GlyphWiki による観智院本『類聚名義抄』字形の再現とその改善

池田　証壽（北海道大学）
李　媛（関西大学）
劉　冠偉（東京大学）
鄭　門鎬（北海道大学）

概要：平安時代漢字字書総合データベース（HDIC）プロジェクトでは，観智院本『類聚名義抄』（以下，名義抄）の全文テキストデータベースを 2022 年 3 月に公開した．名義抄は，平安時代の日本語の重要資料であり，その価値は，多数の和訓を収録し日本語の語彙史の資料となること，和訓と漢字の多くに声点が施され音韻史の資料となること，漢字には「正」「俗」等の字体注記が豊富で文字史の資料となることにある．名義抄は古写本として伝存するが，その活字翻刻本はいまだ存在しない．これは，名義抄に難字が極めて多いことによるが，Unicode と GlyphWiki とを併用すれば，必要な字形を作字することが可能である．ただ，手書きの漢字を明朝体化するに際して，包摂の範囲を明示的に示す必要がある．本稿では，その対処法を述べるとともに，作字プラットフォームの改善点を報告する．
キーワード：日本古辞書，写本文献，外字，HDIC

Design and Improvement of Kanchi’inbon Ruiju Myogisho by GlyphWiki

Shoju Ikeda (Hokkaido University)
Yuan Li (Kansai University)
Guanwei Liu (The University of Tokyo)
Munho Jung (Hokkaido University)

Abstract: In March 2022, the Integrated Database of Hanzi Dictionaries in Early Japan (HDIC) project released a full-text database of the Kanchi’inbon Ruiju Myogisyo (hereinafter referred to as Myogisho). Myogisho is an important source of Japanese language in the Heian period. Its value is that it contains a large number of Japanese vocabularies and serves as a material for the lexical history of the Japanese language. It serves as a material for Sino-Japanese history as a lot of tone marks reserved, and the Hanzi has a wealth of annotations on the style of characters such as orthodox 正 and common 俗, making it a material for the history of writing. Myogisho has been handed down as an old manuscript, but no printed copy exists yet. This is due to the fact that there are a lot of difficult characters in the Myogisho, but by using Unicode and GlyphWiki together, it is possible to compose the necessary glyphs. However, when rewriting handwritten Hanzi into Mincho, it is necessary to explicitly indicate the scope of unification. In this paper, we describe how to deal with this problem and report on improvements to the character creation platform.
Keywords: Dictionaries in Early Japan, Manuscripts, external character (Gaiji), HDIC

１．まえがき

平安時代漢字字書総合データベース（HDIC）プロジェクトでは，観智院本『類聚名義抄』（以下，名義抄と略すこともある）の全文テキストデータベースを 2022 年 3 月に公開した．名義抄の掲出字数は約 42,000 字であるが，難字・僻字の極めて多い古写本であるため，Unicode で入力可能なものはその 80%程度であり，しかも原本の筆画の細かな部分は表現できていない．さらに残り 20%は IDS により構成する漢字部品を示すにとどまり視認性を欠いている．

原本画像を併用する方法もあるが，誰もが自由に使えるという点で課題が残る．しかし Unicode と GlyphWiki とを併用すれば，必要な字形を作字することが可能である[1][2]．辞書研究の領域では，GlyphWiki の作字の実装例として，鈴木慎吾の篇韻データベースが挙げられる[3]．ただ，手書きの漢字を明朝体化するに際して，包摂の範囲を明示的に示す必要がある．以下，その対処法を述べるとともに，作字プラットフォームの改善点を報告する．

©2022 Information Processing Society of Japan

２．研究資料

２.１ 資料解説

名義抄は，平安時代の日本語の重要資料であるが，その価値は，多数の和訓を収録し日本語の語彙史の資料となること，和訓と漢字の多くに声点が施され音韻史の資料となること，漢字には「正」「俗」等の字体注記が豊富で文字史の資料となることの三点である．言語資料として利用する前提として，名義抄の依拠資料に関する研究（出典研究），名義抄の諸本の系統関係を明らかにし祖本の本文を復元する研究（系統研究）の蓄積も少なくない．しかし，名義抄は古写本として伝存するが，その活字翻刻本はいまだ存在しない．活字翻刻本の不在は，名義抄の利活用を制限することとなっている．また近年の中国では，域外漢籍研究の観点から，日本の漢文古辞書に着目した研究も増えている．

