

外字実態調査の今後の活用

平成24年3月15日

富士ゼロックス株式会社
富士ゼロックスシステムサービス株式会社

全体成果物について

①文字情報基盤漢字同定リスト

MJ022001 - MJ023000

MJ022014	住基ネット: J+B99B 戸籍統一: 348820 UCS: U+839F UCS実装:
苔	
	字形一致: 3件 デザイン差: 1件 類似文字: 13件

字形一致	苔	苔	苔							
デザイン差	苔									
類似文字	苔	苔	苔	苔	苔	苔	苔	苔	苔	苔
類似文字	苔	苔	苔							

同一UCS一覧

MJ022013		
苔		

②不明文字リスト

(漢字を含む不明文字)

滢	邛	祿	标	峴	晃	裨
晔	尧	糞	綫	蛭	严	欢
队	运	辽	焜	垠	标	罗
导	焮	焮	峯	庠	弑	惜
溢	焯	灿	尢	廔	鬪	糙
岩	选	骅	桦	稳	烧	暨

(記号)

AM	PM	cc	Co,	Ltd.
(協)	(名)	(社)	(資)	(財)
(庁)	(官)	(安)	(長)	(振)
(健)	1/2	1/3	1/4	2/3
③	④	⑤	⑥	⑦
⑬	⑰	⑲	⑳	←
(福)	(遠)	(学)	(資)	(退)

(変体仮名)

𑖀	𑖁	𑖂	𑖃	𑖄
𑖅	𑖆	𑖇	𑖈	𑖉
𑖊	𑖋	𑖌	𑖍	𑖎
𑖏	𑖐	𑖑	𑖒	𑖓
𑖔	𑖕	𑖖	𑖗	𑖘

外字を提供頂いた市区町村向け成果物について

レイアウト等変更あり

①同定結果リスト

(1-866/866文字) 全体結果

外字	MJ文字図形名	補足
1_E0DC	MJ021691	デザイン差
𪛗	𪛗	住基: 344780 戸籍: 344780 UCS: U+82EB UCS実装:
1_E0DD	MJ006365	デザイン差
𪛘	𪛘	住基: 000870 戸籍: 000870 UCS: U+4E3D UCS実装:
1_E0DE		
𪛙		
1_E0DF	MJ009126	類似文字
𪛚	𪛚	住基: J+5805 戸籍: 059450 UCS: U+5805 UCS実装: O
1_E0E0	MJ022632	デザイン差
𪛛	𪛛	住基: 358940 戸籍: 358940 UCS: U+84FC UCS実装:

外字	MJ文字図形名	補足
1_E0E6		
𪛜		
1_E0E7	MJ025036	類似文字
𪛝	𪛝	住基: J+8CC0 戸籍: 416510 UCS: U+8CC0 UCS実装: O
1_E0E8		
𪛞		
1_E0E9	MJ009401	類似文字
𪛟	𪛟	住基: J+58FD 戸籍: 064850 UCS: U+58FD UCS実装: O
1_E0EA	MJ027290	類似文字
𪛠	𪛠	住基: J+9435 戸籍: 464820 UCS: U+9435 UCS実装: O

