ヲコト点データベースの改良と書誌情報の追加

堤　智昭，山田　貴弘，萩原　泰地（東京電機大学），
田島　孝治（岐阜工業高等専門学校），
小助川　貞次（富山大学）

本稿では，訓点の一種であるヲコト点をまとめたヲコト点図を，関係する訓点資料の書誌に関する情報と共に電子化を行う．一般的に知られているヲコト点図のみならず，研究者が訓点資料を解析して作成したヲコト点図を，作成に用いられた訓点資料の情報と共に電子化しデータベース化を試みた．本研究は，訓点研究者が研究を目的に作成したヲコト点図をコンピュータ上で処理・データ共有するための環境を整備することを目的としている．そのため，これまで我々が開発を行ってきたヲコト点データベースを，新たに電子化したヲコト点に適するように改良を行った．さらに，データの共有・連携を行うため，データベースを他システムから利用可能にする WebAPI の実装を行った．

Improvement and Addition of Bibliographic Information of Wokototen Database.

Tomoaki TSUTSUMI / Takahiro Yamada /Taichi Hagiwara　(Tokyo Denki University)
Koji TAJIMA (National Institute of Technology, Gifu College)
Teiji KOSUKEGAWA(University of Toyama)

The purpose of this research is to develop an environment for processing and data sharing on a computer a Wokototenzu created by a researcher for the purpose of research. In this paper, we digitized Wokototenzu with information on the bibliographies of the relevant data. We tried digitizing the Wokototenzu created by the researcher together with the information of the subject of study used in the creation and making it into a database. Furthermore, we have improved the Wokototen database to fit the new data. And we implemented WebAPI to share and Wokototen data. It makes the database available to other systems.

１．はじめに

日本語史の研究として，訓点資料と呼ばれる古典籍を対象とした研究が行われている．訓点は漢籍，仏典，国書などの漢文本文に対して日本語で理解するための記号で奈良・平安・鎌倉時代を中心にそれ以降も広く用いられた．これまでに我々は，訓点を対象に研究を行い，ヲコト点を含む訓点資料を電子化するための構造化記述手法を提案し，ヲコト点図の電子化・データベース化を行ってきた[1][2][3][4].

本稿では，一般的に知られているヲコト点図のみならず，研究者が作成したヲコト点図など様々なヲコト点図をデータベース化しネットワークを通じて共有することを目的としたシステム開発について述べる．はじめにこれまでに電子化した主要ヲコト点 26 種以外のヲコト点を電子化しする．具体的には，築島裕[5]の点図集不載点使用文献仮名字体表・ヲコト点図に記載されたヲコト点，及び中田則夫[6]に記載された点図を電子化した．電子化したヲコト点を共有するため，データベースシステムにヲコト点の情報を入力する．その時，従来のヲコト点データベースでは，点図作成者の情報や，点図作成に使用した訓点資料の情報を関連付けて保存できないため，データベース構造の改良を行い，新しいデータベースシステムとして実装した．また，ヲコト点のデータを他システムからも利用可能とするため，データベースに搭載された情報を，HTTP プロトコルを用いて取得するためのアプリケーションインターフェースを実装した．

2.ヲコト点の概要

2.1 訓点資料とヲコト点

ヲコト点は，平安時代から鎌倉時代に広く用いられた訓点の一種である．漢字の周辺もしくは内部に付与され助詞や助動詞の代用として用いられる．ヲコト点はどの位置に書き加えられるかによってその読み方が異なる．また，加えられた点の形状も「・」「ー」「丨」等多岐にわたりその形状によっても読み方が変わる．ヲコト点は使用する流派や加点者によって様々な種類が存在し，その種類は 100 種類を超えると言われている．