名義抄のデータベース化の難しい点は，難字・僻字が極めて多いことである．これは Unicode と GlyphWiki とを併用すれば，必要な字形を作字することはそれほど難しくない．難しいのは，手書きの漢字を明朝体化するに際して，包摂の範囲，包摂の粒度をどのようにするかである．名義抄の出典研究・系統研究では，微妙な字形差が問題となる可能性があるうえに，研究者にはそうした利用のニーズが多い．一方，漢字にどのような音注（反切，片仮名音注，声点），漢文意義注，和訓注が施されているかを他の字書類と比較参照して考察できなければ本文を読んだことにならない．一般の利用では緩やかに包摂した方が一般書籍・著作での利便性が高いし，なにより将来的に Unicode に採用される可能性が高まる．

整理すると，名義抄の資料価値は，主として，語彙史，音韻史，文字史の三点にあり，名義抄の文献学的研究は，諸本系統研究，出典研究，内容研究の三点にある．どの点を重視するかによって求める字形が相違するということである．文字史としての資料価値，本文の内容研究に課題が多いと考えているので，その方面での研究に有用な字形設計方針をとることとした．

２.２ 構成・収録字数・文字コード化

名義抄は仏・法・僧の構成で，仏をさらに仏上・仏中・仏下本・仏下末に四分，法を法上・法中・法下に三分，僧を僧上・僧中・僧下に三分し，冒頭に目録一冊を加えて全十一冊である．冊別の部首収録の詳細は表 1 にまとめた．

表 1　名義抄における冊別の部首収録

仏上(10)	仏中(10)	仏下本(10)	仏下末(10)
人彳走匚麥一丨十身	耳女舌口目鼻見日田肉	舟骨角貝頁彡髟手木犬	牛片豸乙兀収八大火黒
法上(10)	法中(10)	法下(20)	
水氵言足立豆卜面歯山	石玉色邑阜土心巾糸衣	示禾米丶宀勹穴雨門口尸虍广鹿疒歹子斗臥寸	
僧上(8)	僧中(24)	僧下(8)	
艸竹力刀羽毛食金	厶瓜网皿瓦缶弓攵矢斤矛戈欠又支殳皮革韋車羊馬鳥隹	魚虫鼠龜鬼風酉雑	

