

文字情報基盤整備事業の進捗状況と 2015年度実施項目（案）

2015年2月

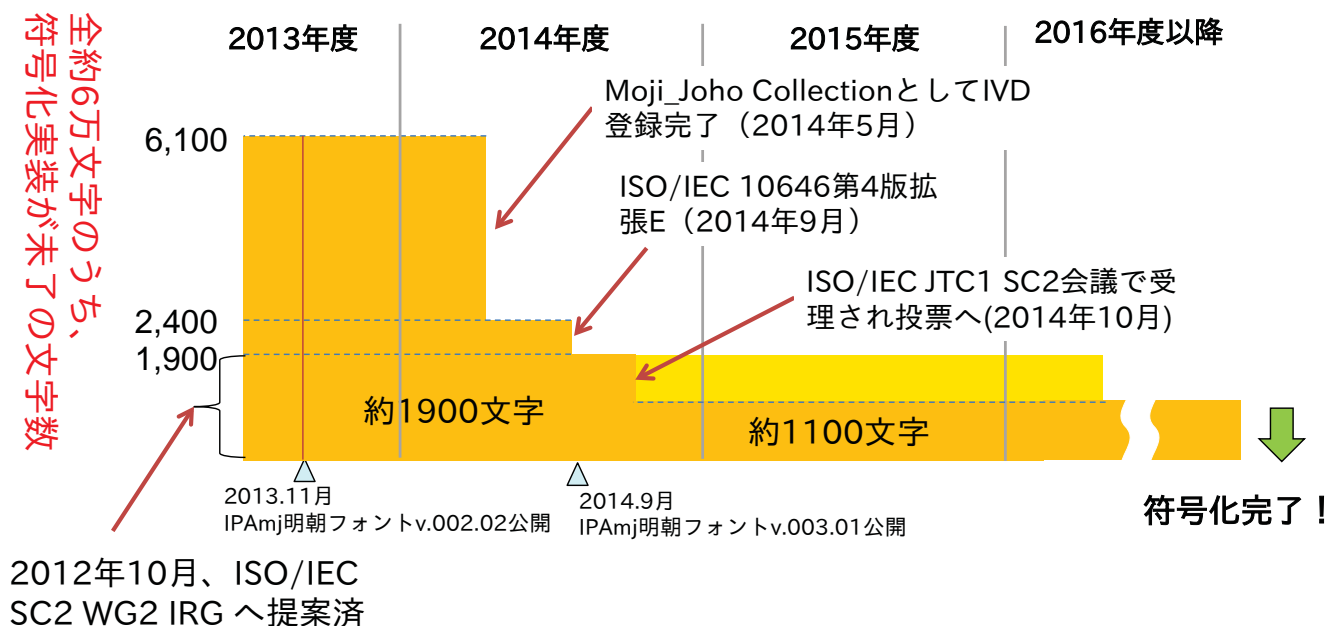
独立行政法人 情報処理推進機構

1

前回委員会（2014年2月）以降の動き **IPA**[®]

日付	事象	概要
3月13日	業界団体の設立	文字情報基盤の普及へ向けた業界活動を推進するため、「IVS技術促進協議会」を改組し、「文字情報技術促進協議会」へ
3月20日	文字情報基盤導入ガイド類の公開	IPAから公開 <ul style="list-style-type: none"> 文字情報基盤導入パンフレット 文字情報基盤導入ガイド Ver.1.0 文字情報基盤導入テクニカルスタディ Ver.1.0
3月24日	「電子自治体の取組みを加速するための10の指針」	総務省自治行政局から自治体へ向けた指針 <ul style="list-style-type: none"> 国の取り組みとして文字情報基盤を紹介
4月25日	「電子行政分野におけるオープンな利用環境整備に向けたアクションプラン」	各府省情報化統括責任者（CIO）連絡会議決定 <ul style="list-style-type: none"> 導入ガイドの参照 縮退マップの整備
5月16日	Moji_Joho collectionのIVD登録	文字情報基盤の文字図形を収録した” Moji_Joho Collection” が、Unicode コンソーシアムから公開
6月18日	セミナーの実施	幕張メッセで開催されたInterop会場にて「行政の電子化への活用が進む『文字情報基盤』」を開催
6月30日	文字情報基盤データベース構築開始	文字に係る情報をRDF形式で格納し、公開するデータベース。2015年夏に仮運用開始予定。本格運用開始は2016年4月。
9月1日	ISO/IEC 10646:2014の発行	文字情報基盤の文字のうち、約500文字が規格に追加
9月19日	MJ文字情報一覧表, IPAmj明朝のバージョンアップ	IVD登録、ISO/IEC10646の発行を反映 標準に従って使える文字が約4,000文字増加
10月3日	ISO/IEC JTC1 SC2委員会開催	<ul style="list-style-type: none"> 文字情報基盤から提案中の約1900文字のうち、約800文字が投票へ。 文字情報基盤のグリフの集合の定義を10646規格書に加えることにつき、提案が受理された → 投票処理中
1月7日	変体仮名規格案の準備状況を情報規格調査会SC2専門委員会へ報告	<ul style="list-style-type: none"> 対象文字の選定状況、規格化方針案を報告し、議論