©2018 Information Processing Society of Japan

2.2 ヲコト点図

構成要素の違いによって様々な種類が存在するヲコト点には，ヲコト点の「種類」ごとに「位置」,「形状」,「読み」をまとめたヲコト点図と呼ばれる資料が存在する．ヲコト点図には，図 1 に示すようにヲコト点の形状と読み，付与された位置の組み合わせから，そのヲコト点がどのような読みを持つかが記載されている．ヲコト点図に書かれている黒い四角の枠は，任意の漢字をイメージしている．一つの枠には，比較的似通った「形状」のヲコト点が複数個まとめて図示される．このまとまりを「壺」と呼ぶ．ヲコト点図は複数の「壺」から構成される．

ヲコト点図は，当時の訓点使用者が作成したもの以外にも存在し，現在も新しいものが作られている．例えば，訓点研究者が実際の訓点資料に記載されたヲコト点をまとめた点図などがそれに当たる．中田祝夫[6]と築島裕[5]は，それぞれ代表的なヲコト点 26 種をヲコト点の歴史的変遷・発達を考慮し 8 つの群（第 1 群点，第 2 群点，第 3 群点・・・第 8 群点）に分類している．このように訓点研究者が研究を目的に作成した点図はコンピュータ上でデータ処理やデータ共有できることが望ましい．

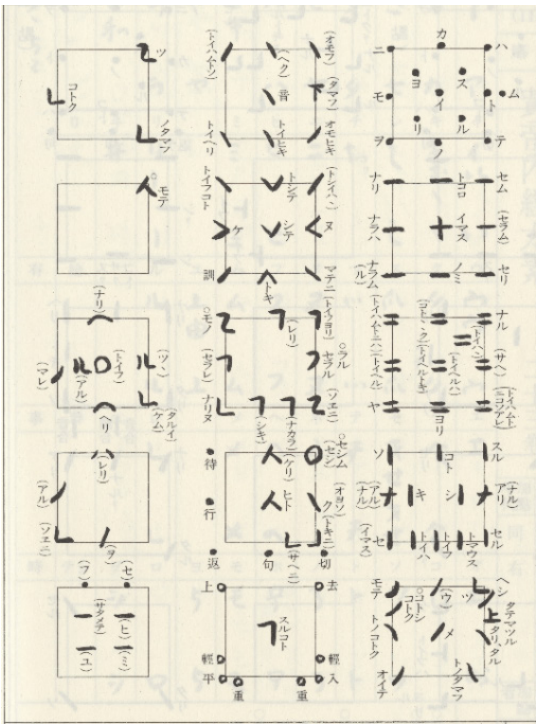


図 1 ヲコト点図の一部([6]より引用)

3 電子化するデータの特徴と電子化方針

築島裕[5]に記載されたヲコト点図は，特定の時代・学派に使用されたヲコト点をまとめたものと，特定の訓点資料に記載されたヲコト点をまとめたものの 2 種類が存在する．前者は主要ヲコト点 26 種と呼ばれ，用いていた学派ごとに喜多院点や西墓点といった名前がつけられ，まとめられている．今回電子化を行う対象となるのは後者であり，妙法蓮華経 8 巻に記載されたヲコト点をまとめた点図，のように訓点資料ごとに一つの点図としてまとめられている．築島裕[5]にはこのような点図が 192 種類記載されており，今回はそれら全てを電子化対象とした．

2 種類どちらのヲコト点図も，ヲコト点の形状と付与された位置の組み合わせから，そのヲコト点がどのような読みを持つかが記載されている．比較的似通った「形状」のヲコト点が複数個まとめて壺ごとに図示されており，壺のまとまりとしてヲコト点図が構成されている点も同様である．そのため，ヲコト点に関する電子化データは従来の研究[3]と同様のデータ構造とし，位置や形状の粒度も同様とした．

主要ヲコト点 26 種の点図と特定の訓点資料に記載されたヲコト点をまとめた点図で大きく異なる点は，ヲコト点が記載された訓点資料や，点図の作成者の情報が必要となるか否かである．そこで今回は，その点図のもととなった「訓点資料名」「巻次」「点図作成者」をヲコト点と関連付けて電子化することとした．また，同一名称の資料が複数現存しており，それぞれに記載されたヲコト点が異なることも考えられる．そのため，その訓点資料が「所蔵」された場所に関する情報も電子化することとした．さらに，実際のヲコト点図から作成した点図であるため，墨点・朱点・角点などもととなった訓点資料に記載された点の色や付け方に関する情報も電子化対象とした．