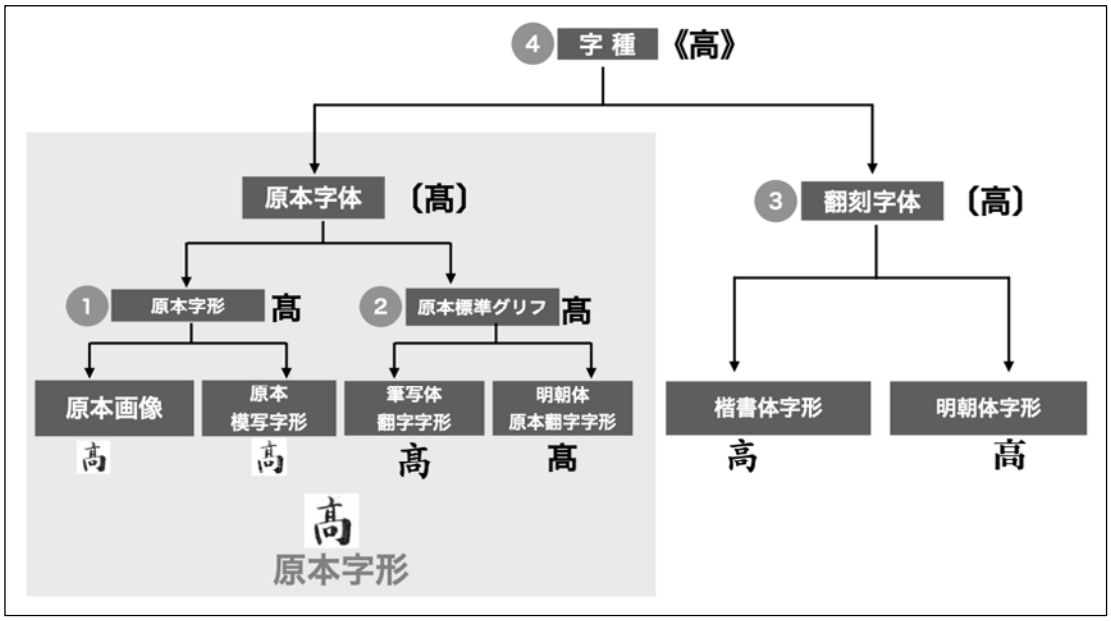


図 1　字種字体字形記述システム

©2022 Information Processing Society of Japan

観智院本類聚名義抄の掲出項目は, 約 32,000 とされる. 詳細の算出については, 酒井[4]や池田[5]の算出方法及び結果が参考になる. そして, 観智院本類聚名義抄データベース (Kanchi'inbon RuijuMyogisho Database, 略称 KRM)の整備に伴い, 池田ほか[6]の論考においては名義抄の基礎データを再集計し, 掲出項目数は 32,607 項目, 掲出字数は 42,340 字となっている. 「人部」を例に, その文字コード化の実態を表 2 で示した.

表 2　人部収録字の文字コード化の内訳[1]

部首	Unicode	IDS	その他	合計
人部	821 71.3%	233 20.2%	97 8.4%	1,151 100%

3. 観智院本類聚名義抄作字の位置付け

3.1 字種字体字形記述システム

現在の大規模文字集合 Unicode 等は, 基本的に現代社会の文字生活に適応するものであるため, 古文献・歴史・国語学の多くの人文学研究分野では対応できない漢字が多くある. 大規模文字集合に含まれない字体の場合, 字体レベルでの表現は, 特に漢字文献の範囲で見ると, まだ不十分である.

HDIC における観智院本類聚名義抄データベースの翻刻, 電子テキスト化, および字形再現は, 「高」の例で図示すると図 1 のようになる. 最初の段階の翻刻および電子化は, 図 1 の右側の翻刻字体のブロックのイメージで, 研究代表者の池田によって実現されてきた.

高田[7]の字形・字体・字種と異体字モデル[7]の論考を参照し, 字種・字体は抽象字形であるから, 厳密には図形化不能なものであると考えられるが, 便宜的に,「高」を代表させて, それぞれ《 》・〔 〕で括って表している. 「高」・「高」は, 高田[7]によれば, 字形が違うのはもちろんであるが, 字体も違うと考えられる.

このように考えると, 従来の唯一の翻刻字形（字体）の対応しきれないことが分かる. 体系的な翻刻システムで, 写本の内容を記述することが重要になる. そして, このような体系的な翻刻システムは, 文献研究, 言語研究, 漢字情報処理の多様なニーズに応えることを可能にする.

諸本研究・書承研究の視点から見ると, 書写者の相互関係を明らかにすることが重要なので, ①〜②のあたりで写本文献字形の記述に重点を置く必要がある. 出典研究・言語研究の視点から見ると, 字形を解釈し, 出典の相互関係が重要なので, ②〜③のあたりで写本文献字形の記述に重点を置く必要がある. 漢字情報処理の視点から見ると, Unicode の範囲で字形情報を記述し, 電子テキストによる検索と表示が重要なので, ③〜④のあたりで写本文献字形の記述に重点を置く必要がある.

3.2 各階層のメリットとデメリット

①原本字形

原本画像は, 文献を撮影した画像ファイルで, 忠実に文献の内容を再現できるが, 画像処理や権利関係で論著への利用に難易度が高い. 現在の各研究機関・図書館で公開されるデジタルアーカイブがその実例にあたる.

原本模写字形（筆写体）は文献資料もしくはその原本画像を元に, 模写したものである. 原本模写字形はほぼ忠実に原本字形を再現するができ, 論著での利用の代替手段として有効である. ただし, データ共有の面では意義が少ない. 江戸・明治時代における各種文献資料の影写本がその実例にあたる. さらに, 個人レベルで個別の原本字形の模写作業もその実例として挙げられる.

