



INFORMATION-TECHNOLOGY PROMOTION AGENCY, JAPAN

文字情報基盤構築事業の概要

2011年7月22日

独立行政法人情報処理推進機構

国際標準推進センター

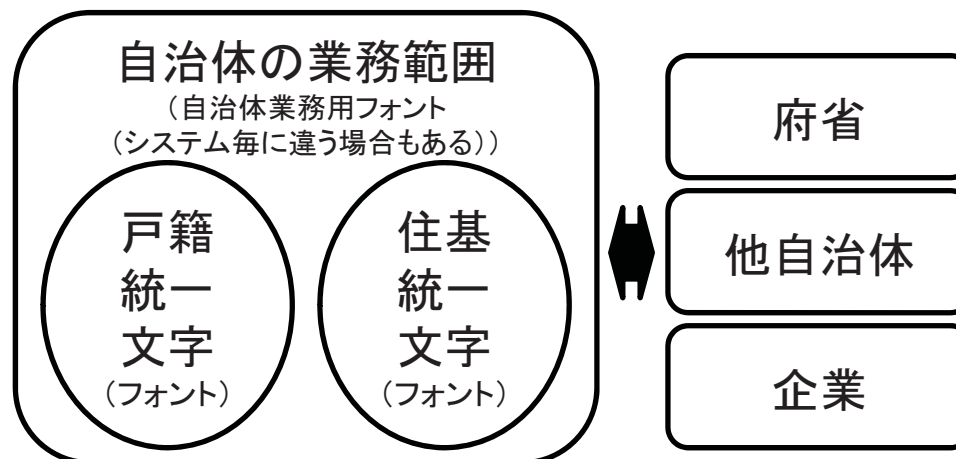
文字を巡る環境

- 多くの関係者の要望を整理すると以下の通りである。
 - 住民
 - ・ 個人のアイデンティティとして正確に表記すること
 - ・ 行政コストを下げるべき
 - 企業
 - ・ 業務効率を落とさないこと
 - ・ 行政機関の電子申請などの要求水準に対応できること
 - ・ 現在一般に使われる以外の文字がむやみに広がるのは困る
 - 行政機関
 - ・ 住民を正確に確認できること(戸籍業務、証明業務等)
 - ・ 行政効率を落とさないこと
 - 外字作成コストが減った代わりに、氏名の同定コストがかかったのでは意味がない
 - ・ 誰もが扱える文字であること
 - ・ すべての文字を使えるようにするのではなく、業務により文字の範囲を制限することも重要
 - 情報転送を受ける行政機関
 - ・ 正確な文字情報が伝わること
 - 文字ビジネス関連企業
 - ・ 民業圧迫になるのは困る

公的主要フォントと一般に使われる文字

■ 各種フォントセットは存在するが、利用制限が設定されていたり、広く活用することが困難である。そのため、行政機関内に複数のコード体系が存在している。

- 戸籍統一文字 **56040字**
 - ・ 戸籍業務に利用範囲を限定
- 住基統一文字 **21039字**
 - ・ 本庁および支所、その他定める範囲
 - ・ 市民病院や図書館では使用できない
 - ・ 情報交換を目的とする



一般の社会生活において、現代の国語を書き表す場合の漢字使用の目安

常用漢字 **1945字** (新常用漢字試案 2136字)

印刷用文字

表外漢字 **1022字**

人名用漢字(戸籍法施行令)

人名用漢字 **2930字**

常用漢字 1945字

漢字の表(一) (常用漢字の異体字でないもの) 776字

漢字の表(二) (常用漢字の異体字であるもの) 209字

情報交換用文字

JIS **10038字**

第一水準 2965字

第二水準 3388字

第三水準 1249字

第四水準 2436字

Unicode 3.1ではJIS X 0213が対応
日本語の文字コードで運用する必要性の高い文字。計11,233字

日本漢字能力検定

1級 **約6000字**
 準1級 約3000字
 2級 約1945字 他人名用漢字

上記が読み書き可能

従来の文字セットの意義

- 本事業で、4つの文字セットがシームレスに連携できるようになる。
- それぞれの意義・経緯を踏まえて利用方法を考える必要がある

邊邊辺 …

邊邊邊

戸籍統一文字(56040字)

戸籍のオンライン手続に使用することを目的として整理した文字
(辞書をベースに整理)

邊邊邊 …

邊邊邊

住民基本台帳ネットワーク統一文字(21039字)

多くの住民が氏名に使う文字を整理

邊邊

JIS漢字コード(10050字)

