

説文小篆の符号化計画に対する Variation Selector 提案とその反応

鈴木 俊哉[†]

2014年9月のISO/IEC JTC1/SC2/WG2会議にて、ISO/IEC 10646への説文小篆の追加提案が行われた[1]。この当初の提案では説文解字のいくつかの版本を整理し、版本ごとの異体字は連続した符号位置になるように並べるものであった。しかし、2015年10月のWG2会議では、この計画は単一の版本(藤花樹本)に基づいた単なるフォント作成に縮小されており、さらに異体字は将来的には別文字符号化すればよいが、当初の計画のように連続した符号位置になるような空き符号位置は予約しないため、すべて後ろの符号位置にすればよいという設計に変わっていた[2]。筆者らは『説文解字繫傳』の版本比較研究における字形差の弁別基準に大きな開きがあることを指摘し、具体的な字形差を個別に検討して符号位置の付与を判断することは困難で、Variation Selectorによる字形差の取り扱いの可能性を述べた。このVariation Selector提案をISO/IEC JTC1/SC2/WG2に送付したが、現段階での反応を報告する。

Proposal of Variation Selector for Small Seal Script and the Responses.

suzuki toshiya[†]

On September 2014, Chinese and Taiwanese experts submitted their proposal to include Shuowen Small Seal into ISO/IEC 10646. Their proposal on 2014 was designed to be a merged collection of major version of Shuowen Jiezi, and the per-version variants are expected to be placed at continuous codepoints. But the revised submission on 2015, the project was reduced to be a simple font production based on one version of Shuowen Jiezi. Also it removed the discussion how to handle per-version variants from the submission. As a result, the project to standardize Shuowen Jiezi is either lacking the criteria for the unification, as ceased project for Oracle Bone, in spite of the situation that the user community of Shuowen Small Seal is the largest one of the Old Hanzi users. In the last report, we summarized the distinctions of the “distinctive glyph difference” among the scholars working for Xu Kai, and reported they were too varied to unify as a single and consistent unification rule. We submitted a proposal to apply source-based variation selector for small seal script, to ISO/IEC JTC1/SC2/WG2. In this report, the proposal and the responses to it is reported.

1 はじめに

『説文解字』は100年頃に後漢の許慎によって書かれた、漢字字書としてはほぼ最古のものである[4][5]。許慎は字義の解釈のため、当時通行していた隸書ではなく、象形的な性格が残る小篆字形を用いた。より古い甲骨資料を得ている現代では、小篆に基づいた議論に限界があることは明らかだが、現代漢字では古漢字を指示できないという標準化提案の動機は許慎の意識とも共通しているといえるかもしれない。現在見ることができる説文は900年ほど後の宋代の刊本で、五代末の南唐から宋にかけての学者である徐鉉(大徐)が校訂した版(一般に大徐本と呼ばれる)である。現在説文解字として広く参照されているのはこの宋刊本か、さもなければこれを底本にした明末以降の翻刻本や校訂本である¹。

2012年に中断した甲骨文字の標準化計画において「何を目的とする標準化なので、どのような機能が実現できなければならないか」明らかになることはなかったが[12]、説文小篆の標準化においてもこの論点はやはり明らかにされ

ていない。そのため、どのような字形差を分離するのか、あるいは統合するのかについて十分な議論がされていない。

たとえば、図1のように、明らかに部品の数が違うものを分離する判断は自明としても(図1上)、同一版本の中で統一が取れていない字形について異体字として符号化する(図1下)という方針には議論の余地がある。



図1: 2014年の提案で分離を予定されていた字形差の例[1]

上: 段注本に見える「水」を付加した異体字を区別する

下: 藤花樹本に見える「世」の字形差を区別する

さらに、図2のように説文解字の中には字形・字注がほぼ同じであるにも関わらず複数回出現する文字などがあり、これらについても議論はされていない。

189	00939		右	口	22	Zhengzhuān
560	02078		右	又	76	Zhengzhuān

図2: 2015年の提案で分離を予定されている異体字の例[2]