②原本標準グリフ

原本翻字字形（筆写体）は, 元の文献資料の原本字形から, 研究者により解釈される字形である. これは原本字形から, 明朝体化が可能な段階まで抽象化するステップとも言える. 各種の古典籍の索引, 例えば名義抄の正宗敦夫の索引[8]がその実例にあたる.

原本翻字字形（明朝体）は, 論著での利用に必要であり, 作字の量と質に難しい面がある. 書承研究では原本模写字形に近いレベルでの明朝体が必要となるが, 出典研究では翻字字体に近いレベルでの明朝体が必要である.

さらに, 未だに符号化されていないものについて, 将来的に Unicode エンコーディング提案する際に, 提出可能な字形を考えると, 翻字字体に近いレベルでの明朝体が必須である. 本稿で報告する GlyphWiki によって作成された名義抄の掲出字形のグリフがその実例にあたる. すなわち, 名義抄の GlyphWiki の作字は, 原本標準グリフの明朝体による実現形である. フォントの統一のために GlyphWiki で作字し, 既存フォントを再利用することとする. 新たなグリフの作成で対応するものは, GlyphWiki による作字が必須となり, どこ

[1] IDS(Ideographic Description Sequence, 漢字構成記述文字列),「その他」は主に次の五種類の状況を含む：虫損で判読不能(□), IDS 方式記述不可・筆画・構造が複雑または不明瞭(=), 省略・代用符号(−), 踊字(ゝ), IDS 方式記述不可または煩雑・筆画・構造に明瞭さを欠く(β方式).

©2022 Information Processing Society of Japan

まで統一的なフォント作成が可能かは大きな課題となる．

③ 翻字字体（明朝体）

翻刻字体は現存の Unicode 漢字で処理できるものである．検索・表示が容易，表示は使用フォントに依存する．日本国語大辞典の表記（図 2）によるものがその実例にあたる[9].

図 2　日本国語大辞典 第二版「朝顔」の例

3．3 翻刻プロセス

従来の文献翻刻においては，暗黙の文字包摂の問題が議論されてきた．豊島[10]は文字計量研究における「明示的で操作可能な文字の同定規準」の重要性を強調し，文字調査では「暗黙の文字包摂を排除出来ない」と指摘した．「暗黙の文字包摂」について，池田[11]は「文字調査者が漢字の常識を働かせて，異なる字形・字体の文字を同じ文字として処理してしまうことである」と定義する．上述の字種字体字形記述システムは，従来の文献翻刻における「暗黙の文字包摂」を明示的に示すことに有効である．

従来の翻刻は，それぞれ研究者の知見によって進めてきたが，研究者脳内の思考過程（暗黙の包摂プロセス）を模式化すると，次のようなイメージになる（図 3）．

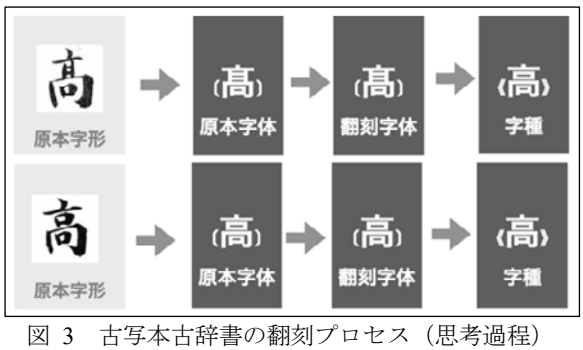


図 3　古写本古辞書の翻刻プロセス（思考過程）

従来の翻刻作業は，このような符号化の制限の下で，研究者によって，古い文献の原本字形を通行字体に包摂したり，手書きの字形や外字を作成したり，翻刻作業を行ってきた．そこには二つの問題点がある．一つ目，研究者の知見で行った翻刻は，暗黙の包摂が避けられない．二つ目，多くの歴史上の字体は文献に記述されても，大規模文字集合に含まれないので、記述する手段がなく，記録することができないのが現状である．

今回，我々の研究グループの名義抄に関する作字作業は，漢字字形共有サイト GlyphWiki を利用し，原本字形の書き振り等の影響を捨象し，研究者の判読した字体の骨組みを表現することである．この作業によって，明確な包摂規準を確立することが可能となる．そして，名義抄の編纂段階で記述されようとした字体を，記録することができる．