実用上の情報交換の必要性から、出現頻度などを元に文字を選定

辺

常用漢字(2136字)

法令、公用文書、新聞、雑誌、放送など、一般の社会生活において、
現代の国語を書き表す場合の漢字使用の目安を示すものである。
(個々人の表記にまで及ぼそうとするものではない。)

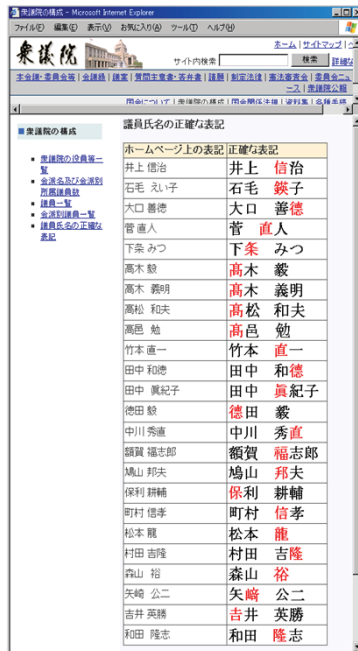
多様性・アイデンティティ



交換性・意味

外字関係者数の推定と組み合わせ例

- 外字氏名の方がどのくらいいるかの正確な数字は計測されていない。
- 氏名表記情報が正確に公開されている衆議院議員名簿を元にとすると、国民の5%が氏名内で外字を使用していると推定される。



議員氏名の正確な表記	ホームページ上の表記	正確な表記
井上 信治	井上	信治
石毛 い子	石毛	咲子
大口 善徳	大口	善徳
菅 直人	菅	直人
下条 みつ	下条	みつ
高木 毅	高木	毅
高木 義明	高木	義明
高松 和夫	高松	和夫
高邑 勉	高邑	勉
竹本 直一	竹本	直一
田中 和徳	田中	和徳
田中 真紀子	田中	真紀子
徳田 毅	徳田	毅
中川 秀直	中川	秀直
額賀 福志郎	額賀	福志郎
鳩山 邦夫	鳩山	邦夫
保利 耕輔	保利	耕輔
町村 信孝	町村	信孝
松本 龍	松本	龍
村田 吉隆	村田	吉隆
森山 裕	森山	裕
矢崎 公二	矢崎	公二
吉井 英勝	吉井	英勝
和田 隆志	和田	隆志

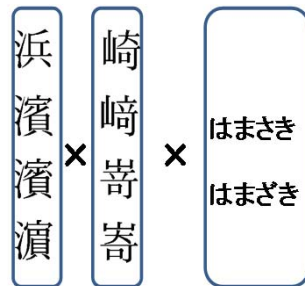
衆議院議員480人中

正確な表記として画像の文字を公開している議員 24名(5%)
 そのうち名字に画像の文字を使用している議員 9名(2%)
 (参議院で同様に検証すると、氏名5%と名字3%)
 (2009年10月20日)



国民の5%が氏名の表記に関する課題を持っていると推定される
 また、2%の人が名字の表記に関する課題を持っていると推定される

注: 常用漢字表前書きにある「字体の違いではなくデザインの差」と言う
 視点で各文字を見ると数は少なくなる



4 × 4 × 2 = 32通り

浜崎という名前は、「浜」で4通り、「崎」で4通りの文字があり、更に読み方が2種類あるので、32通りの組み合わせが考えられる。



同じ漢字氏名にも多くのパターンがあり、検索をすることが困難である

- 実際の電子申請システムでは、利便性やシステムの接合性などを考慮して、文字の制限を実施している(外字がないことも一因)
 - e-Gov電子申請システム利用規約(使用可能な文字)
 - ・ 第10条 本システムにおいて使用可能な文字は次の各号に掲げる文字とする。
 - ・ 一 JIS X 0201として規格化されている英数字及び記号を含む1バイト文字。
 - ・ 二 JIS X 0208として規格化されている2バイト文字。
 - ・ 三 JIS第一水準漢字及びJIS第二水準漢字。
 - e-Taxで使用できる文字コードは、JIS X 0221をUTF8で符号化した以下の文字コードが使用できます。
 - ・ 「基本ラテン(JIS X 0201と互換性のあるもの。ただし、文字タブ(0009)、改行(000A)及び復帰(000D)以外の制御文字(0000～001F、007F)を除く)」、「平仮名」、「片仮名」、「CJK統合漢字」、「CJK互換漢字」、「CJK用の記号及び分音記号」、「半角形・全角形(ただし、半角カナ(FF66～FF9F)を除く)」、「ラテン補助(ただし、制御文字(0080～009F)を除く)」、「矢印」、「一般句読点」、「罫線素片」、「幾何学模様」、「基本ギリシャ」、「キリール」、「数字記号」、「数字の形」、「囲み英数字」、「囲みCJK文字／月」「CJK互換文字」
 - ・ なお、e-Taxで使用できない文字については、利用者の判断により、代替文字を選択していただくこととなります。
 - 警視庁 警察官・職員採用 インターネット申込みFAQ
 - ・ Q3-1 氏名が常用漢字ではないのですが、外字入力できますか？
 - ・ A3-1 外字は入力できませんので、申込み時は正しい氏名に近い漢字またはひらがなで入力してください。第一次試験当日には受験票の氏名を赤字で修正したものを持参してください。