[†] 広島大学 Hiroshima University

¹ 日本では『説文解字』の各版本の概要は[4][5]が良く参照される。中国での認識は[6][7]が参考となる。明末清初の汲古閣本については[8][9]、清代の小徐本の流通については[10]を参照されたい。また以上に漏れる断片について[11]も参照されたい。

2 ISO 文字符号における字形の取り扱い

ISO/IEC 10646 は文字の表示図形の詳細には立ち入らないという発想で設計がはじまり、漢字については ISO/IEC JTC1/SC2/WG2 の下に Ideographic Rapporteur Group という特別な作業グループを設け、追加提案される漢字に対し既存の符号化済漢字、あるいは、同時に提案されている他の漢字との統合可能性を検討するという開発体制になっている。

2.1 現代漢字における微細な字形差の取り扱い

ISO/IEC 10646 の最初の版において、既存の各国の漢字符号との往復互換性を保証するため、原規格分離(字形の上では統合できるものであっても、既存の文字符号規格で分離されている場合、一対一対応を保証するため、統合しない)という原則が用いられた。この原則は、一応は最初の版の統合漢字(U+4E00~U+9FA5、いわゆる URO)に限られており、その後の追加漢字では適用しないものとされた。

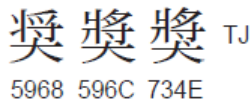


図 3: 本来統合するが原規格分離によって分離された例[13]

その後、意味的な使い分けがなく、統合可能な字形差しかないが、何らかの需要によって別符号位置で処理しなければならないという場合は、統合漢字ではなく、互換漢字(U+F900~U+FAFF)として扱う方針がとられた(もともとの互換漢字は、韓国の漢文字符で同じ字形で読みが異なるものに複数の符号位置を与えたことによるものであったが、その後、統合可能な字形差しか持たないベンダ外字などを取り込むためにこのような用途になった)。

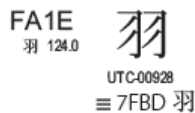


図 4: 本来統合するが互換漢字で別位置符号化された例[13]

互換漢字はあくまでも移行措置として規格に取り入れられたもので、正規化を行った場合に統合漢字に置き換えられるなどの制限がある。これは、統合できる字形差しかない文字の使い分け需要は将来的に収束すると考えられたためであるが、その後、日本の行政において統合可能な字形差しか持たない外字が大量に別コードで運用されており、それらを全て互換漢字として処理しようとすると数千個の符号位置が必要となることが明らかになった。大量の互換漢字を新たに追加することは望ましくないという意見もあり、ISO/IEC 10646 の 2008 年版以降では、統合漢字に枝番号を付加することによって字形の詳細を指定する技術として、Unicode が定めていた Ideographic Variation Selector(IVS) を ISO/IEC 10646 で正式に参照することとなった。

2.2 その他の文字での微細な字形差の取り扱い

このように、ISO/IEC 10646 における漢字は初版から絶え間なく拡張が続けられており、標準規格側で「微細な字形差をどのように扱うべきか」という問題への対応にも流動性がある。しかしながら、歴史的な文字の場合、漢字のように絶え間なく追補されるわけではなく、最初の段階の設計が大きな影響を与える。西夏文字に関しては互換文字や Variation Selector(VS)が議論されたことがあるが[14]、現時点では漢字以外の VS はモンゴル文字、パスパ文字、マニ文字、数学記号、携帯電話絵文字(正確には携帯電話絵文字のうち従来の記号類と統合され、それらが携帯電話絵文字とデザインが大幅に異なるもの)などごく少数にとどまっている。

また、携帯電話絵文字の VS も、ベンダごとの絵文字のデザイン差を指定するためのものではなく、携帯電話絵文字以前からあったテキスト環境用のシンボルと統合されたものについて、携帯電話風のファンシーな画像を指定するためのものである。



図 5: 携帯電話絵文字のための Variation Selector の例[15]