②同定不可能リスト

(1-332/3790文字) 同定不可能文字

外字	MJ文字図形名	補足
1_E004		
𪛡		
1_E00D		
𪛣		
1_E017		
𪛦		
1_E030		
𪛩		
1_E031		
𪛫		

外字	MJ文字図形名	補足
1_E0A7		
𪛯		
1_E0AB		
𪛰		
1_E0B5		
𪛳		
1_E0B9		
𪛷		
1_E0BB		
𪛹		

③同定結果CSV

A	B	C	D	E	F	G	
1	外字	MJ文字図形名	入力区分名	住基ネット統一文字コード	戸籍統一文字番号	UCS	UCS実装
2	*****_1_E000	MJ009234	字形一致	J+5869	061460	U+5869	○
3	*****_1_E001	MJ021701	類似文字	J+B94F	344840	U+82F1	
4	*****_1_E002	MJ011690	デザイン差	J+60E7	121630	U+60E7	○
5	*****_1_E003	MJ021839	デザイン差		346650	U+8938	
6	*****_1_E004		同定不可能				
7	*****_1_E005	MJ007854	デザイン差	J+5353	031810	U+5353	○
8	*****_1_E006	MJ011014	類似文字	J+5EB5	105940	U+5EB5	○
9	*****_1_E007	MJ023971	類似文字	J+88B4	389290	U+88B4	○
10	*****_1_E008	MJ020333	不明	J+7E23	310470	U+7E23	○
11	*****_1_E009	MJ007179	不明	J+512A	014720	U+512A	○
12	*****_1_E00A	MJ029806	不明	J+9E97	539450	U+9E97	○
13	*****_1_E00B	MJ009376	デザイン差	J+58E4	064220	U+58E4	○
14	*****_1_E00C	MJ030060	類似文字	J+9F4B	548940	U+9F4B	○
15	*****_1_E00D		同定不可能				
16	*****_1_E00E	MJ019130	類似文字	J+7A3B	282290	U+7A3B	○
17	*****_1_E00F	MJ030060	類似文字	J+9F4B	548940	U+9F4B	○
18	*****_1_E010	MJ026447	類似文字	J+9152	451550	U+9152	○
19	*****_1_E011	MJ023079	類似文字	J+85E4	366460	U+85E4	○
20	*****_1_E012	MJ026200	類似文字	J+908A	445400	U+908A	○
21	*****_1_E013	MJ007682	デザイン差		027550	U+52CE	○
22	*****_1_E014	MJ023044	類似文字	J+85CF	365090	U+85CF	○
23	*****_1_E015	MJ030185	類似文字	J+9299	161510	U+9299	○
24	*****_1_E016	MJ059779	デザイン差	J+B436		U+2AD89	
25	*****_1_E017		同定不可能				

※同定結果CSVの項目について

- 外字・・・市区町村で利用している外字コード
- MJ文字図形名・・・外字に同定されたMJ文字図形名
- 同定区分
 - 字形一致、デザイン差、類似文字、同定不可能
- 住基ネット統一文字コード、戸籍統一文字番号、UCS
 - 外字に同定されたMJ文字図形に付されているそれぞれの番号
- UCS実装・・・UCS実装の有無

全体成果物から判明した課題

文字情報基盤漢字同定リスト及び不明文字リストを確認していくと、本来氏名に利用できる文字でありながら、文字情報基盤漢字に搭載されていない文字が存在することが判った。

成果物	課題
文字情報基盤漢字同定リスト (類似文字に分類された文字)	① 氏名に利用できる新字体文字が存在する
	② 出現頻度の高い俗字が存在する
不明文字リスト	③ 漢和辞典に掲載されている文字が存在する
	④ 戸籍関連通達及び法務局通知の文字が存在する
	⑤ 変体仮名が多数存在する
	⑥ 行政用途に利用されている記号等が存在する

文字情報基盤の行政利用文字の網羅性を向上させ、外字の発生しない文字情報基盤を実現するために、次回以降の文字情報基盤漢字への文字追加検討の際には、実際の行政での利用状況等を確認しながら、今回の結果を参考にして頂きたい。



外字を提供頂いた市区町村の活用シーン①

① 外字の整理に活用

外字を提供頂いた市区町村向け成果物である同定結果リスト及びCSVには、住基ネット統一文字及び戸籍統一文字と一致した（字形一致及びデザイン差）文字について、住基ネット統一文字コード、戸籍統一文字コードを掲載したので外字の整理に活用可能です。

・住基ネット統一文字と一致した外字について

住基ネット統一文字と一致した外字は、削除し本来の住基ネット統一文字コードに変更し整理することができます。整理する際には、データの変更を伴いますので十分な確認が必要と考えますが、これら外字を全て解消できると、全国で約24万文字の外字解消につながります。

※調整中の数値です

・戸籍統一文字と一致した外字について ※調整中の数値です

戸籍統一文字と一致した外字（約7万文字）は、現に戸籍で利用されている文字が住基ネット統一文字に存在しなく外字を作成し利用している文字と推測されます。外字の整理には、十分に戸籍を確認し注意して行う必要があります。