ただし，全く新しい試みとなるため，これまで文系研究者の内にある翻刻規準（暗黙の包摂規準）を，明確に描き出すのは容易なことではない．具体的に，細かくどのようなグリフの作成が原本標準グリフの規準に達するかは，核心となる問題である．完璧な規準をすぐに提示することは難しい．作業・検討を推進するとともに整えていかなければならないものだと考える．本稿は，その枠組み，検討方法，そして問題所在等を報告したい．

4．作字作業のプロセス

名義抄の作字作業は 2019 年 12 月に開始した．事前準備・点検は 1 名（前半 2 年間は李，後半は鄭），作字プラットフォームの開発・管理は劉が担当し，池田は全体を総括した．作字担当者は北海道大学文学研究院の院生・研究員数名，京都大学文学研究科の院生 1 名である．担当者は，古辞書に関する知識を持つが，Unicode と漢字の包摂などに関する知識には差があった．作業の期間・内容などを箇条書きすれば次のとおりである．

第一段階　2019 年 12 月から 2020 年 2 月まで．
紙媒体の作業台帳，約 2,000 字のグリフを作成．
第二段階　2020 年 7 月から 2021 年 2 月まで．
オンライン対応，HDIC Viewer における作字プラットフォームへ転換，約 17,000 字のグリフを作成．
第三段階　2021 年 7 月から 2022 年 3 月まで．作字プラットフォームの改善，新たに 19,000 字を作成．
第四段階　2022 年 4 月から 2022 年 8 月まで．
作字担当者が入れ替わり，再度作字作業トレーニングを実施．新たに 5,116 字を作成（図 4）．

©2022 Information Processing Society of Japan

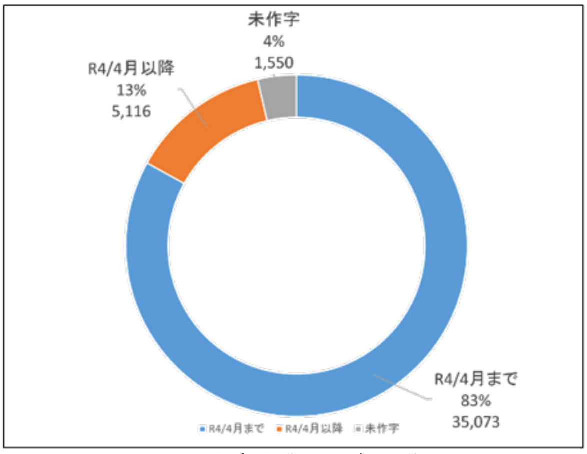


図 4　作字作業の現在の進捗

５．作字作業の詳細

５．１　作字対象

作字対象は全部で次の三種の状況がある．そのうち，新たに作字が必要とされる状況は次の二種である（①②）．

① Unicode で符号化されていない掲出字
② Unicode に既存の漢字で翻刻字体として対応したが，原本字形とのずれが大きい掲出字

さらに，既存の Unicode 明朝体字形で十分に表現できる掲出字に対しても，所在情報を付与することで，すべての原本字形を表現することを試みる（③）．

③ Unicode の明朝体字形で十分に表現できる掲出字

次の例 1 は「弄」の俗字で，①に該当する Unicode で符号化されていない掲出字となる．KRM では IDS 方式で処理した．

例 1 ⿰亻⿱壬下　俗弄字
(仏上 4　人部)

そして，次の例 2 は「参」の異体字であり，②に該当する Unicode に既存の漢字で翻刻したが，原本字形とのずれが大きい掲出字である．

例 2 ―(人)参　カノニケクサ
(仏上 1　人部)

さらに，次の例 3 は「伿」の例で，既存の Unicode 明朝体字形で十分に表現できる掲出字であり，③の状況に該当する．所在情報を付与することで，その名義抄の字形再現データを再作成した．

例 3 音易　コヽロ　モノウシ　憒也
(仏上 2　人部)

５．２　「人部」の全体像

これまでの作字作業について，「人部」を例に，その詳細，実態等をまとめる．人部の総字数は 1,151 字である．そのうち，Unicode で符号化された字で翻刻したものは 821 字で，全体の約七割を占める．IDS 等で表現したものは 233 字で，約二割を占める．そのほかの符号類のものは 97 字で，約一割を占める．