4.ヲコト点図のデータベースの構造

本データベースは，ヲコト点図データベースを元に改良し，別のデータベースシステムとして実装を行う．データベース構造は図 2 に示すように，ヲコト点に関するヲコト点テーブル，訓点資料に関する書誌情報を保存する書誌情報テーブル，点図作者に関する情報を保存する点図作成者テーブルの 3 つを持つものとした．従来のヲコト点データベースと同様に，ヲコト点図に関するテーブルにヲコト点一つを 1 レコードとして入力する．このテーブルは,「点の ID」「形状」「読み」「X 軸の位置」「Y 軸の位置」「壺数」の項目を持ち，それらにヲコト点一つの情報を保存する．それに加えて,「書誌情報 ID」と「点図作者 ID」の項目をもち，書誌情報テーブルと点図作成者テーブルそれぞれの該当するレコードを呼び出すために利用する．

書誌情報テーブルは，一つの書誌を 1 レコード

©2018 Information Processing Society of Japan

とし，「書誌情報 ID」「訓点資料名」「巻次」「所蔵」「点の色」「群」の 6 つの項目を持つこととした．「書誌情報 ID」はユニークな値とするため，入力順に 1 から数字の連番を割り振ることとした．3 章でも述べたとおり，「訓点資料名」のみでは同一の書誌が存在し，レコードが一意に特定できない可能性があるため，「所在」を入力することで「石山寺」に所蔵されている「菩提心論」のように一意に特定できるようにした．その書誌に記載されたヲコト点の色に関する情報も，書誌に関する情報の一つと考え「点の色」項目を作り，記載されていたヲコト点の色を入力することとした．

点図作成者テーブルは，作成者一人を 1 レコードとし，「点図作成者 ID」「点図作成者名」の 2 つの項目を持つこととした．「点図作成者 ID」は「書誌情報 ID」と同様に，ユニークな値とするため，入力順に 1 から数字の連番を割り振ることとした．

5. ヲコト点図データベース検索システム

本システムは図 3 に示すように，従来の Web ブラウザを用いたヲコト点の検索，閲覧システムを備えるとともに，WebAPI による外部システムとの連携も可能な構成とした．従来のヲコト点データベースとは，検索するデータベースが異なるため，別のシステムとして実装している．

5.1Web ブラウザを用いた検索・閲覧システム

Web ブラウザ用の検索システムは，[4]のヲコト点データベース検索システムをベースに検索項目と検索するデータベースを変更し，プロトタイプを試作した．4 章に示した構造のデータベースへデータを入力し，そのデータベースに適した検索項目へ変更を行っている．従来のヲコト点を条件とした検索に加えて，書誌情報による検索も可能とするため，検索項目は「読み」「形状」「位置」「訓点資料名」「所蔵」「点図作成者名」「群」の 7 つとした．

5.2WebAPI

ヲコト点データベース外のシステムと連携するため，HTTP プロトコルを用いて点図データを取得可能な WebAPI を実装した．本機能により，ヲコト点データベース以外のシステムへ，点図情報を容易に提供することができる．具体的には，訓点のついたテキストを自動で書き下すシステムなどへのデータ提供を想定している．データは，XML と JSON の 2 種類の形式で提供することとした．どちらのデータも，1 ファイルにつき 1 つの点図データを記載することとした．ヲコト点は実際のヲコト点図と同様に，壺単位で管理される構造とした．XML と JSON のどちらでデータを

