

CHISE-wiki における HNG カード画像利用の試み

守岡 知彦¹

概要：HNG の基になった「石塚漢字字体資料」のカード画像を CHISE-wiki 上で試験的に公開したのでその概要について報告する。また、京都大学人文科学研究所蔵の開成石經の拓本画像と「石塚漢字字体資料」のカード画像の統合も試みた。

1. はじめに

現在、著者は漢字字体規範データベース (Hanzi Normative Glyphs; HNG) [1] [2] の CHISE 文字オントロジー [3] [4] への統合を試みている。この HNG と CHISE の統合は、

- (a) HNG の見出し字を CHISE における UCS 統合漢字の抽象文字オブジェクトにリンクする
- (b) HNG の各資料の各字体の代表字形を CHISE における字形オブジェクトとして CHISE 文字オントロジーの包摂関係や異体字・類字関係のグラフの中に位置付ける

という2つの方法で行っている。[5]

(a) は機械的に実現可能な方法で、現在、2007 年度時点で公開されていた HNG の字体データの全てが CHISE-wiki [6] 上で公開されている。

これに対し、(b) は HNG の各代表字形が CHISE の既存の字体オブジェクトや抽象文字オブジェクト等に包摂可能かどうかや IWDS-1 [7] 的に統合漢字に包摂可能かどうかといった点について検討する必要がある、字形が不鮮明なケースも多々あるため、漢字に関する知識が必要であり手間もかかる。このため、全てを網羅することは簡単ではないため、まず初唐標準字体を代表する資料として「妙法蓮華經卷五 (今西本)」と「妙法蓮華經卷三 (守屋本)」の統合を行った。次に、開成石經の中から論語の統合作業を行うことにし、現在、52% の統合が完了している。

HNG の各資料の字体用例の代表字形は HNG の元になった「石塚漢字字体資料」の紙カードに貼られた字形用例の中から字体を代表する1つを切り出してデジタル化したものであり、字体の変遷を知るという目的からいえばこの紙カードを見る必要はあまりないといえ、オリジナルの

HNG でもこの紙カードの画像は公開されていなかった。しかしながら、HNG が同字体とみなした字形用例の中には必ずしも字形デザイン差とは言い切れないような差異やある傾向性を共有したクラスター (あるいは、クラスターに分けることが可能な弁別的特徴) を見いだすことができるケースもあり得る。また、複数人で作業していたために生じてしまった判定規準の揺れもあり得るし、別の包摂ポリシーに基づいて HNG の字体認定基準を再検討したい場合もあるかもしれない。また、HNG の代表字形が不鮮明な場合に、他の字形用例と比較検討する場合にも紙カードの画像は有用であるといえる。

こうしたこともあり、昨年 11 月に開催されたシンポジウム「字体と漢字情報」—HNG 公開 10 周年記念— の場で、石塚晴通氏は「石塚漢字字体資料」の紙カードを公開する方針を表明され、この方針に基づき、著者は CHISE-wiki 上でこの画像情報の公開を試みた。この際、京都大学人文科学研究所蔵する開成石經の拓本画像を HNG の開成石經の紙カード画像とリンクすることも併せて試みた。本報告ではこの試みの概要について説明する。

2. HNG の紙カード画像

「石塚漢字字体資料」は、漢字の字体には時代・地域 (国) による標準が存在し、その標準は時代・地域により変遷することを示すために、漢字文化圏の各時代の標準的な文献 (漢籍・仏典・国書等の典籍) を中心に、比較のため、私的な文献も交えて選定し、漢字字体の用例調査を行い、紙カードとして蓄積していたものである。

この用例調査では、選定された文献 (以下では「ソース」と呼ぶことにする) 毎に、そのソースに出現する字形を字種毎にまとめて、1つの紙カード (図 1) にまとめる形で行われていた (字種は「大字典」に基づいて整理されていた)。但し、ある字種に属する字形の数が多すぎて1枚の紙カードに張り付けられない場合は複数のカードを用いて

¹ 京都大学人文科学研究所
Institute for Research in Humanities, Kyoto University

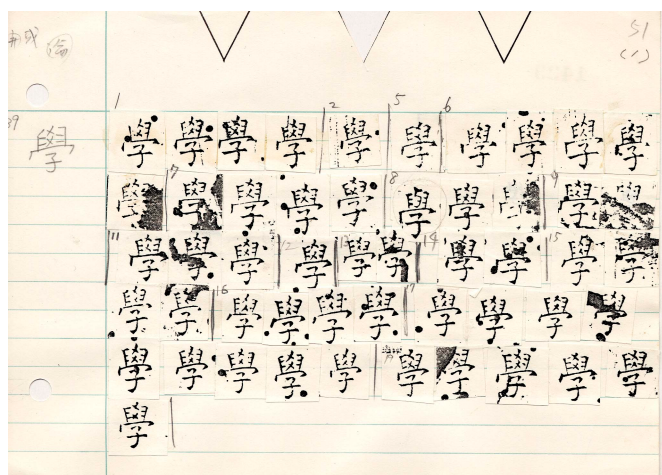


図 1 紙カードの例 (開成石經論語-學)