2



参考：国際会議の日程

- SC 2 会議
 - 20th – 2015-10-19/23, Matsue, Japan, collocated with WG 2
 - 21st – 2016-10(tentative), Russia (tentative), USA(backup)
 - 22nd – 2017-10(tentative), China (pending confirmation)
- WG 2 会議
 - 64th – 2015-10-19/23, Matsue, Japan
 - 65th – 2016-10(tentative), Russia (tentative), USA(backup)
 - 66th – 2017-10(tentative), China (pending confirmation)
- IRG 会議
 - 44th - 2015-06-15/19, Seoul, Republic of Korea
 - 45th - 2015-11-16/20, HKSAR

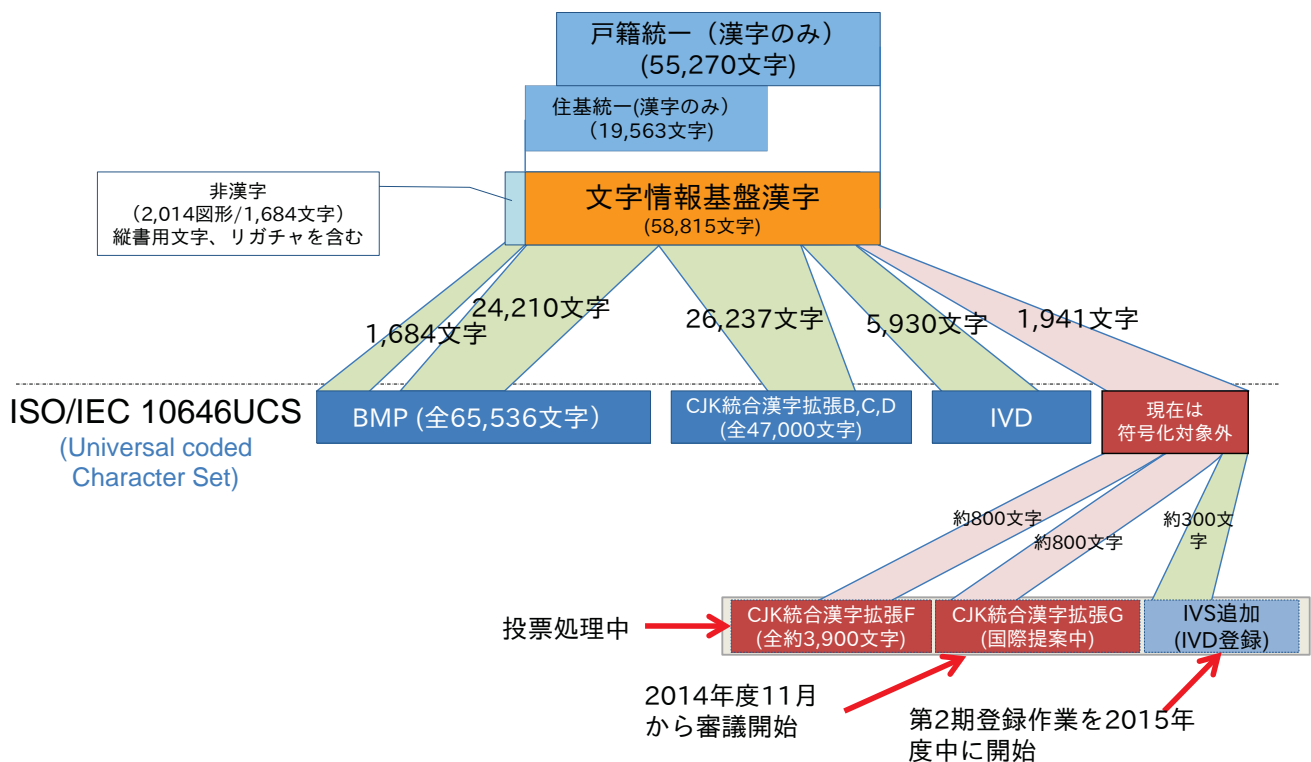
- 2014年9月19日に公開
- MJ文字情報一覧表Ver.004.01
 - 国際規格の更新に伴う掲載情報の追加
 - 行政実務で用いられる字辞典検字番号情報の追加
- IPAmj明朝フォントVer.003.01
 - 2014年9月発行のISO/IEC 10646 第4版に対応
 - 国際標準規格の最新版に対応し、前バージョンと比較して約4,200文字が利用可