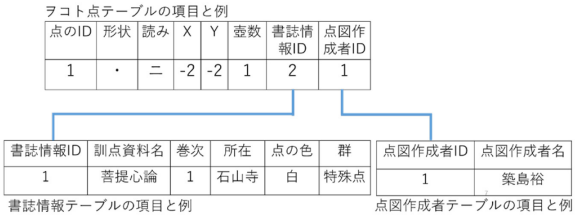


図 2　データベースの項目とデータ例

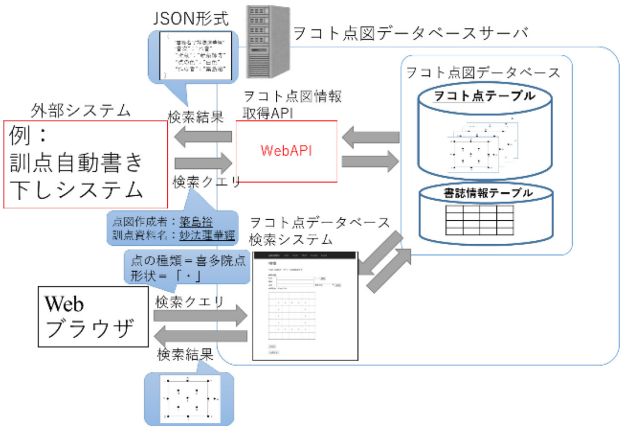


図 3　検索システムの構成

取得した場合でも，壺の配下に複数のヲコト点が存在する，という実際のヲコト点図の考え方を踏襲した設計となっている．

5.2.1 クエリパラメータ

本 WebAPI は，点図の情報を XML 形式か JSON 形式で出力できる．どちらの形式を返却するかは，ユーザが送信する Accept ヘッダの値によって決定する．具体的には，「application /json」が送られてきた場合には，JSON 形式で返却し，「application/xml」が送られてきた場合には，XML 形式で返却する．WebAPI を経由して取得できる情報は，点図データ，データベースの元になった点図の XML データ，書誌情報，データベースに搭載されている点図一覧の 4 種類である．

データベースに搭載された点図名の一覧は「/names」エンドポイントへアクセスすることで取得できる．その点図名は「/metadatas」エンドポイントへアクセスし点図に関する書誌情報を取得するときに使用できる．「/metadatas」エンドポイントはアクセス時のパラメータとして，取得する点図の名前と群を指定する「name」「gun」パラメータを持つ．返却値は「name」「gun」パラメータに入力された情報に合致する点図の書誌情報と，点図情報の取得に用いる URL である．ここで取得した URL にアクセスすると，点図データを取得するための URL と，データベースの元になった点図の XML データを取得するた

©2018 Information Processing Society of Japan



図 4　検索システム画面例

表 1　XML の要素一覧

要素名	説明
<tenzu>	ルート要素である.
<info>	メタデータ要素である．以下に点図の情報を表現するための要素と属性を持つ.
<holder>	所蔵に関する情報を記載する要素である．点図の元になった訓点資料がどこに所蔵されているかを記載する．値は文字列で入力する.
<age>	点図が加点された年代に関する情報を記載する要素である．元号も値に入るため，文字列で入力する.
<tsubo>	点図の何壺目かを表すための要素である.
<ten>	点一つを表すための要素である.

めの URL の 2 つを取得できる. 前者は「/data/{id}」エンドポイント，後者は「/data/{id}/raw」エンドポイントへアクセスするための URL となっており，{id}にはデータベースに入力されている点図一つ一つに振られた固有 ID が入力されている．このように，本 API では，点図の概要を確認してから実際の点図を見る，という手順を想定し，「/metadatas」エンドポイントから書誌情報を取