最初のステップは，未符号化の IDS 等で表現したものを作字したが，次の三つの問題点が見えてきた．

1 明朝体翻刻字体・原本標準グリフの混在
2 Unicode 符号化字と原本字形のずれ
3 原本画像利用の制限

これらの 3 点を解決するために，すべての原本字形を，原本標準グリフ（明朝体原本翻字字形）のレベルで再現することを目指す．

５．３　作字結果の検証

作字者やユーザーによって，原本標準グリフに対する理解は異なる．作字者のレベルで，原本標準グリフに対する理解にどれぐらいの違いがあるのかを確認・調査した．

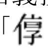
①写本字形の視認差

作業開始の段階，作字担当者全員に同じ原本画像を示し，同時に字形を作成してもらったが，次の例 4 の字について写本字形の視認差が生じた（表 3）．

例 4 「⿰亻享」俗停字
(仏上 3)

表 3 人部「⿰亻享」に関する視認差

担当者	担当者 1	担当者 2	担当者 3	担当者 4
作字 字形				

この例について，注文内容，前後の掲出項目，そして名義抄の諸本等で確認した結果，担当者 3 の字形「  」が最も相応しい．

これは写本の字形のみに注目することと，作字担当者の予備知識の差に原因があると考えられる．解決策として，原本の前後の項目を確認し，同一グループと判定する項目の注文，字体注記等の内容を参照することとした．さらに，これらの内容を手掛かりに，ほかの辞書や資料で調べる必要性についても確認した．また，現存する諸本の範囲で，対応する掲出字の有無も確認する．

©2022 Information Processing Society of Japan

②作字字形のデザイン

次の例 5 に記した作字は，どちらかといえば手書き字形に近似する．GlyphWiki の KAGE システムのデータの詳細を確認すると分かるように，多くの部品要素を取り入れ，細かく調整したような作業の様子が窺える．作字担当者に確認したところ，なるべく原本字形を忠実に再現するために，筆画を細かく分解して作字した結果ということであった（図 5）．

例 5　 「⿰亻草」ハヤウス

（仏上 19）

図 5　GlyphWiki 字形エディターにおける「⿰亻草」の KAGE データ

これは前述（第 3 節）の原本翻字字形（筆写体）に該当する作字と言える．また、これは原本字形から，明朝体化が可能な段階まで抽象化するステップとも言える．

なお，符号化されていないものとして，将来的に Unicode エンコーディング提案する準備段階と考えると，翻字字体に近いレベルでの明朝体が必要となる．ゆえに，このような書写字形のゆれを捨象し，なるべく均一に構成する部品要素を調整しなければならない．そのためには、できるだけ既存の部品を利用して作字することが必要となる．

５．４ 作字作業の実際

作成したグリフは，次の二種に大別される．

A 現存の Unicode 漢字で処理できるもの

これは通行字体で対応できる掲出字である．さらに次の二つの状況があるが，実際の作業では両者を区別する必要がない．人部では A に該当するものは 448 字がある．

A-1 通行字体と差異がなく，包摂の処理も問題にならないもの．

例 6 ：　

A-2 通行字体と差異があるが，現行各種の包摂規準で処理可能なもの．

例 7 ：　

A-2 について，筆画の交わる位置等の違いは，現行各種の包摂規準で処理可能であるが，そのうち，一部の包摂規準に関して，特に部品の構造が異なる場合，新たなグリフの作成で対応した．次はその例示である（図 6）．

図 6　JIS X 0208 包摂規準の 145 番と 149 番

B 新たなグリフの作成で対応するもの

人部では B に該当するものは 373 字がある．これらは次の二つに分けられる．

B-1 各種資料に問題となる部品が取り上げられているもの．

例 8：

この類のものは，概ね GlyphWiki の既存グリフで処理した．

B-2 各種資料で確認できないが，筆画増減，部品構造・配置に差があるもの．

例 9：

この種の掲出字について，既存データは IDS 方式，＝方式，β 方式の三種で処理した．IDS 方式は，主に筆画・構造が明瞭な字形を処理した．＝方式は，IDS 方式不可，筆画・構造が複雑または不明瞭な字形を処理した．β 方式は，IDS 方式不可または煩雑，筆画・構造に明瞭さを欠く字形に対応した．