いる (図 2)。ある字種に複数の字体が認められる場合も 1 つ (もしくは一連の) 字種毎の紙カードにまとめられており、字体の数が紙カード上部の三角の切込み位置で表現されている。

HNG はこの紙カードの情報を電子化したものを基礎に構成されている。

HNG ではこの紙カードをスキャンしソース毎に 10 進 4 桁の番号を付した画像ファイルにしている。

字種もまたこのソース毎に 10 進 4 桁の番号で管理されている。即ち、ある字種の紙カードが 1 枚の場合はその紙カードの画像ファイルの番号が字種の番号となる。もし、ある字種の字形用例が複数の紙カードで構成されている場合は、先頭の紙カードの画像ファイルの番号がその字種の番号となる。そして、この字種の番号とソースの ID との対によってユニークな ID を構成することができる。

字体は字種の ID に対して、a, b, ... という枝番を付与することで識別している。

このように、HNG では字種や字体の ID は紙カードの番号に基づいて生成されており、字体や字種と紙カードは紐づけられている訳である。

3. 設計

従来、CHISE と HNG の統合は HNG における代表字形を CHISE の字形粒度の文字オブジェクトとして表現し、CHISE の文字オブジェクトの世界に取り込むことで実現していた。すなわち、『文字』の世界、CHISE が用いている有向グラフ型データベースエンジン **Concord** [8] の用語でいえば、character ジャンルだけを使って表現していた。

HNG の代表字形に「石塚漢字字体資料」の紙カードの画像を紐づけることを考えた場合、最も単純なのは、紙カードの画像をどこかの Web サーバーに置いておき、HNG 字形オブジェクトに対応する紙カードの画像をリンクすることである。

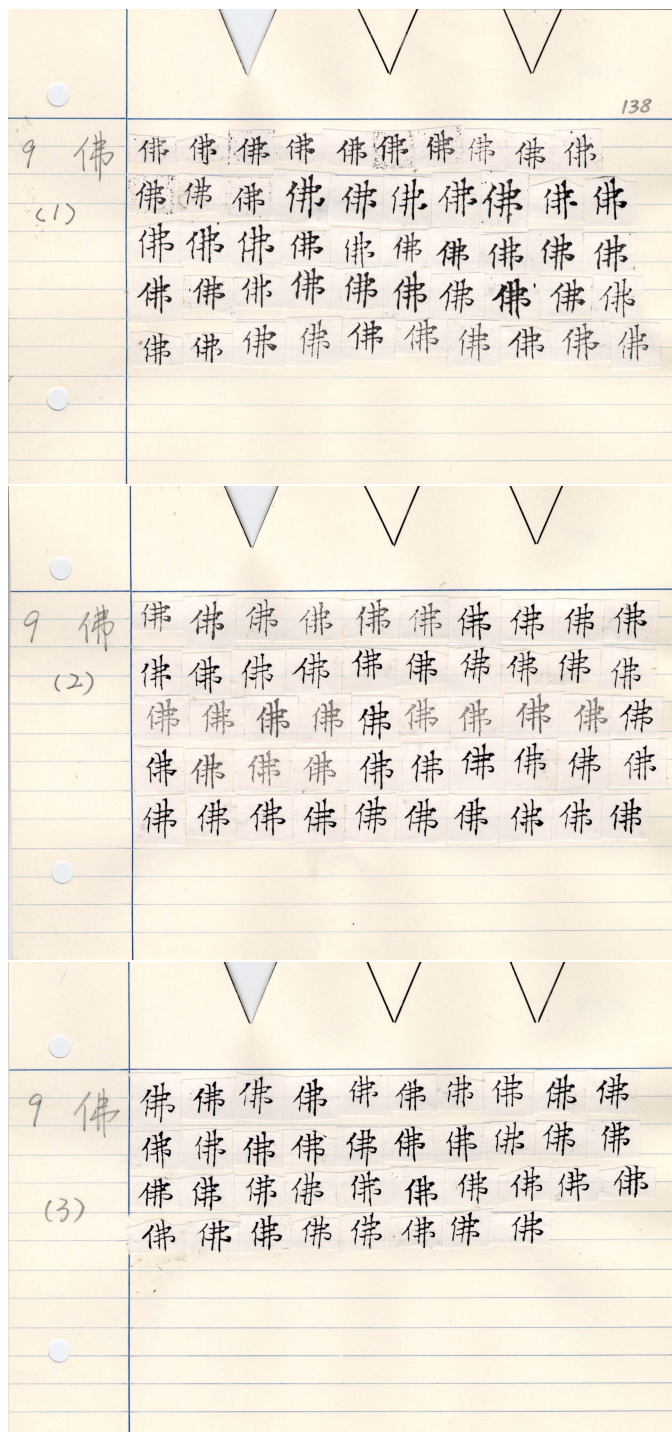


図 2 複数の紙カードの例 (守屋本妙法蓮華經卷三-佛)