2.3 字形を指定しない Variation Selector

VS は基本的には異なる字形を使い分けるための、処理内容によっては無視しても構わない(あるいは無視すべき)枝番号である。しかし、最初に登録された IVS 集合である Adobe-Japan1 には既に既存実装との互換性のために、同じ字形に対して複数の IVS がわりあてられていた。さらに、この後追加された Hanyo-Denshi IVS は Adobe-Japan1 とほぼ同じ字形に対しても(字形差の判別基準が整合しないため)別の IVS をわりあてることとなった。その結果、現状では基底文字あるいは IVS が異なればなんらかの字形差がある、という状態ではなくなっている。



図 6: 字形差判定基準が異なるため、IVS を共用しなかったが、部分的に見るとよく似た字形に異なる IVS が振られる例[16]

Adobe が主導するオープンソースフォントである PanCJKV プロジェクトでは、中国・香港・台湾・日本・韓国へのローカライズを進めているが、ローカライズに際して共用できる字形が多いにも関わらず、どの地域でどの字形が適切かという情報には必ずしも法則性がない。このため、「字形単位でなく、ローカライズ対象地域の単位で枝番号を割り当てる VS」が提案されている[17]。この手法は具

体的な字形の詳細に対しての要求というよりは、字形に対して何らかのカスタマイズが必要であることと、その対象を指示することだけに力点を置いた VS の利用である。現時点で特に言及されていないが、今後、規格票の例示字形が改訂された場合にはそれに追従して字形を変更する意図があると思われる。

Variation Selector	Code Point	Region Code
VS246	U+E01E5	VN (Việt Nam)
VS247	U+E01E6	KP (DPRK)
VS248	U+E01E7	KR (ROK)
VS249	U+E01E8	JP (Japan)
VS250	U+E01E9	MY (Malaysia)
VS251	U+E01EA	MO (Macao SAR)
VS252	U+E01EB	HK (Hong Kong SAR)
VS253	U+E01EC	TW (ROC)
VS254	U+E01ED	SG (Republic of Singapore)
VS255	U+E01EE	CN (PRC)
VS256	U+E01EF	XK (Kángxi)—pseudo-region

PanCJKV で使用する IVS 一覧

VNKPKRJPMYMOHKTWSGCNKK

字字字字字字字字字字字
 骨骨骨骨骨骨骨骨骨骨骨
 曜曜曜曜曜曜曜曜曜曜曜

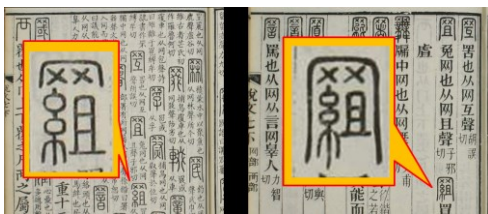
各 VS を適用した結果

図 7: PanCJKV のローカライズ用 IVS 提案[17]

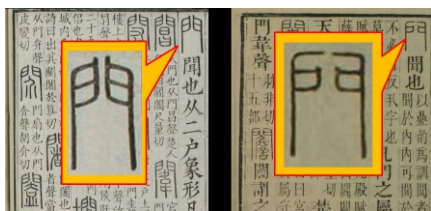
「一」はローカライズによる字形差が必要ないが、「字」については 2 種類、「骨」
 「曜」については 3 種類の字形があり、どの地域にどの字形を割り当てるべきか
 は規則性がない。このため、地域単位で VS を設定する。

3 説文小篆における字形差

『説文解字』見出し字の字形差とはどのようなものであろうか。冒頭に述べた例の中の「世」の異体字までも符号化するとすると、以下のようなものが考慮対照となるだろう。



版本ごとに部首の重なり方が異なる例「世」(藤花樹本(右)・陳昌治本(左))



版本ごとに異なるデザインポリシのある例「門」(藤花樹本(右)・段注本(左))