※CJK統合漢字(20902字):日本中国・台湾韓国の各漢字の間で統合化操作が行われたもの

医師等資格確認サービスの例(厚生労働省)

医師等資格確認検索

http://licenseif.mhlw.go.jp/search/

こちらのブックマーク バーにブックマークを追加すると簡単にページにアクセスできます。

厚生労働省 Ministry of Health, Labour and Welfare

医師等資格確認検索画面へ

■目的

この医師等資格確認検索システムは、厚生労働省に現在登録されている医師及び歯科医師(以下「医師等」という。)のうち、医師法(歯科医師法)に規定する2年に1度の届出を行っている医師等について、漢字等の氏名を入力することにより検索を可能とするものです。

■検索結果

この検索システムでは、検索結果として、職種(医師・歯科医師の別)、氏名、性別、登録年及び該当する者に限り行政処分に関する情報(処分の種類及び期間)がご覧いただけます。
 なお、行政処分に関する情報については、医業(歯科医業)停止期間が経過した後は、再教育研修が未修了の場合を除き、当該情報は表示されません。
 ※注 行政処分により医業(歯科医業)停止中の医師等は、医業(歯科医業)が禁止されているため、一切の診療行為ができません。

■検索方法

医師等の検索は、氏名が揃っている場合のみ可能です。氏名の片方のみ又はフリガナでは、検索を行うことができません。なお、検索に当たっては、姓と名の間に空白を設ける必要があります(例 厚生 太郎)。また、通称名が登録されている場合(外国籍を持つ医師等は、通称名を利用されている場合があります)は、通称名での検索も可能です。ただし、検索結果に本名と通称名を同時に表示することはありません。

氏名に正字以外の字(例「高」「廣」など)を用いている医師等の検索は、添付の異体字リストを参考に、対応する正字で検索することができます。
 また、対応する正字がない字(添付ファイルの外字リスト)については、該当する字の箇所を「?」として検索することができます(例 厚? 太郎)。
 (異体字リスト) (外字リスト)

■留意事項

この検索システムでは、次のいずれかに該当する医師等は検索できません。
 ① 医籍(歯科医籍)の氏名に対応しているため、旧姓等の使用により、登録名と使用している氏名が異なる医師等
 ※ 訂正の手続きが済んでいない医師等への御案内 → (医籍訂正)(歯科医籍訂正)

○異体字検索漢字リスト

正字	異体字等	正字	異体字等	正字	異体字等
亜	亞	宇	宇	綠	綠
悉	悉	雨	雨	媛	媛
庵	庵	羽	羽	燕	燕
				淵	淵
章	章	英	英		
連	連	榮	榮	桜	櫻
為	為	衛	衛	翁	翁
飴	飴	曳	曳	応	應
蔚	蔚	穎	穎	横	横
巷	巷	益	益	歐	歐
逸	逸	悅	悅	鶯	鶯
陰	陰	円	圓	奥	奧
陰	陰	延	延	温	溫
隱	隱	遠	遠		
		遠	遠	花	花
				花	花

正字	外字
?	邦 源 璣 閃 備 逸 原 葉 嚴
	邱 鑑 覺 格 歡 靜 偶 殷 藤
	筆 襪 蘇 華 鈞 潘 莊 捷 兩
	爾 裕 承 張 洵 翰 純 甫 賈
	權 薛 勳 克 范 楊 淳 述 毓
	畢 蔚 董 徐 翰 吉 呈 穆 倫
	葛 蘭 緝 為 禹 鼎 齡 遵 梁
	程 董 蘊 估 淵 愈 彥 芬 狀
	會 祿 蔽 皓 兼 均 俊 周 茂
	廖 范 祿 照 杏 昭 黃 祺 祿
	蘇 兌 斐 格 霄 為 簡 顏 樸
	蔣 具 鮮 祚 珺 祐 高 運 啟
	綏 高 裕 勳 鎮 嘉 廣 香 冰
	翁 懷 述 靖 米 文 萬 滿 燕
	趙 敬 彥 藏 蔣 英 閃 政 鄭