しかしながら、紙カードに対してメタデータを記述したり、なんらかの言及 (アノテーション) を行ったり、別の何かと関連づけようとする場合、紙カードを単なる画像ファイルとするだけでは不十分であり、Linked Data におけるノードとして扱うのが望ましいといえる。つまり、紙カードの画像とは別に抽象的な概念ないしはものとしての紙カードを示すオブジェクトを想定するのが望ましい。

よって、HNG 字形オブジェクトと HNG 紙カードオブジェクトと HNG 紙カード画像オブジェクトという 3 種類

のオブジェクトを設け、HNG 字形オブジェクトと HNG 紙カードオブジェクト、HNG 紙カードオブジェクトと HNG 紙カード画像オブジェクトをリンクすることにする。

4. 実装

4.1 画像リソースと画像配信サービス

一般に、ある画像ファイルを Web 上で配信する方法は複数ありうる。同じファイルを複数の異なるサーバーに置く場合もありうるし、画像形式を変換する必要がある場合もある。また、ファイルサイズの大きな高精細画像を配信する場合、複数解像度タイル画像として扱った方が良くかもしれない。現在、デジタル人文学の世界を中心として、このような画像配信用 Web API の仕様として IIIF Image API [9] が普及しつつあり、相互接続性を考慮した場合、このような標準的な API を用いることも考慮すべきであるといえる。しかしながら、Web API は Web の世界の変容とともに変わっていく可能性も大いにありうるし、新たな画像配信の仕組みや形式に対応する必要に迫られる場合もあるかもしれない。

こうしたことを鑑み、画像ファイルとその配信サービスを区別したモデル化を行い、抽象的な画像ファイルを示すオブジェクトの属性としてその内容を配信可能なサービスを列挙するという方法を採用することにした。

4.1.1 IIPImage server

画像配信サービスの実装としては、JPEG ファイルをそのまま配信する方法に加えて、IIIF Image API をサポートするために IIPImage [10] server Version 1.0 を採用した。IIPImage server は、その名の通り、IIP (Internet Imaging Protocol) に基づいて画像を配信するサーバーである。IIPImage server は IIP の独自拡張の一種として IIIF Image API をサポートしており、1 つのサーバーで IIP と IIIF Image API の両方をサポートすることができる。^{*1} よって、IIIF Image API だけでなく IIP もサポートすることにした。

画像配信用サーバーの OS には Debian GNU/Linux 8 “jessie” の AMD64 版を用いたが、jessie の IIPImage server は 0.9.9 であったため、次期リリースバージョンのためのテスト中のパッケージ群 (testing) からパッケージを借りることにした。これは次の手順で行った：

4.1.1.1 ターゲットリリースの既定値の指定

```
APT::Default-Release "jessie";
```

という行を含んだ `/etc/apt/apt.conf.d/99target` を作成する。

4.1.1.2 source.list に testing を追加

```
/etc/apt/source.list に  
deb http://ftp.jp.debian.org/debian \
```

```
testing main contrib non-free  
deb-src http://ftp.jp.debian.org/debian \  
testing main contrib non-free
```

の2行を追加する。

4.1.1.3 apt-get update の実行

4.1.1.4 iipimage-server のインストール

```
% sudo apt-get install iipimage-server/testing
```

後は利用する Web サーバーでの設定を行えば利用可能となる。Apache を用いる場合、各種設定は `/etc/apache2/mods-enabled/iipsrv.conf` で行うことができる。

4.1.1.5 iipsrv.conf の設定

```
<Directory "/iipsrv/"> となっていたら、  
<Directory "/usr/lib/iipimage-server/"> と直してやる必要がある。
```

また、`FILESYSTEM_PREFIX` を必ず設定すること。例：

```
FcgidInitialEnv FILESYSTEM_PREFIX \  
"/opt/iipimage"
```

この例の場合、IIPImage server で用いる配信用タイル画像ファイルは `/opt/iipimage` 以下に置くことになる。もし、この環境変数を設定しない場合、全ファイルシステムがアクセス可能になるのでセキュリティ上問題があるといえる。

また、第三者にサービスを公開したい場合は環境変数 `CORS` を設定すること。

4.1.1.6 複数解像度タイル画像の作成

[11] を参考に、VIPS [12] を利用して、IIPImage server で用いる Tiled Multi-Resolution TIFF (複数解像度タイル画像) を作成した。

4.1.1.7 アクセス用 URL

画像配信用サーバーのホスト名を `hng.chise.org` にしたため、IIPImage server のエンドポイントは

```
http://hng.chise.org/iipsrv/iipsrv.fcgi
```

となる。この後に IIP のコマンドを付けることで各種サービスが利用可能となる。

IIIF Image API は IIIF というコマンドを用いることで実現されており、

```
http://hng.chise.org/iipsrv/iipsrv.fcgi?IIIF=
```

以下に IIIF Image API を記述することで利用可能となるが、Linked Data 等での利用の便宜を計るため、

```
http://hng.chise.org/images/iiif/
```

に IIIF Image API 用のエンドポイントを設けることにした。

4.1.2 画像リソースオブジェクト

ある画像ファイルを抽象化したものとして「画像リソース」オブジェクトを設けた。これは Concord の `image-resource` ジャンルのオブジェクトとして実現している。