図 8: 版本ごとに異なる字形の例

3.1 異体字の別符号化とその排列

2003 年ごろの、全ての説文関係文献をもとに統合された文字表を作るという計画[18][19]の際の文字表は以下に示すようなものであった。

Serial No.	Rep. Script/Glyph	Original Shape/Glyph	Source	Period/ Epoch	Area/ Terrain	Material	Radical	Glyph Determ.	Corresp. Modern Char	Notes
1	上	上	《説文》大徐本			紙本文献	上	上	上	大徐本以古文为正篆
2	一	二	《説文》段注本			紙本文献	二	上	上	段注改古文正篆
3	上	上	《説文》大徐本			紙本文献	上	上	上	篆文形体
4	社	社	《説文》大徐本			紙本文献	示	社	社	
5	社	社	《説文》大徐本			紙本文献	示	社	社	古文
6	社	社	《説文》段注本			紙本文献	示	社	社	段注改古文
7	𠂔	𠂔	《説文》大徐本			紙本文献	又	𠂔	𠂔	
8	𠂔	𠂔	《説文》大徐本			紙本文献	又	𠂔	𠂔	籀文
9	𠂔	𠂔	《説文》大徐本			紙本文献	又	𠂔	𠂔	或体
10	𠂔	𠂔	《説文》大徐本			紙本文献	又	𠂔	𠂔	
11	𠂔	𠂔	《説文》大徐本			紙本文献	又	𠂔	𠂔	古文

図 9: 当初計画で想定していた小篆の文字表[18]

この文字表から推察するに、版本 A, B, C がある場合の文字符号は以下のように並ぶであろう。

文字 00001	見出字 00001 版本 A, B, C 型
文字 00002	見出字 00002 版本 A, C 型
文字 00003	見出字 00002 版本 B 型
文字 00004	見出字 00003 版本 A 型
文字 00005	見出字 00003 版本 B, C 型
...	...

図 10: 2004 年当時の小篆符号の構造

特定の版本にしか興味のない利用者にとっても、番号が飛び飛びにはなるものの、単純なソートが可能である。しかし、2015 年からの計画では以下ようになる。

文字 00001	見出字 00001 版本 A(B,C)型
文字 00002	見出字 00002 版本 A(C)型
文字 00003	見出字 00003 版本 A 型
...	...
文字 11108	見出字 11108 版本 A 型
文字 11109	見出字 00002 版本 B 型
文字 11110	見出字 00003 版本 B, C 型
...	...

図 11: 2015 年計画での小篆符号の構造

灰色部分については[2]では具体的な提案はない

このような設計を取るためには、版本 A が十分広い需要をカバーできること、また、版本 B, C の字形をどのように

区別するかの議論を行う必要がある。バージョン A の字形と、緩やかな統合規準を使うならば、大半のバージョンの字形を最初の段階で収容することができるが、図 8 のような字形差を区別しようとした場合、利用者が「どのバージョンはどの字形を使っているか」を熟知していなければならない。「門」のようにバージョン毎で統一されたデザインポリシーがあるような例は可能であるとしても、「羅」のような他の情報と何の関連もない字形差を記憶できるかどうかはかなり疑問がある。

もちろん、現代漢字にも単字では判断できない使い分けは存在する。たとえば日本国内の規格であれば「1981 年制定常用漢字に含まれていた漢字のシンニョウは 1 点につくるが、それ以外は 2 点につくる」というような、ある程度広く区分を共有する集団の場合である。現在の提案フォントがそのような利用者集団を抱えているか、あるいはその期待が持てるかには以下のような疑問がある。

- 広く参照されているのは藤花樹本よりも、陳昌治本や段注本である(たとえば『漢語大字典』が掲出する小篆字形は陳昌治本による)。
- 藤花樹本は宋刊本に最も近いという位置づけで選定されているが、疑問がある。
 - 藤花樹本が出版された清代は影印出版ができず、広く参照できる宋刊本が無かったが、現在は宋刊本影印のほうが広く流通している。
 - 藤花樹本の底本がはっきりしておらず、現存する宋刊本と異なる(図 12)。
- 提案フォントは藤花樹本に何らかの修正を行っているが、その根拠がはっきりしない(図 12)。