※調整中の数値です

また、類似文字（約62万文字）と同定されている外字についても、住基ネット統一文字と一致した外字と同じ様に、外字の整理に活用可能です。

※調整中の数値です

・類似文字中の誤字及び簡体字（約47万文字）について

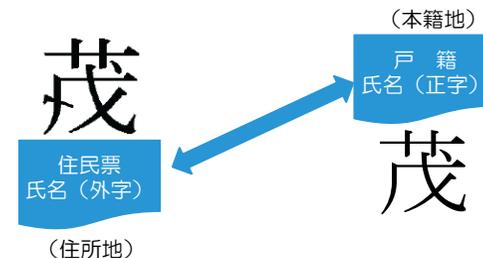
まず、類似文字中の誤字については、既に戸籍の電算化などで正しい文字に置き換わっている可能性があります。戸籍を確認して文字の確認をしたのちに、住民票等のデータを修正することで、外字解消につながります。

類似文字中の簡体字については、改正住民基本台帳法の一部を改正する法律（平成21年法律第77号附則第3条）によりその多くが正字等に置き換えられている可能性があります。仮住民票等を確認して文字の確認をしたのちにデータを修正することで、外字解消につながります。

※住基ネット統一文字と一致した外字

珺	珺	MJ017205	字形一致	住基 戸籍 UCS UCS	MJ021641	字形一致	住基 戸籍 UCS UCS
邓	邓	MJ029223	字形一致	住基 戸籍 UCS UCS	MJ022528	字形一致	住基 戸籍 UCS UCS
熙	熙	MJ016505	字形一致	住基 戸籍 UCS UCS	MJ018879	字形一致	住基 戸籍 UCS UCS
椛	椛	MJ014130	字形一致	住基 戸籍 UCS UCS	MJ021647	字形一致	住基 戸籍 UCS UCS

※類似文字中の誤字確認



※類似文字中の簡体字確認

張 → 張

外字を提供頂いた市区町村の活用シーン②

② 文字環境の見える化に活用

外字を提供頂いた市区町村向け成果物である同定結果リスト及びCSVを用いて、市区町村内の各種業務システムや市区町村外とのデータ連携に係る外字の見える化に活用可能です。

- 提供外字と文字情報基盤漢字の見える化

同定結果リストでは、提供外字と文字情報基盤漢字の対照を見える化した上でフィードバックしています。市区町村内の文字環境の見える化を行う際の参考として活用する事が可能です。

- 各種業務システムの文字環境見える化

同定結果リストと同様の内容をCSVで提供していますので、このCSVと既存住基システム：住基ネットの文字対照表、さらには既存住基システム：各種業務システムとの文字対照表などを組み合わせることで、各種業務システムと文字情報基盤漢字の対照を見える化することができ、今後の文字整理方針の策定に役立てることができます。

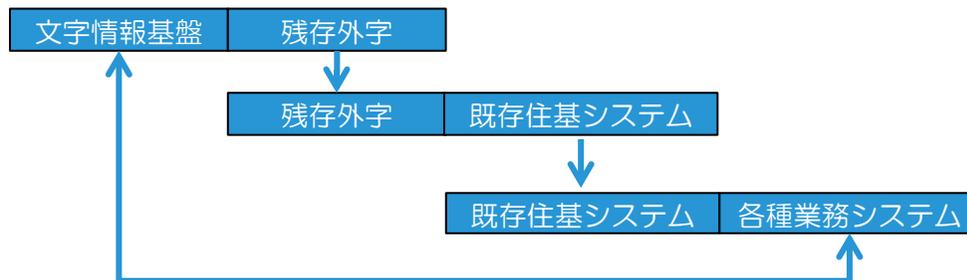
- 低解像度外字の再作成時の字形参照

文字情報基盤漢字と一致した外字には、MJ図形名が付されているので、低解像度外字の再作成時に文字情報基盤文字のデザインを参照して利用することができます。証明書などに印字する文字のデザインが簡単に綺麗にできるので、住民との文字トラブルの解消にもつながります。

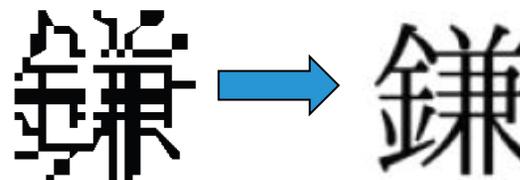
※外字の見える化

1-866/866文字) 全体結果					
外字	MJ文字図形名	補足	外字	MJ文字図形名	補足
1_E0DC	MJ021691	デザイン差	1_E0E6		
苦	苦	住基: 344780 戸籍: U+82EB UCS実装:	书		
1_E0DD	MJ006365	デザイン差	1_E0E7	MJ025036	類似文字
丽	丽	住基: 000870 戸籍: U+4E3D UCS実装:	贺	贺	住基: J+8CC0 戸籍: 416510 UCS: U+8CC0 UCS実装: ○
1_E0DE			1_E0E8		
1_E0DF	MJ009126	類似文字	1_E0E9	MJ009401	類似文字
坚	堅	住基: J+5805 戸籍: 059450 UCS: U+5805 UCS実装: ○	环		
			寿	寿	住基: J+58FD 戸籍: 064950 UCS: U+58FD UCS実装: ○