^{*1} IIP と IIIF Image API を混在して使用することもできる。

このオブジェクトは
=id 素性 ID を示す。シンボル型。必須。
=location 素性 画像資源の識別子 (URL) を示す。文字列型。必須。
=location@iif 素性 IIF Image API での URL を示す。文字列型。オプション。
=location@iip 素性 IIP での URL を示す。文字列型。オプション。
image-width 素性 画像の幅 (ピクセル数) を示す。整数型。必須。
image-height 素性 画像の高さ (ピクセル数) を示す。整数型。必須。
name 素性 画像資源の名前を示す。文字列型。必須。
<-image-resource 素性 この画像を表示とするオブジェクトを示す。オブジェクトのリスト型。オプション。といった素性を持つ。

4.1.3 HNG カードオブジェクト

「石塚漢字字体資料」の紙カードを表現したものとして「HNG カード」オブジェクトを設けた。これは Concord の hng-card ジャンルのオブジェクトとして実現している。

このオブジェクトは
=id 素性 ID を示す。シンボル型。必須。
=hng-id 素性 ID を示す。シンボル型。必須。
name 素性 紙カードの名前を示す。文字列型。必須。
representative-glyph 素性 対応する HNG 代表字形オブジェクトを示す。オブジェクトのリスト型。オプション。
->image-resource 素性 紙カードを写した写真の画像リソースを示す。オブジェクトのリスト型。必須。といった素性を持つ。

=id 素性と =hng-id 素性の値は 18-1234 のようにソース番号とカード番号の 10 進数を「-」で繋げたものを用いた。

4.1.4 HNG 代表字形オブジェクト

従来から CHISE 文字オントロジー中に設けていた HNG 代表字形オブジェクト (Concord の character ジャンルのオブジェクト (文字オブジェクト)) に

sources@HNG/card 素性 対応する HNG カードオブジェクトを示す。オブジェクトのリスト型。
を追加した。

HNG 代表字形の ID と紙カードの ID は元々同じものであり、機械的に変換可能であるので、この情報は冗長であるが、CHISE-wiki (EgT[13]) は Concord 中に存在する Linked Data のグラフに書かれたことしか知らないため、特別な処理を行わなければ紙カードの画像を表示したりリンクすることができない。しかしながら、このような冗長な情報をあらかじめ静的に生成しておくことはメンテナンスコストを上げるため、HNG 代表字形オブジェク

トの CHISE-wiki 頁 (例: <http://www.chise.org/est/view/character/rep1.hng-kar=12340>) にアクセスした際に動的に生成するようにしている。

5. 拓本文字データベースの統合

HNG はさまざまな字形用例のデータを収録しているが、「石塚漢字字体資料」の紙カードをベースにしているという技術的な理由と、利用許諾の問題という2つの制約のために、全文画像を表示することができなかった。

一方、「拓本文字データベース」では京都大学人文科学研究所所蔵の拓本画像に座標データ付の文字情報が埋め込まれており、字形用例から元の拓本画像を表示可能である。しかしながら、拓本文字データベースでは字形用例の整理がなされておらず、HNG のように字体毎に整理して表示したり分析することができない。

こうしたことから、両者を繋げることで HNG に対して全文画像表示機能を提供するとともに、拓本文字データベースに対して HNG へのマッピングを提供することを目指す。

こうしたことを実現するためには HNG と拓本文字データベースの両方で共通したソースが存在しなければならない。そして、それに該当する可能性があるのが開成石經の孝經と論語、周易であったので、これらを対象に検討を行った。

HNG と「拓本文字データベース」は共に開成石經のデータを持っているが、元になっている拓本は異なっており、前者が東洋文庫所蔵の拓本を用いているのに対し、後者は京都大学人文科学研究所所蔵の拓本を用いている。よって、厳密に言えば、拓本上の字形は同一とはいえないが、同じものの拓本には違いがなく、そう大きな差異はないであろうと仮定してこの両者を対応づけることにした。

5.1 拓本文字データの表現

5.1.1 元データ

拓本文字データベースでは各拓本に含まれる文字の情報が画像マークアップされている。拓本中の各文字は拓本画像の座標を使ってその領域が示されるとともに、その文字の UCS での符号位置が記されている。

5.1.2 拓本画像の画像リソースと配信サービス

拓本画像の配信サービスとしては、紙カード画像と同様に、IIPImage server を用いることにした。そして、画像リソースオブジェクトによって拓本画像を表現することにした。但し、拓本画像では、=location@iif 素性と =location@iip 素性に加えて、=location@djvu 素性を用いて拓本文字データベースでの URL も示すようにした。

また、5.1.3 節で述べる <-image-segment 素性の逆関係素性 ->image-segment と、5.1.4 節で述

べる `<-segmented-glyph-image` 素性の逆関係素性 `->segmented-glyph-image` が生成される。前者はある拓本画像に含まれる部分画像（文字画像）の一覧を示し、後者は出現字形の一覧を示す。