00342		00423	
00378		00429	
06966		00563	
08586		00948	

図 12: 宋刊本、清刊本の実際の字形と提案フォント

左より大徐本の見出し字の通し番号、岩崎本(宋刊本)、平津館本(清刊本)、陳昌治本(清刊本)、藤花樹本(清刊本)、提案フォントの字形。提案フォントは必ずしも藤花樹本の字形を忠実に示してはならず(左)、藤花樹本と提案フォントの両方とも宋刊本と異なる場合がある(右)。

この状況は、PanCJKV のような「マクロなコンテキスト情報はありうるが、そこからどの文字に対してどの字形を選定すればよいかの具体的な情報は無い」という状況に近い。そこで、小篆の標準化の速度に影響しない方法として、小篆字形に対して、PanCJKV のような「字形差を根拠としない VS」の適用を提案した[20]。

4 説文小篆のバージョンベース Variation Selector

筆者が提案した VS は、個別の字形差の判断を避けるため、PanCJKV の VS と同様にバージョンごとに設定するものとした。従って、まったく同じ字形のものでもバージョンの数だけ VS をわりあてる。提案にあたり、参考実装として小篆に対応する既存の統合漢字の符号位置に IVS を付加すると各バージョンの字形(藤花樹本を U+E0100、段注本を U+E0101、岩崎本を U+E0102、平津館本を U+E0103、陳昌治本を U+E0104 にわりあてた)が選択できるフォントを作成した。現在の Web ブラウザの多くは IVS と CSS の @font-face 宣言をサポートしているため、図 13 のように非常に簡易にこれを利用できる。

```
<html>
<head>
<style>
<!--
@font-face {
font-family: "SWVS";
src: url("SWVS-20160427-2240.ttf");
}
table, td, th {
border-collapse: collapse;
border: solid 1px black;
}
td {
font-family: "SWVS";
font-size: 36;
}
-->
</style>
</head>
<body>
<table>
<tr>
<th>&#x7F51;<br>(U+7F51)</th>
<td>&#x7F51;&#x0E0100;</td>
<td>&#x7F51;&#x0E0101;</td>
<td>&#x7F51;&#x0E0102;</td>
<td>&#x7F51;&#x0E0103;</td>
<td>&#x7F51;&#x0E0104;</td>
</tr>
```

羅 (U+262FD)										
置 (U+262FE)										
羅 (U+2632C)										
罍 (U+7F72)										
龍 (U+7F77)										
置 (U+7F6E)										
罍 (U+7F6F)										
罍 (U+8A48)										
罍 (U+7F75)										
罍 (U+2632D)										

図 13: バージョンベース VS を実装した小篆フォントと HTML での利用

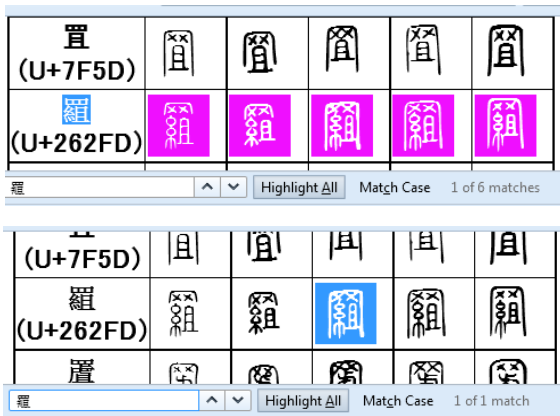


図 14: 参考実装フォントを用いた異体字群に対する検索例

上: 基底文字のみ入力、下: VS も含め入力

本来 VS は表示以外の用途では無視されるべきものだが、現時点では VS を含めたビットパターンで検索される処理系も多く、図 14 のように VS を指定することで特定のバージョンの異体字を絞り込むことができる。