表 2　XML の属性一覧

属性名	説明
name	書誌名，点名を記載する属性である．値は文字列として入力する．<info>要素で使用する.
source	書誌の出典を記載する属性である．値は文字列として入力する．<info>要素で使用する.
gun	8 群分類のいずれに該当するかを記載する属性である．値は文字列として入力する．<info>要素で使用する.
major26	主要ヲコト点 26 種に該当するかどうかを記載する．値は Boolean として入力する．<info>要素で使用する.
index	点図の何壺目か記載するための属性である．<tsubo>要素で使用する.
shape	形状を記載するための属性である．<ten>要素で使用された場合は，そのヲコト点の形状を示す. <tsubo>要素で使用された場合は，その壺の中でもっとも頻出する形状を記載するための属性である．この項目は，点図自体が似た形状の点を壺としてまとめていることに由来して，作成した.
mean	ヲコト点の読みを記載する属性である．<ten>要素で使用する.
x	ヲコト点の位置のうち X 座標に関する値を記載する属性である．<ten>要素で使用する.
y	ヲコト点の位置のうち Y 座標に関する値を記載する属性である．<ten>要素で使用する.

得し，そこから実際の点図情報を提供する URL へ誘導するように作成した．また，固有 ID は点図データから SHA-256 によるハッシュ計算をして求めている. そのため誤植などが見つかり点図の内容が修正され書き換わった場合はアクセス ID が変わることになるが，これは同一のアクセスリソースを持つデータは中身が同一である，という設計思想に基づいている．また，それぞれのエンドポイントでは，「indent」パラメータに true

©2018 Information Processing Society of Japan

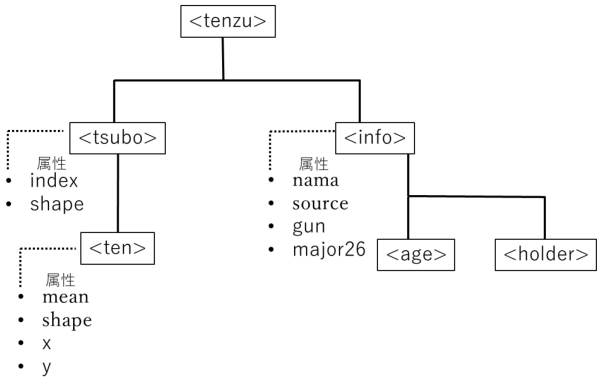


図 5　XML データ構造

を指定することで，返却値となる XML/JSON データを整形するかどうかを指定できる．

5.2.2XML データ構造

XML に記載するデータのうち，子に構造を持つものと，同じ階層に複数の同じ構造が出現するものは要素(Element)を使うこととした．それ以外の子を持たず兄弟に同じ構造がないデータは属性（Attribute）を使うこととした．XML の要素の一覧を表 1 に，属性の一覧を表 2 に示す．また構造は図 5 に示すように，<tenzu>要素をルート要素とし，その下に<info>要素と<tsubo>要素を持つ．<info>要素の下には<holder>要素と<age>要素を持つ．<tsubo>要素の下には<ten>要素を持つ．<info>要素は書誌名や点図の出典，群番号など点図に関する情報を記載する．<tsubo>要素は，ヲコト点図の 1 壺を表す要素である．その子要素である<ten>要素にはヲコト点一つの読み，形状，位置を記載する．

5.2.3　JSON データ構造

JSON データは XML のデータ構造を JSON 向けに修正する形で定義した．一つの JSON ファイルは，info オブジェクトと複数の tsubo オブジェクトを持ち，tsubo オブジェクトの下に ten オブジェクトの配列が挿入される構造とした．info オブジェクトがもつキーを表 3 に，tsubo オブジェクトが持つキーを表 4 に，ten オブジェクトがもつキーを表 5 に示す．info オブジェクトはヲコト点に関する書誌情報を入力するためのオブジェクトである．tsubo オブジェクトはヲコト点の一つの壺を表すためのオブジェクトである．この tsubo オブジェクトの子要素として，ヲコト点一つ一つを表す ten オブジェクトを配列として保持する．全ての ten オブジェクトは tsubo オブジェクトの子となるため，親である tsubo オブジェクトにおいてその壺の中で最頻出の形状を shape キーに保持することで，多くの ten オブジェクトの形状（shape キー）を省略可能とした．