５．５ 作字の質と量

上述した「A 現存の Unicode 漢字で処理できるもの」は，検索・表示が容易に対応できる．しかし，表示は使用フォントに依存し，パソコン環境によって，安定的な字形表示は難しい．一方で，「B 新たなグリフの作成で対応するもの」のフォントと齟齬がよく生じるので，表示・印刷に統一感を欠くこととなり，翻刻本文の質，完成度に影響が生じる．解決策として，フォントの統一のために GlyphWiki 作字することとする．KRM を参照し，所在情報に基づく ID を付与することで，既存フォントを再利用し，対応するグリフを作成する．

そして，「B 新たなグリフの作成で対応するもの」は，GlyphWiki による作字が必須で，どこまで統一的なフォント作成が可能か，質の向上・維持・管理が重要である．既存する各種包摂規準との対応，新規包摂規準の設定が必要となってくる．

©2022 Information Processing Society of Japan

６．作字プラットフォームの改善

作成されたグリフと作字担当者が増えるとともに，今まで利用してきた作字プラットフォーム（以下，旧プラットフォームとする）に改善すべき点を見つけた．次の 3 点について述べる．

６.１ GlyphWiki ユーザー管理の改善

GlyphWiki で登録された字形は基本的に誰でも編集できるが，「ユーザー占有グリフ」として登録された字形はそのユーザーしか編集できない[12]．第一段階と第二段階で作成された字形は，作業者各自のアカウントの「ユーザー占有グリフ」で登録した．これらの字形のグリフ名にはユーザー名の情報も含めるため，作成字形数の増加と作業者の交代とともに，グリフ情報の把握が困難となった．

改善する方法は，作業者各自がユーザー登録する代わりに，事前に準備した GlyphWiki アカウントを割り当てた．同時作業の人数を考慮して，「hdic-tanki01」から「hdic-tanki10」までの 10 個の GlyphWiki アカウントを登録して利用した．さらに各アカウントと作字担当の範囲をバインドすることにより，GlyphWiki に登録されたグリフ名の確認と作字プラットフォームの管理が容易となった（図 7）．

図 7　担当範囲ページ

６.２ 担当範囲ページの細分化

第二段階（2021 年 2 月）までの旧プラットフォームでは部首単位で表示しており，作業範囲を「〇部首の〇番目から〇番目まで」のように作字担当者に伝えたため，担当箇所の確認や進捗管理などはやや困難だった．

改善策として，作業範囲ページの表示を動的にした．作業範囲ページの URL に「事前用意したアカウント名」，「部首」，「部首内の開始番号」，「部首内の終了番号」を設定して，担当ページの表示内容はウェブサイトのルーティングによって変更できるようになった．担当範囲の変更や伝達はより簡便になった（図 8）．次は URL の一例：「https://viewer.hdic.jp/krm-glyphwiki-tanki/hdic-tanki03/阜/1/50」．

図 8　作業担当ページ

６.３ 字形校正の支援

作字の品質を保ち，かつ進捗を管理するため，点検担当の李と鄭が作成済みの字形を校正したうえ，GlyphWiki アカウントの「hdic」のユーザー占有グリフに移管することにした．

①GlyphWiki に新しいグリフページの作成，②担当者が作成した字形のデータを新しいグリフに移管，③原本字形の確認，④問題個所の修正の一連の作業が必要である．

第二段階までは，①と②の作業を個別に GlyphWiki で行った．第三段階以降は，校正作業の効率向上のため，①と②の操作を作業担当ページに集成した専用サイトを構築した（図 9）．さらに，④の場合でもワンクリックで GlyphWiki の字形編集ページに移動できるようになった（図 10・11）．