4.1 説文小篆は現代漢字と別に符号化すべきか

さて、この提案は説文小篆を現代漢字とは異なる新たな文字としてで符号化した上で、その文字符号に VS を付加するという提案であるが、もし VS を使うのであれば、説文小篆そのものを既存の現代漢字の表示字形の一群として捉えることはできないだろうか。

図 13 に示した参考実装において、基底の符号位置として未符号化文字の提案に一般的に用いられる PUA ではなく、対応する統合漢字を選んだのは、Web ブラウザでは表示している文書の中では Web フォントを用いることができても、検索フォームなどの文書外部のインタフェースでは Web フォントを用いることができない制限に配慮したものであるが、基底文字の議論が十分にされているかには疑問がある。

筆者が古漢字の標準化に関わってから意見を尋ねた日本の専門家の反応は「説文小篆が標準符号化されれば便利かもしれないが、自分は現代漢字で十分指定できるので、自分は使わない」といった、反対はしないが積極的な意味はないという反応が多かった。使用に興味のある専門家の中でも、完全に現代漢字と分離し小篆としてしか表示されないような符号位置を期待する声も少数ながらあったが、対応する現代漢字との相互運用に配慮してほしい、といった意見のほうが目立った。おそらく、既存の小篆フォントが「現代漢字の一書体」として実装されていることや、外字コードでの表示と文書データとしての交換性の得失を判断した結果現代漢字で表記したという状況も少なくないからであろう。

5 提案に対する反応、今後の課題

提案に対する反応は、VS の設計よりも、まず現代漢字を基底とした VS で小篆を付加すべきかどうかという議論に流れた。現代漢字とは別個に符号化すべきだという意見は、大きくは 2 つの背景による。

- 説文小篆を表示しようとして制作したデータが、フォント不備などによって現代漢字にフォールバックしてしまうのはデータ制作者にとって許容できないのではないか。たとえば、グジャラート文字はデヴァナガリ文字から発展した文字だが、グジャラート文字のフォントが無いからといってデヴァナガリ文字にフォールバックしてしまうようなシステムは、グジャラート文字の利用者にとって、ゲタ(=)にフォールバックしてしまうシステムよりも受け入れ難いものだろう。
- 現代漢字に古漢字を「既存の漢字よりも、よりのもの字形に忠実に楷書化したもの」が何度も提案されつづけており、もっとも適切な現代漢字とのマッピングを作ることは非常に困難ではないか。

前者の評価は古漢字のデータがどのように使われているかに依存する。ここで例に挙げられたグジャラート文字とデヴァナガリ文字は、どちらもグジャラート語とヒンディー語の両方を表記できるが、現代では「グジャラート語を表記するのにグジャラート文字を用いる利用者」にとってはヒンディー語を表記するのに適切な文字はデヴァナガリ文字であり(たとえば Unicode の CLDR[21]ではグジャラート文字を使用する言語にはヒンディー語はなく、デヴァナガリ文字を使用する言語にはグジャラート語は入っていない)、異なる文字で表示すると内容の理解に影響を与える懸念がある。一方、説文小篆の場合は、説文小篆で表記されるものと、現代漢字で表記されるもの間に明確な線引きはできない。小篆が定められた秦代から既に篆書と隸書は同じ言語を書く別書体として平行して運用されており、また、さまざまな古字書で楷書体を示しながら篆書としていることも少なくない。さらに、最初から小篆でかかれた長文のテキストよりも、現代漢字のテキストを小篆で書き換えたものや、現代漢字の文献の中に部分的に小篆が混じっているもののほうが遥かに多い。そのため、内容主導で小篆と現代漢字のどちらが適切かを判断できるとしても、それが実際の文献ではどちらで示されているかと一致するかはかなり疑わしい。派生関係にはあるが独立した用字系として運用されているものと対比するのは適切でないと思われる。

現代漢字と小篆の関係については、数学用アルファベット(U+1D400~U+1D7FF)のように「文字符号上で独立して運用してきた例はないが、今後独立して符号化したほうが良いかどうか」という判断を求められている状況に近いの