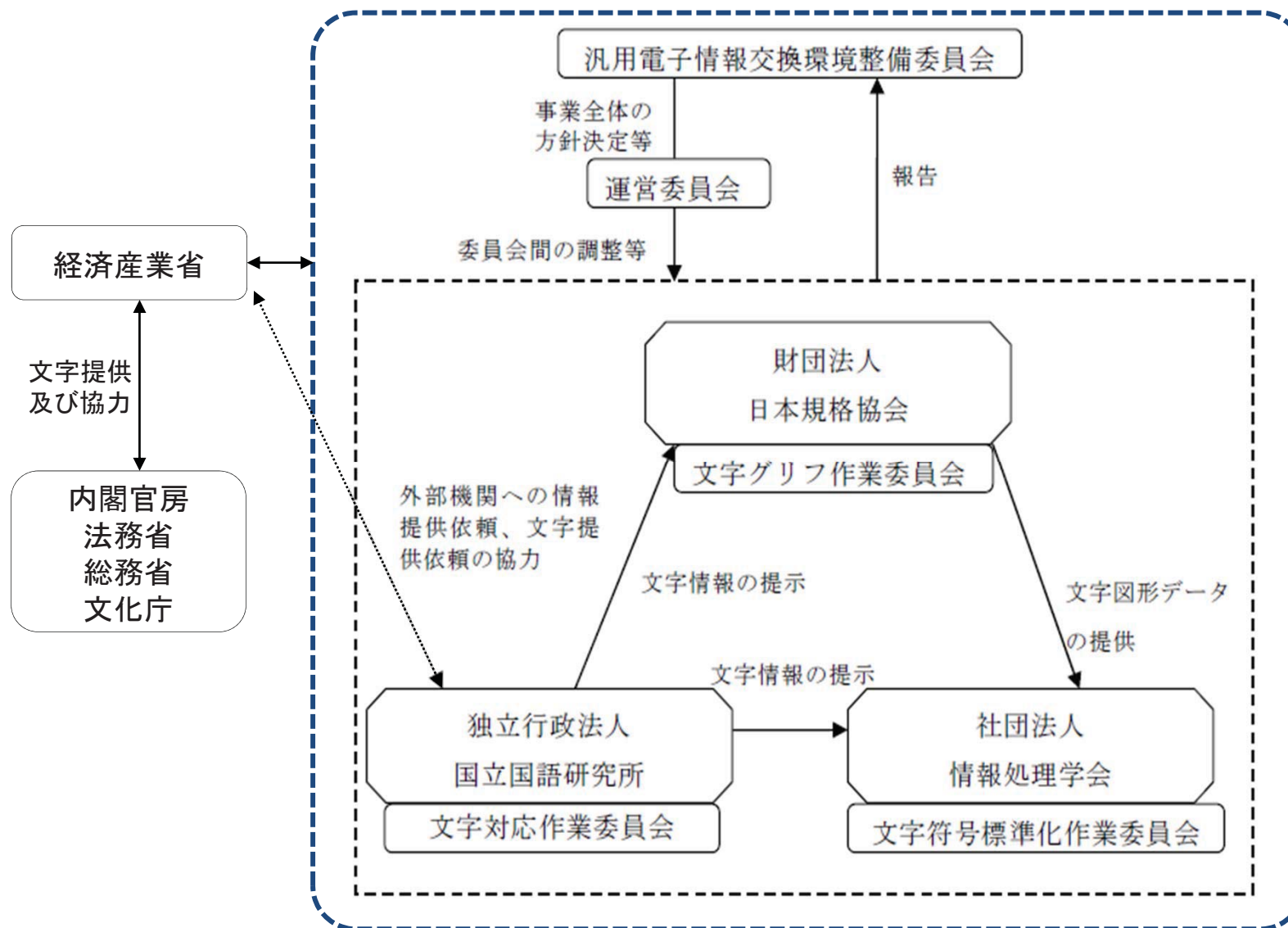
これまでの経緯

文字に関するこれまでの計画

- これまで多くの計画や戦略で目標が定められてきたが、文字の多さ、同定の困難さ、権利関係の複雑さ等から実現は大幅に遅れた。
 - 新たな情報通信技術戦略 工程表 平成22年6月22日 高度情報通信ネットワーク社会推進戦略本部
 - ・ 文字も検討課題。
 - 次世代電子行政サービス基盤等検討PT中間報告書 平成21年12月21日
 - ・ 「公共サービス連携基盤(仮称)」の中では、統一された文字コードで外字を扱えるようにするために、文字図形番号を用いる。また、文字の重複の排除や対象とする文字の範囲など、標準とする文字に関する基本的なルールを整理した上で、文字情報データベースを用意する。
 - IT政策ロードマップ 平成20年6月11日 IT戦略本部決定
 - ・ 国・地方公共団体等の相互連携を推進するため、文字コードやデータ形式の標準化を推進するほか、諸外国における技術参照モデル(TRM)の設定等も参考に、オープンな標準の採用を促進する。
 - IT政策パッケージ-2005 平成17年2月24日 IT戦略本部決定
 - ・ ③文字コード標準化の推進とデータベースの運用(経済産業省及び関係府省)
 - 住基ネット統一文字、戸籍統一文字の重複関係を整理、体系化したデータベースを2005年末までに開発する。また、現在構築中の登記統一文字について、2006年度末までに、このデータベースとの整合性を図ることとする。さらに、データベースの具体的な運用方法について関係府省で検討を行い、2005年末までのできる限り早期に結論を得る。
 - 電子政府構築計画2003年 平成15年7月17日 各府省CIO連絡会議決定
 - ・ (4)文字情報・コードの整備
 - 経済産業省及び関係府省は、行政情報化の共通基盤の一環として、将来的な国際標準との整合を視野に入れつつ、官民が汎用的に利用できる文字情報データベースの整備を進め、2005年度末(平成17年度末)までに構築する。