図 3 に画像リソースオブジェクトの E_gT での表示例を示す。ここでは `->segmented-glyph-image` 素性の値である出現字形オブジェクトをそれに対応する文字として表示しているため、ブラウザの検索機能を使ってテキストに含まれる文字列を検索することができる。この例では `->segmented-glyph-image` 素性の値が省略されているが、「…」をクリックすれば省略されている残りの部分も表示できる。

また、画像リソースオブジェクトが `=location@iiif` 素性を持っている場合、E_gT の頁上部の主題表示部に OpenSeadragon [14] を使って `=location@iiif` 素性の値で示される画像を表示するようにした。このため、画像をドラッグして表示位置を変えたり、OpenSeadragon のアイコンを押して拡大縮小したり全画面表示することも可能である。

5.1.3 部分画像オブジェクト

出現字形の画像を表現するために、元となる画像の中の矩形領域で示される部分画像を表現するための Concord オブジェクトを設けた。これを「部分画像オブジェクト」と呼ぶことにする。

部分画像オブジェクトは、HNG の紙カード画像や拓本画像と同様に、`image-resource` ジャンルのオブジェクトとするが、全体画像の場合に加えて、

`image-offset-x` 素性 全体画像中での部分領域の始点の X 座標を示す。整数型。必須。

`image-offset-y` 素性 全体画像中での部分領域の始点の Y 座標を示す。整数型。必須。

`<-image-segment` 素性 この部分画像オブジェクトに対応する全体画像オブジェクトを示す。オブジェクトのリスト型。必須。

`<-image-resource` 素性 この画像を表示とするオブジェクトを示す。オブジェクトのリスト型。オプション。

`=location@djvuchar` 素性 拓本文字データベースでの部分画像の URL. 文字列型。必須。

という素性を付与することにした。

図 4 に部分画像オブジェクトの E_gT での表示例を示す。ここで、先頭に表示されている部分画像をクリックすると全体画像での出現位置にジャンプすることができる。また、`<-image-resource` 素性の値部をクリックすると、5.1.4 節で述べる出現字形オブジェクトに飛ぶことができる。

5.1.4 出現字形オブジェクト

拓本中の各文字（字形）を表現するために、出現字形オブジェクトを設ける。これは HNG 字形オブジェクトと

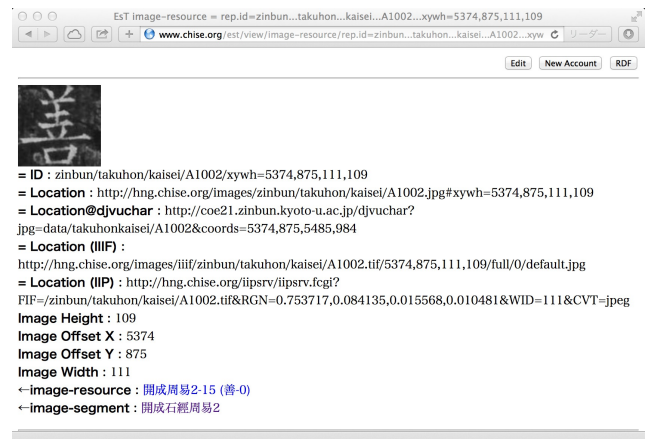


図 4 部分画像の E_gT での表示例

は異なり、文字オブジェクト（`character` ジャンルのオブジェクト）とはせず、新たに `glyph-image` ジャンルの Concord オブジェクトを設けることにした。このオブジェクトは

`=id` 素性 ID を示す。シンボル型。必須。

`character` 素性 出現字形に対応する UCS の抽象文字を示す。オブジェクトのリスト型。オプション。

`name` 素性 出現字形の名前を示す。文字列型。オプション。

`<-glyph-image@zinbun/takuhon` 素性 対応する HNG カードオブジェクトを示す。オブジェクトのリスト型。オプション。

`<-segmented-glyph-image` 素性 この字形が出現した拓本画像リソースオブジェクトを示す。オブジェクトのリスト型。必須。

`->image-resource` 素性 この出現字形の画像リソースを示す。オブジェクトのリスト型。必須。

といった素性を持つ。



図 5 出現字形の E_gT での表示例

図 5 に出現字形オブジェクトの E_gT での表示例を示す。ここでは、`character` 素性に「→ UCS 抽象文字」、`<-glyph-image@zinbun/takuhon` 素性に「→ HNG カード」という素性名の表示を設定している。また、部分画像オブジェクトの場合と同様に、先頭に表示されている字形画像をクリックすると全体画像での出現位置にジャンプす