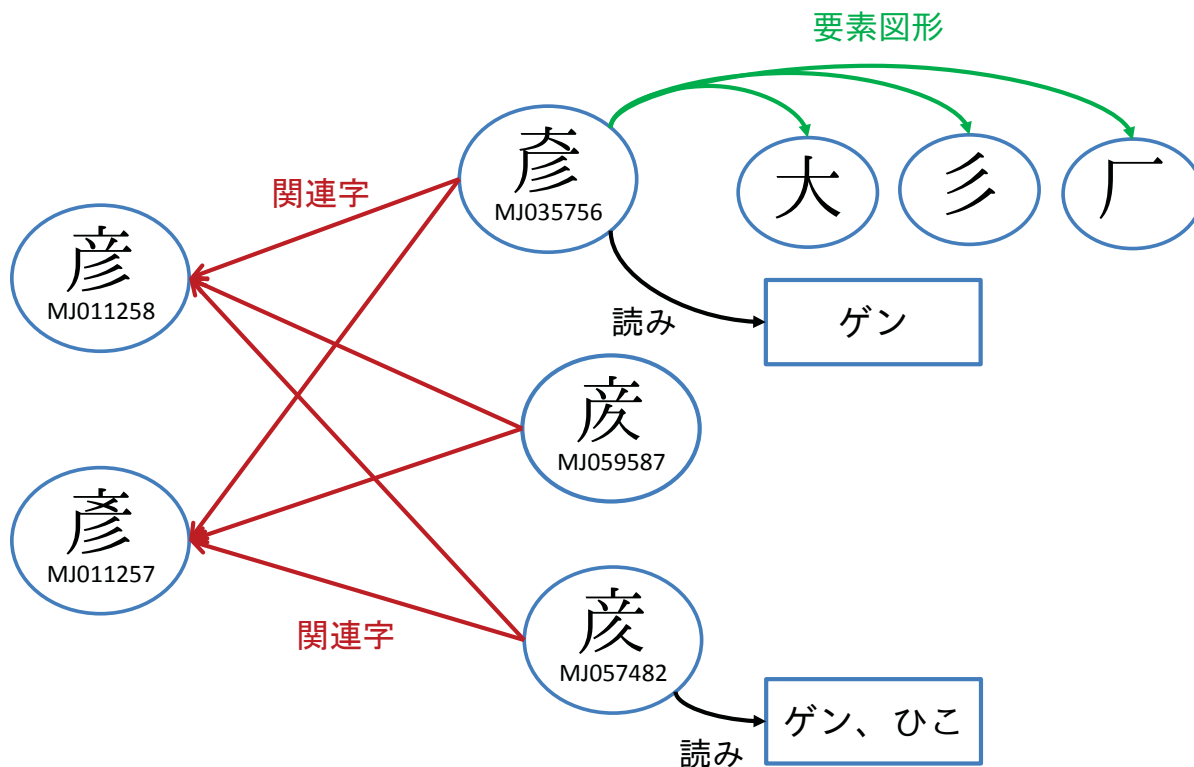
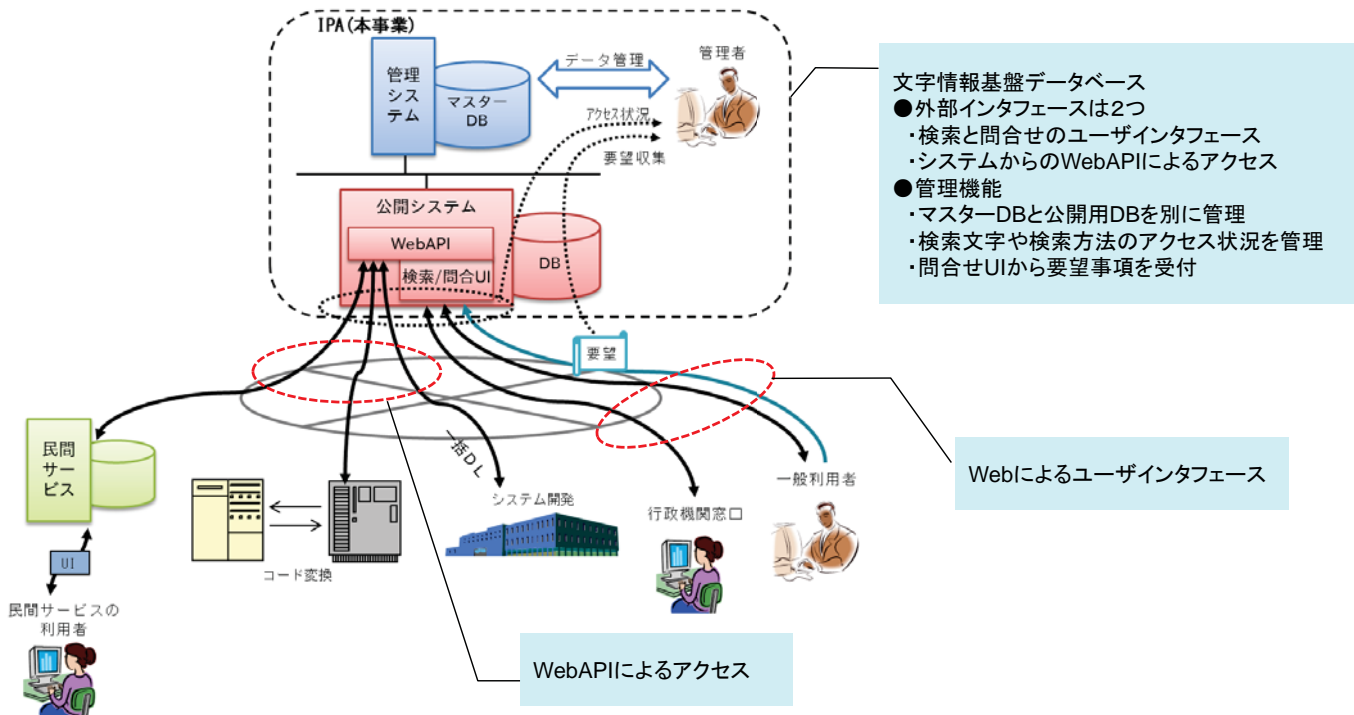
- 国際規格更新に係る項目追加
 - IPAmj明朝Ver.003.01で実装した「Moji_JohoコレクションIVS」
 - IPAmj明朝Ver.003.01で実装した互換漢字に対するSVS(Standardized Variation Sequence)
- ISO/IEC 10646第4版に伴う「対応するUCS」の更新
 - 新たに付与したもの・・・505文字
 - 符号位置を変更したもの・・・2文字
- 以下の字辞典検字番号の追加
 - 講談社新大字典
 - 角川大之源

