Eiga Monogatari (栄華物語) **	Text 20: Omoto-kyo (大本教)
Text 8: Okagami (大鏡) *	Text 21: Itto-en (一燈園) **
Text 9: Azuma Kagami (吾妻鏡)	Text 22: Tensho Kotai Jingu-kyo (天照皇太神宮教)
Text 10: Gukansho (愚管抄) *	Text 23: Rissho Kosei-kai (立正佼成会)
Text 11: Jinno Shotoki (神皇正統記) *	Text 24: Tsubaki Ookami Yashiro (椿大神社)
Text 12: Taiheiki (太平記) *	Text 25: Manyousyu (万葉集) **
Text 13: Daijingu Jin'iki (大神宮神威記)	

【参考文献】

- [1] Ikuo Oketani, Delmer Brown, Shigeo Kikuchi, Chizuko Saito “A Collaborative Research between Japan and the US on Designing and Constructing the Full Text Coordinated Retrieval System of Japanese Historical Documents using the Internet” The 2003 PNC (Pacific Neighborhood Consortium) Annual Conference and Joint Meetings, Bangkok, Thailand, Proceedings(CD-ROM), pp.1-18, November 7-9, 2003.
- [2] 桶谷猪久夫, Delmer Brown, 藤本雅彦, 大久保祐子, 『Web を利用した歴史史料英日全文連携検索システムの開発 - 日米共同研究について - 』, 情報処理学会研究報告, Vol.2003, No.107, 2003.10.24
- [3] 桶谷猪久夫, Delmer Brown, 才藤千津子, 新谷廣一, 『Web を利用した歴史史料の英日全文連携検索システムの開発』, 大阪国際大学紀要「大阪国際論叢」第 16 巻, pp.1-20, 2002.11.30,
- [4] 河村秀根・益根 『書紀集解(二)(三)(四)』, 臨川書店, 1969, stored in UCB East Asian Library.
- [5] W. G. Aston ”NIHONGI: Chronicles of Japan from the Earliest times to A.D. 697”, Printed by the Japan Society, 1896.
- [6] 補増 六国史巻三, 『続日本紀上巻』, 朝日新聞社, 1940, pp.1 - 455
- [7] 補増 六国史巻四, 『続日本紀下巻』, 朝日新聞社, 1940, pp.1 - 516
- [8] Translated and annotated by J.B. Snellen “Shoku Nihongi : Chronicles of Japan, continued, from 697-791 A.D.” (Vol.1-3) In The Transactions of the Asiatic Society of Japan. Second Series. Vol.11. 1934. Asiatic Society of Japan; 151-239
- [9] Translated and annotated by J.B. Snellen “Shoku Nihongi : Chronicles of Japan, continued, from 697-791 A.D.” (Vol.4-6) In The Transactions of the Asiatic Society of Japan. Second Series. Vol.14. 1937. Asiatic Society of Japan; 209-278
- [10] Japanese Historical Text Initiative : <http://sunsite.berkeley.edu/JTHI/>.