図 9　字形登録確認専用サイト

©2022 Information Processing Society of Japan

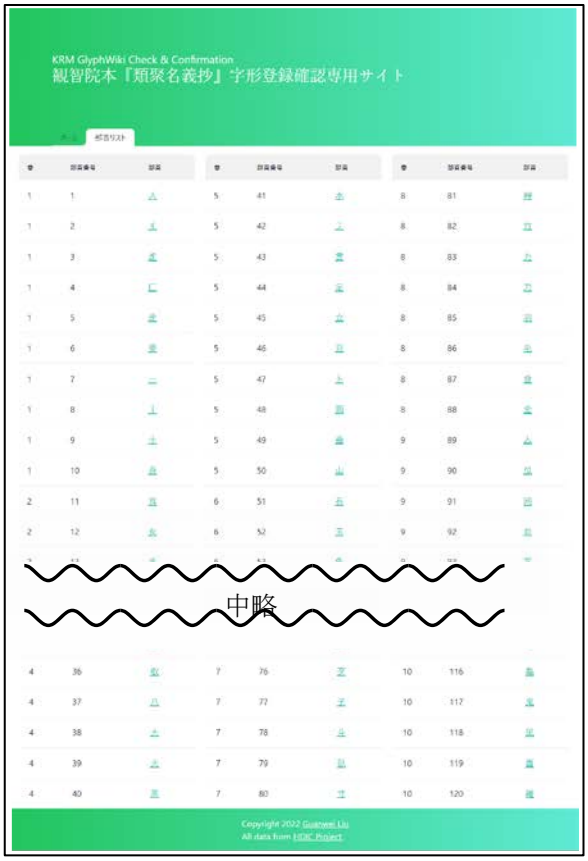


図 10　部首リスト

図 11　部首別字形登録確認ページ

7．あとがき

以上，HDIC における GlyphWiki 名義抄字形の再現とその改善点を報告した．この手法は，古写本の字形再現を必要とする分野で活かされることが期待される．例えば，色葉字類抄については，藤本灯がデータベースの構築を進めているので，そちらのプロジェクトへの実装も検討したい．作字プラットフォームの改善策として，作成した字形の登録確認専用サイトとそのソースコードを公開している[2]．

謝辞

本研究の一部は，科研費 JSPS19H00526，21K18013，21H00529，21K18364，東京大学史料編纂所「データ駆動型歴史情報研究基盤の構築」の助成を受けたものである．

参考文献

[1] 李媛，劉冠偉，池田証壽：観智院本類聚名義抄に関する作字作業をめぐって，東洋学へのコンピュータ利用第 32 回研究セミナー，2020，p.27-48.
[2] 劉冠偉，李媛，池田証壽： GlyphWiki による観智院本類聚名義抄字形の再現， じんもんこん 2020 論文集，2020，p.203-208.
[3] 鈴木慎吾 :『切韻』諸本テキスト一覧システムの構築について，じんもんこん 2018 論文集, 2018, p.117-122.
[4] 酒井憲二：類聚名義抄の字順と部首排列．本邦辞書史論叢，三省堂，1967，p.191-258.
[5] 池田証壽：観智院本類聚名義抄の掲出項目数と掲出字数，北海道大学文学研究科紀要，2008，Vol. 124，p.137-151.
[6] 池田証壽，劉冠偉，鄭門鎬，張馨方，李媛：観智院本『類聚名義抄』全文テキストデータベース：その構築方法と掲出項目数等の計量，訓点語と訓点資料，2020，Vol. 144，p.129-105.
[7] 高田智和：字形・字体・字種と異体字，日本語学，2013，Vol. 32，No. 5，p.180-191.
[8] 正宗敦夫：漢字索引，類聚名義抄 第貳巻．風間書房, 1955.
[9] JapanKnowledge『日本国語大辞典 第二版』の見本ページによる．
https://japanknowledge.com/contents/nikkoku/book_second03.html，（参照 2022-10-30）
[10] 豊島正之：書評 横山詔一・笹原宏之・野崎浩成・エリク=ロング『新聞電子メディアの漢字－朝日新聞 CD-ROM による漢字頻度表－』．日本語科学，1999，Vol. 6，p.91-102.
[11] 池田証壽：コーパスによる文字の研究，日本語学．明治書院，2003，Vol. 22，No. 5， p.210-220.
[12]“GlyphWiki:グリフを登録する”．
https://glyphwiki.org/wiki/GlyphWiki:グリフを登録する，(参照 2022-08-31)

[2] https://github.com/toyjack/krm-glyphwiki-confirm

©2022 Information Processing Society of Japan