- UCS、IVSの符号化実装の追加
 - ISO/IEC 10646 第4版に対応し、拡張Eに該当する496文字をUCS符号化実装
 - IVD Moji_JohoコレクションのIVSを10,710文字実装
 - 互換漢字に対するSVSを89文字に対し実装
- 前のバージョンと比較して、利用可能な文字が4,152文字増加

現在のUCS符号化状況

IPAmj明朝フォントVer.003.01の符号化状況





誤字俗字・正字一覧表

「正字等」漢字番号	「正字等」MJ図形名 (Subject)	アザ心差	「誤字等」漢字番号	種別と表示記号 (Predicate)	「誤字等」MJ図形名 (Object)	アザ心差
1554-1-1	MJ011752		1554-1-1-1	民事2842誤字		
1554-1-1	MJ011752	1				
1554-1-1	MJ011752	1				
1554-1-2	MJ010612		1554-1-2-1	民事2842誤字		
1554-1-2	MJ010612		1554-1-2-2	民事2842誤字		
1554-1-3	MJ010114		1554-1-3-1	民事2842誤字		
1554-1-4	MJ009693		1554-1-4-1	民事2842誤字		
1554-1-4	MJ009693	1				
1554-1-4	MJ009693		1554-1-4-2	民事2842誤字		
1554-1-5	MJ016350		1554-1-5-1	民事2842俗字(▲) MJ057990		

※他の種別 → 民事2842別字 (●)

誤字俗字・正字一覧表における整理

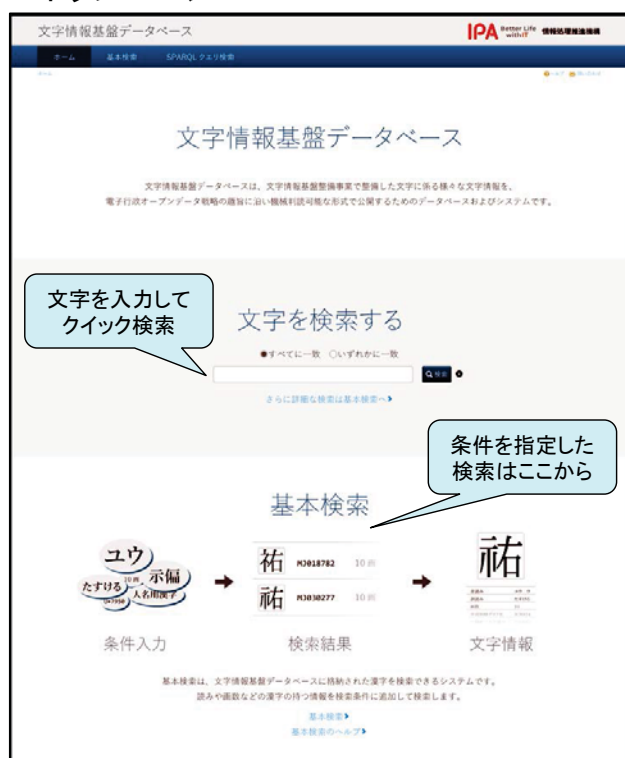
辞書における関係性の整理

代表字 異体字 関係

Subject	Predicate	Object
8351	本字	8352
8351	別体	8353
8351	別体	8354
8351	別体	8355
8351	別体	8356
8351	別体	8357
8351	別体	8358
8351	別体	8359
8351	別体	8360
8351	別体	8361
8351	別体	8362

検索機能	概要
1 基本検索	以下に示すような画数、読み、部首等の項目を指定した検索 ・ 常用漢字 ・ 音読み ・ 戸籍法上の人名漢字 ・ 訓読み ・ 総画数 ・ JIS X 0213水準 ・ 康熙部首・部首内画数
2 コード検索	以下に示すような各種コードを指定した検索 ・ MJ文字図形名 ・ 平成明朝グリフ名 ・ 戸籍統一文字番号 ・ 住基ネット統一コード ・ 入管正字コード ・ 登記統一文字番号 ・ 対応するJIS X 0213面区点位置 ・ 対応するJIS X 0212区点位置 ・ 対応するUCS ・ 大漢和辞典 検字番号 ・ 日本語漢字辞典 検字番号 ・ 新大字典 検字番号 ・ 大字源 検字番号
3 要素図形検索	利用者がMJ文字図形に含まれる要素図形を指定した検索
4 SPARQLクエリ検索	利用者が直接SPARQLクエリを指定する検索