表 3 info オブジェクトのキー一覧

キー名	説明
name	書誌，点図に関する名称を記載する．
holder	該当の訓点資料が所蔵されている場所の情報を記載する．
age	該当の訓点資料に点が加点された年代を記載する．
source	点図の出典に関する情報を記載する．
gun	該当の点図が 8 群分類のいずれに該当するかを記載する．
major26	該当の点図が主要ヲコト点 26 種のものかどうかを記載する．

表 4 tsubo オブジェクトのキー一覧

キー名	説明
index	点図の何壺目か記載する．値は 1 以上の整数値となる．
shape	点図の形状を記載する．tsubo オブジェクトのキーとして用いられた場合はその壺野中で最も頻出する形状を記載する．
ten	ten オブジェクトの配列である．ヲコト点一つ一つに関する情報はここに記載する．

表 5 ten オブジェクトのキー一覧

キー名	説明
mean	ヲコト点の読みを記載する．
x	ヲコト点の位置のうち X 座標に関する値を記載する．
y	ヲコト点の位置のうち Y 座標に関する値を記載する．
shape	ヲコト点の形状を記載する．この値は省略可能であり，省略された場合は tsubo オブジェクトが持つ shape の値と等しいことを示す．

6.まとめ

本稿では，実際の訓点資料から研究者が作成したヲコト点図など様々なヲコト点図をデータベース化しネットワークを通じて共有することを目的としたシステム開発について述べた．築島裕[5]の点図集不載点使用文献仮名字体表・ヲコト点図に記載されたヲコト点，及び中田則夫[6]

©2018 Information Processing Society of Japan

に記載された点図を電子化し，それらを管理するためのデータベース構造を検討した．これらの点図データは，点図のもととなった訓点資料と連携付けて保存できる構成とした．電子化したデータを Web ブラウザ上で検索できるようにするため，ヲコト点データベースを改良した検索システムを試作した．また，従来の主要ヲコト点 26 種も含め，点図データを他システムから利用できるようにするため，WebAPI によるデータ提供システムを開発した．

今後の課題として，WebAPI を利用するシステムからフィードバックを得て点図データの提供方法を改良することが挙げられる．また，研究者が作成した点図を容易に Web 上にアップロードし，共有するシステムの開発を進めていく予定である．

謝辞

本研究は JSPS 科研費 17K1850606 の助成を受けたものである．また，人間文化研究機構広領域連携基幹研究プロジェクト「異分野融合による総合書物学」の国語研ユニット「表記情報と書誌形態情報を加えた日本語歴史コーパスの精緻化」による成果の一部である．

参考文献

[1] 田島孝治，堤智昭，高田智和：ヲコト点電子化のためのデータ構造と入力支援システムの試作，人文科学とコンピュータシンポジウム「じんもんこん 2012」じんもんこん 2012 論文集，Vol.2012，pp211 -216(2012).
[2]堤智昭，田島考治，高田智和 : 点図情報入力支援ツールによるヲコト点図の電子化，人文科学とコンピュータシンポジウム「じんもんこん 2015」，じんもんこん 2015 論文集，Vol.2015，pp.185-190(2015).
[3]堤智昭，田島孝治，高田智和，小助川貞次 : コンピュータを用いた主要ヲコト点の関係性の解析，人文科学とコンピュータシンポジウム「じんもんこん 2016」，じんもんこん 2016 論文集, 2016, 139-146 (2016-12-02).
[4] 堤 智昭 ，土井 裕絵 ，田島 孝治 ，高田 智和 ，小助川 貞次 : ヲコト点データベースと検索システムの試作，人文科学とコンピュータシンポジウム「じんもんこん 2017」，じんもんこん 2017 論文集，Vol.2017，pp.253-258(2017).
[5] 築島裕：平安時代訓点本論考　ヲコト点図仮名字体表，汲古書院(1994).
[6]中田祝夫：古点本の国語学的研究,講談社(1954).

©2018 Information Processing Society of Japan