※文字情報基盤漢字と各種業務システムの文字環境見える化



※低解像度外字の解消



外字を提供頂いた市区町村の活用シーン③

③ システム間連携やシステム移行時の標準文字に活用

外字を提供頂いた市区町村向け成果物である同定結果リストおよびCSVには、文字情報基盤漢字の文字図形番号及びUCSを掲載しているため、これらの情報を用いて各種業務システムの文字環境見える化を行うことにより、各業務システムの文字を文字情報基盤漢字に関連付けすること（現状では一部について外字も併用することとなるので注意願います）により、文字情報基盤漢字をシステム間連携やシステム移行の標準文字に活用可能です。

・システム間の連携用データの標準文字

文字情報基盤漢字と各種業務システム文字コードを関連付けすることで、システム間連携の標準文字コードに文字情報基盤漢字コードを利用することができるようになり、システム間連携構築の際のベンダーロックを解消することにつながります。

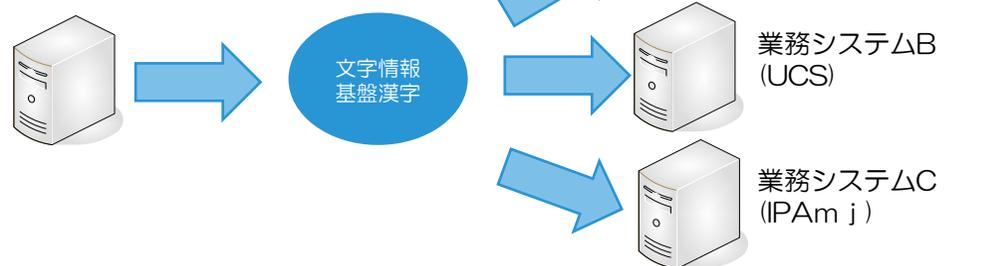
・既存システムの新システムへの移行時の標準文字

既存システムの文字コードと文字情報基盤漢字の対応付けを行ったうえで、本研究会のもう一つの成果物であるデータ移行用の中間標準レイアウトと合わせて活用することで、システム移行時に固有文字コードや固有移行レイアウト等による、ベンダーロックを解消することにつながります。

これらによって、最終目標である自治体クラウドへの円滑なデータ移行が可能となります。

※システム間連携の標準文字

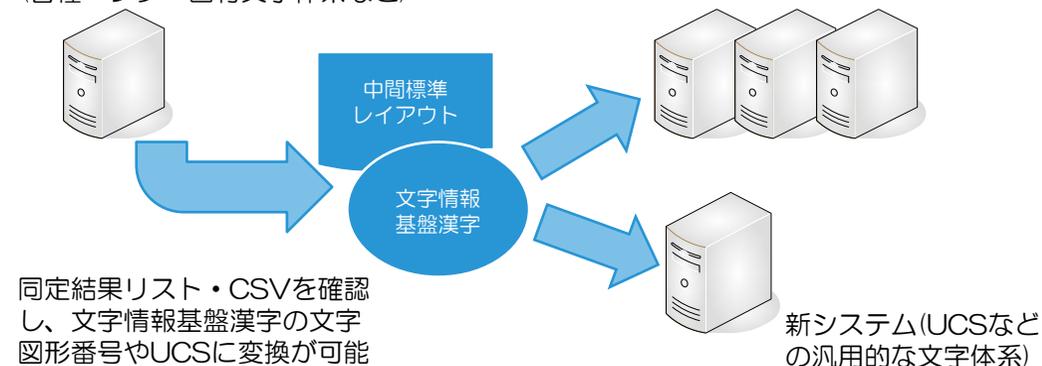
既存住基システム
(ベンダー固有文字体系など)



※システム移行時の標準文字

既存システム
(各種ベンダー固有文字体系など)

自治体クラウド
(文字情報基盤漢字)



FUJI xerox 