トップページ

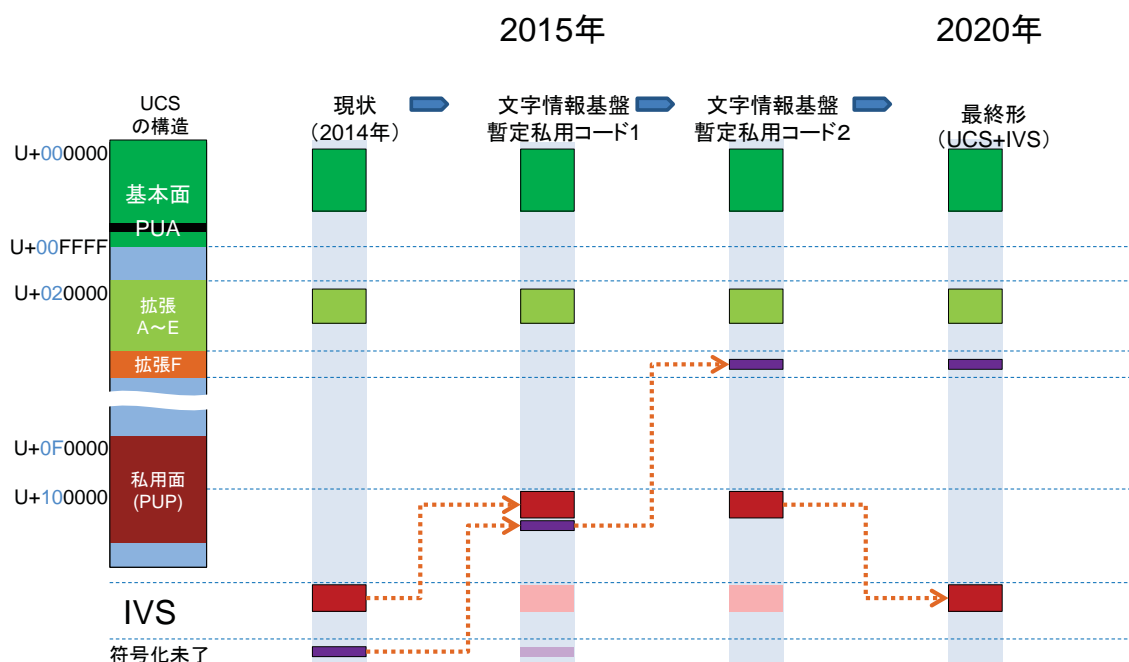


検索結果、文字詳細表示



暫定私用コードの策定と配布

- 暫定私用コードの必要性
 - 導入ガイドと同時公表のテクニカルスタディに明記
- 標準化作業と実装普及までの暫定的な施策
 - UCSにおいて未符号化のMJ文字図形約1900文字を暫定的に私用面にマッピング
 - IVDにMJコレクションとして登録済みのMJ文字図形約6000文字を暫定的に私用面にマッピング
- 文字情報技術促進協議会が策定・配布
 - 利用に当たっては、対象組織の限定等を行わない
 - 私用文字領域利用の際は、送受信間の相互了解が必須
 - 利用に当たっての副作用（コードを共有していない組織への送信制限の必要性、暫定私用コード利用期間終了時の移行処置の必要性等）については、利用者の自己責任を徹底



- 行政実務での必要性
 - 新戸籍法の施行（1948年）以前に生まれた人の名に使用例
 - 現行戸籍法下でも必要。戸籍法施行規則の改正で除籍謄本の保存期間が150年になったので、今後も相当期間必要
 - 現在の戸籍統一文字には168文字の変体仮名に固有IDが付与されている
- 学術用途での必要性
 - 高田智和氏（国立国語研究所）が中心となり、印刷業界、学術関係者からの聞き取り調査を実施
 - 日本語学、歴史学、書誌学、図書館情報学など、人文科学系の研究分野において、情報通信環境下での変体仮名利用への根強い要望があった
 - 研究者有志による討議検討を経て、必要変体仮名300文字弱を選定
- 情報規格調査会SC2専門委員会としての変体仮名レポートリーの選定と国際提案
 - 行政実務用と学術用を整理統合した300文字強を提案
 - 音価と字母の組に固有のIDを付与する方向で検討中



今後の計画(案)

日付	事象
2014年度中	標準提案へ向けた変体仮名文字セット、デザイン案の確定
2015年8月	MJ文字情報一覧表へ追加 (戸籍文字との対応関係確定)
2015年9月	ISO/IEC JTC1/SC2総会及び同WG2会議への提案
2017年度	標準化完了を目指す

- 変体仮名の国際規格化提案
- 提案済みの漢字2000文字標準化フォローアップ
- 文字情報データベースの構築と稼働
- 業界団体との連携
- 普及啓発