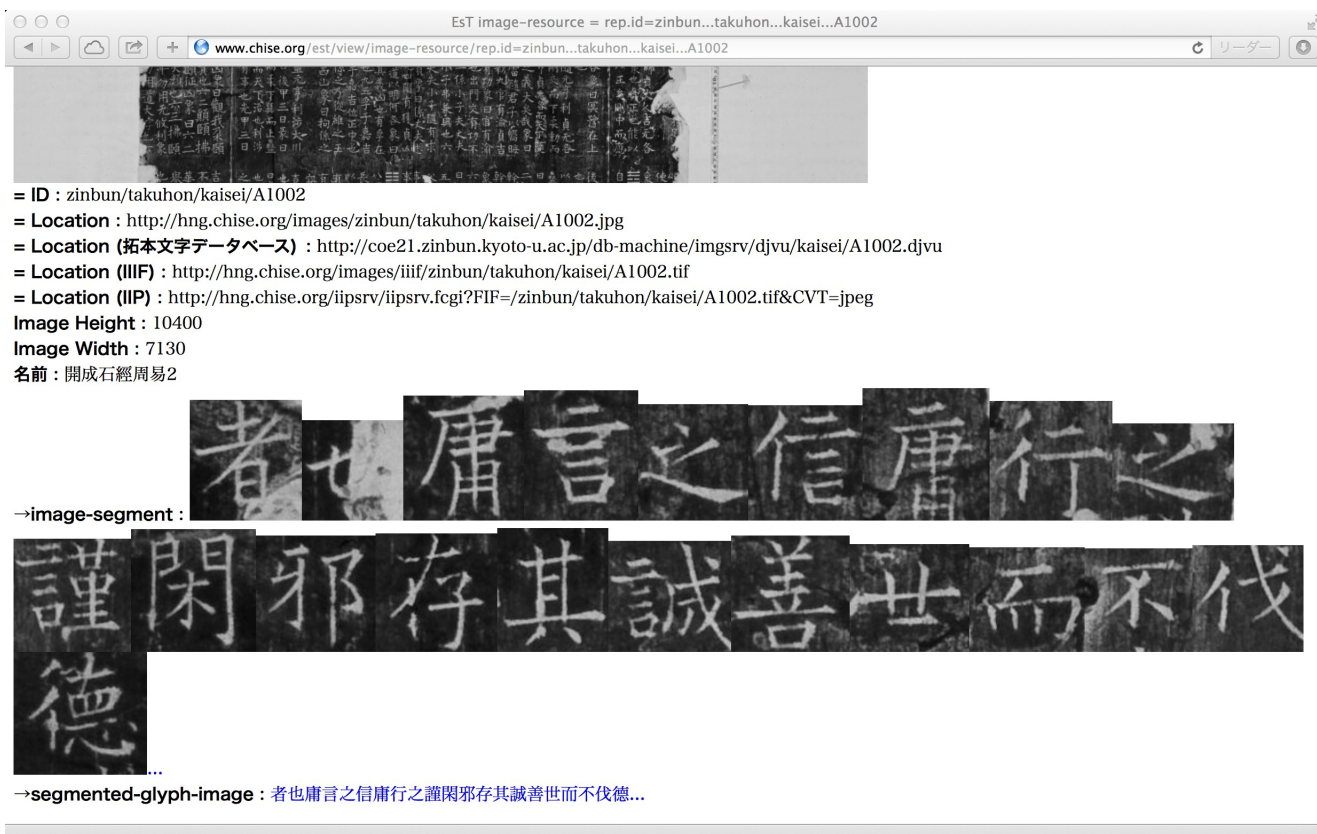


図 3 拓本画像リソースの ESt での表示例 (開成石經周易 2)

ることができる。

5.2 HNG カードオブジェクトを用いた情報統合

拓本文字データベースの出現字形の情報と「石塚漢字字体資料」の字形用例の情報を統合する方法として、ある紙カードの字種に相当する拓本文字データベースの出現字形オブジェクトの集合をその HNG カードオブジェクトの中の素性として表現する方法を採ることにした。そして、このための素性として、5.1.4 節で述べた <-glyph-image@zibun/takuhon 素性の逆関係素性である ->glyph-image@zibun/takuhon 素性を用いることにした。

これは、要するに、拓本文字データベースの開成石經の出現字形を「石塚漢字字体資料」と同様な方法で整理して仮想的な紙カードを作り、それを元々の「石塚漢字字体資料」の紙カードと併置するようなものだといえる。これによって、HNG カード画像 (東洋文庫所蔵の拓本字形) と拓本文字データベースの出現字形 (京大人文研所蔵の拓本字形) を見比べることができる訳である (図 6)。

5.3 HNG 代表字形オブジェクトでの表示

4.1.4 節で述べたように、HNG カードオブジェクトをサポートするために、HNG 代表字形オブジェクトに sources@HNG/card 素性を追加したが、拓本文字データ



図 6 HNG カードの ESt での表示例

ベースの出現字形の情報を統合するために sources@zibun/takuhon 素性 対応する拓本出現字形 オブジェクトを示す。オブジェクトのリスト型。を追加した。

sources@HNG/card 素性と同様に、この素性も HNG 代表字形オブジェクトの CHISE-wiki 頁にアクセスした際に動的に生成 (HNG カードオブジェクトの ->glyph-image@zibun/takuhon 素性の値をコピー) す