ではないかと思われる。数学用アルファベットは規格票字形はもちろん通常のアルファベットと異なる書体で示されているが、通常のアルファベットを数学用アルファベットと同じ書体でデザインしたフォントが規格違反となるわけではない。ただし、数学用アルファベットに関しては互換等価性を用いた正規化が定義されており、処理方法によっては通常のアルファベットと同一視した検索などもできるが、小篆に対して現代漢字への安定した互換等価性をどのように定義するかは自明でない。この透過互換性は後から変更することができないので、仮に現代漢字とは別に符号化するとしても、対応関係の整理は必要な課題として残っていると思われる。

謝辞

本研究は科研費課題番号 24500116, 26330377, の成果を含みます。

参考文献

- [1] TCA and China: ISO/IEC JTC1/SC2/WG2 N4634, "Proposal to encode Small Seal Script in UCS", 2014-09-30.
- [2] TCA and China: ISO/IEC JTC1/SC2/WG2 N4688, "Proposal to encode Small Seal Script in UCS", 2015-10-20.
- [3] 鈴木俊哉, 鈴木敦, 菅谷克行: "説文小篆の字形差はどのように研究されてきたか", 情処研報, DD-101-6, 2016/03/24, p.1-6
- [4] 福田襄之介: 『中国字書史の研究』, 明治書院(1979)
- [5] 頼惟勤監修、説文会編: 『説文入門』, 大修館書店(1983)
- [6] 周祖謨: 「許慎及其説文解字」, 『問學集』, p.710-722.
- [7] 王貴元: 「《説文解字》版本考述」, 古籍整理研究學刊, (1999), 第6期, p.41-43,34.
- [8] 高橋由利子: 「『説文解字』毛氏汲古閣本について」, 汲古
- [9] 高橋由利子: 「段玉裁の『汲古閣説文訂』について」, 中国文化(55), 1997
- [10] 坂内千里: 『経部引用書から見た「説文解字繫傳」注釈考』(2014)
- [11] 倉田淳之助: 「説文展観余録」, 東方学報・京都, 第10冊・第1分 (1939), p.145-154.
- [12] Old Hanzi Expert Group: ISO/IEC JTC1/SC2/WG2/IRG N1836, "Report from the Old Hanzi Expert Group", 2012-02-23.
- [13] ISO/IEC 10646: "Information technology -- Universal Coded Character Set (UCS)", 2014.
- [14] Andrew West: "Re: Comment on PRI 98: IVD Adobe-Japan1 (pt.2)", 2007/03/21, 2016/06/14 閲覧, <http://unicode.org/mail-arch/unicode-ml/y2007-m03/0123.html>
- [15] Unicode Inc.: "Emoji Variation Sequences", 2016/06/14 閲覧, <http://unicode.org/emoji/charts/emoji-variants.html>
- [16] Unicode Inc.: "Ideographic Variation Database", ver. 2012-03-02, 2016/06/14 閲覧, http://www.unicode.org/ivd/data/2012-03-02/IVD_Charts.pdf
- [17] Ken Lunde: "Proposal to accept the submission to register the "PanCJKV" IVD collection", 2016-03-10, L2/16-063, 2016/06/14 閲覧, <http://www.unicode.org/L2/L2016/16063-pancjkv-ivd-collection.pdf>
- [18] China: ISO/IEC JTC1/SC2/WG2/IRG N1119, "Old Hanzi Samples from PRC", 2005-05-18.
- [19] China: ISO/IEC JTC1/SC2/WG2/IRG N1139, "References on Old Hanzi", 2005-05-25.
- [20] Suzuki Toshiya: "Proposal to Apply Source-based Variation Selector in Shuowen Small Seal Encoding", ISO/IEC JTC1/SC2/WG2 N4716, 2016/04/28, 2016/06/14 閲覧, <http://www.unicode.org/L2/L2016/16092-wg2n4716-shouwen-small-seal.pdf>
- [21] Unicode Inc: Unicode Common Locale Data Repository, 2016/06/14 閲覧, <http://cldr.unicode.org/index>