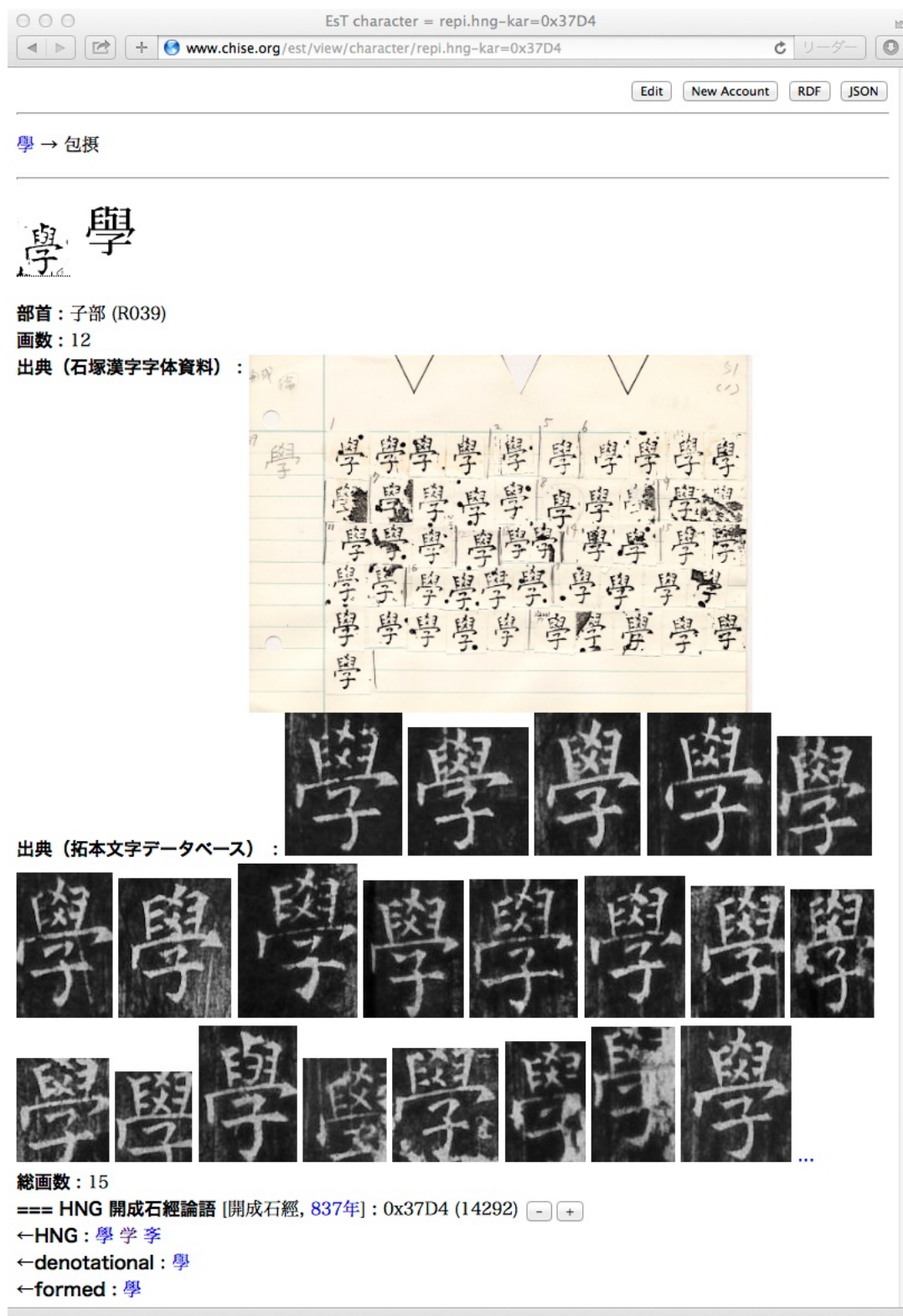


図 7 HNG 代表字形の CHISE-wiki (EgT) での表示例

るようにしている。

図 7 に HNG 代表字形の CHISE-wiki (EgT) での表示例を示す。ここでは、sources@HNG/card 素性に「出典 (石塚漢字字体資料)」、sources@zinbun/takuhon 素性に「出典 (拓本文字データベース)」という素性名の表示を設

定している。

この例では sources@zinbun/takuhon 素性の値としてこの字形とは異なる出現字形が含まれていることが判るが、これは拓本文字データベースでは出現字形の字体への分類作業が行われておらず、単純に UCS 統合漢字の抽象

文字に対応づけているだけだからである。^{*2}

HNG カードオブジェクトを用いた情報統合の場合、「石塚漢字字体資料」の紙カードは字種単位にまとめられているので、拓本文字データベースの出現字形を UCS 統合漢字の抽象文字によって統合しても問題が生じなかったが、HNG 代表字形オブジェクトに統合する場合、このような気持ち悪い事態に遭遇し得る訳である。

しかしながら、開成石経では 1 つの字種において異体字が使われる例は極めて少なく、実際にはこのようなケースはほとんど生じないといえる。

むしろ、現実的に HNG と拓本文字データベースの情報統合で問題となるのは、HNG の見出し字と拓本文字データベースの UCS のコードポイントが一致していないケースである。HNG では「大字典」に基づいた見出し字以外に幾つかの異体字も含めているが、必ずしも網羅的でないために、UCS が複数の異体字を収録している場合に食い違うことがある。また、拓本文字データベースの作成過程において既存の電子テキストを利用して画像マークアップしたために、開成石経での表現とは異なる異本の文字が埋め込まれてしまったケースもある。また、入力ミスと思われるものもあり、異体字処理だけでは解決できないといえる。

こうした不一致のケースはこれまでの所、11 件見つかっているが、こうした場合、拓本文字データベースのテキスト情報を利用して、異体字情報や論語の異同情報、前後関係の情報等を利用することで対応する出現字形を見つけることができた。また、こうしたものが見つかった場合、HNG 代表字形オブジェクトの <-HNG@zinbun/takuhon 素性の値に拓本文字データベースでの UCS の抽象文字を入れることで両者の対応関係を表現するようにしている。^{*3}

6. おわりに

CHISE IDS 漢字検索や CHISE-wiki 等の CHISE の Web サービスの枠組で「石塚漢字字体資料」のカード画像を提供する試みについて概説した。また、拓本文字データベースで公開されている京都大学人文科学研究所所蔵の開成石経の拓本画像と「石塚漢字字体資料」のカード画像の情報統合も試みた。

「石塚漢字字体資料」のカード画像を HNG 代表字形に繋げて公開したことにより、HNG の豊富な字体用例に対して、その元になった字形を確認することができるようになった。このことは漢字字体の歴史的変遷の理解だけでなく、HNG やその元となった「石塚漢字字体資料」のあり方やそのモデルや思想を理解する上でも有用であると考

えられる。また、このカード画像と京都大学人文科学研究所所蔵の開成石経の拓本画像の字形情報を繋げたことにより、同じ開成石経の異なる拓本の字形を比較できるようになり、また、開成石経の拓本の全文画像をすぐに確認できるようになった。

このシステムでは画像配信用プロトコルとして IIIF Image API の利用を試みた。なお、IIIF や IIPImage server に関しては永崎研宣氏の一連のブログ記事（特に [15]）が参考になった。

最後に、今回の実験を行う機会を与えて頂いた、石塚晴通先生と高田智和先生に感謝する。なお、本論文における誤りや誤解は全て私の責任であることはいうまでもない。

参考文献

- [1] 石塚晴通, 池田証寿, 岡崎裕剛: 漢字字体規範データベースとその応用, 東洋学へのコンピューター利用第 17 回研究セミナー, pp. 53–63 (2006).
- [2] 高田智和: 漢字字体と典籍の性格との関係 — 「漢字字体規範データベース」が主張するもの —, 情処研報, Vol. 2013-CH-97, No. 12, pp. 1–4 (2013).
- [3] 守岡知彦: 文字オントロジーに基づく文字処理について, 情処研報, Vol. 2006, No. 112, pp. 25–32 (2006). 2006-CH-72.
- [4] Morioka, T.: Multiple-policy Character Annotation based on CHISE, *Journal of the Japanese Association for Digital Humanities*, Vol. 1, No. 1, pp. 86–106 (2015).
- [5] 守岡知彦: 多粒度漢字構造モデルに基づく字形整理の試み — 漢字字体規範史データベースの CHISE への収録を通じて —, じんもんこん 2015 論文集, 情報処理学会, 情報処理学会, pp. 1–8 (2015).
- [6] 守岡知彦: CHISE のセマンティック Wiki 化の試み, 情処研報, Vol. 2010-CH-87, No. 8, pp. 1–8 (2010).
- [7] : IRG Working Document Series, <http://appsrv.cse.cuhk.edu.hk/~irg/irgws.html>.
- [8] 守岡知彦: Concord: プロトタイプ方式のオブジェクト指向データベースの試み, Linux Conference 抄録集, Vol. 4 (2006).
- [9] Appleby, M., Crane, T., Sanderson, R., Stroop, J. and Warner, S.: IIIF Image API 2.1, <http://iiif.io/api/image/2.1/> (2016). Version 2.1 (Crowned Eagle).
- [10] : IIPImage, [http://iipimage.sourceforge.net/\(2000-2015\)](http://iipimage.sourceforge.net/(2000-2015)).
- [11] : Images (of IIPImage), <http://iipimage.sourceforge.net/documentation/images/>.
- [12] : VIPS, <http://www.vips.ecs.soton.ac.uk/index.php?title=VIPS>.
- [13] 守岡知彦: Wiki 的手法に基づく構造化データの編集について, 人文科学とコンピュータシンポジウム論文集 — 人文工学の可能性 — 異分野融合による「実質化」の方法 —, 情報処理学会シンポジウムシリーズ, Vol. 2010, No. 15, 情報処理学会, 情報処理学会, pp. 33–40 (2010).
- [14] : OpenSeadragon, <https://openseadragon.github.io/>.
- [15] 永崎 研 宣: IIIF を使ってみよう! 人のための IIPImage Server インストール記 (簡易版), <http://digitalnagasaki.hatenablog.com/entry/2016/04/21/214423> (2016).

^{*2} 実際にはさらに異体字をどこかに寄せている場合がある。

^{*3} これにより、拓本文字データベースでの UCS の抽象文字の ->HNG@zinbun/takuhon 素性の値にこの HNG 代表字形オブジェクトが入り、HNG と拓本文字データベースの両方の見出し文字から HNG 代表字形オブジェクトに到達することができるようになる。