 - e-Japan 戦略 II 平成15年7月2日 IT戦略本部決定
 - ・ ⑨行政システム間をはじめ、民・官でのデジタル情報の自由な交換・共有のため、現在外字として利用されている文字について、2003年度中に公開用文字情報データベースの試験的運用を開始し、2005年度までに当該データベースを構築し、これを元に国際整合性も勘案した文字コード規格を整備する。
 - ・ ③アジアにおいて、コンテンツ流通のための権利処理・権利保護等の問題への対応、知的財産侵害品流通の防止、文字コードの標準化を進めると共に、放送・出版等で作成されたコンテンツや、美術館・博物館等の所蔵品のデジタル化及びアーカイブ化の推進、コンテンツの多言語化の推進により、アジア全域におけるコンテンツの積極利用を促進する。
 - 行政情報化推進基本計画 平成6年12月25日 閣議決定
 - ・ 情報システムにおいて使用している外字について、JIS第3水準及び第4水準の制定を待つてその解消を図るとともに、この場合においても残る可能性のある外字について、交換のルールを策定する。

汎用電子情報交換環境整備プログラム(H14-20)

- 5府省連携で本格的にコードの整理を実施した。



22年度事業の概要

文字情報基盤の整備にあたり、国内の関係機関が参加する委員会を構成 (2010年9月—2011年3月)

委員長 林 史典 聖徳大学・聖徳大学短期大学部人文学部長
大学院言語文科研究科長人文学部日本文化学科長教授

委員参加組織

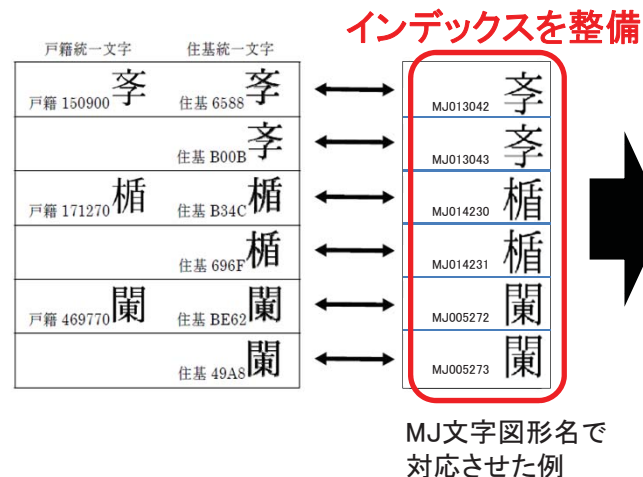
財団法人日本規格協会
社団法人情報サービス産業協会
社団法人コンピュータソフトウェア協会
社団法人電子情報技術産業協会
財団法人日本情報処理開発協会
大学共同利用機関法人人間文化研究機構国立国語研究所
社団法人日本印刷産業連合会
社団法人日本印刷技術協会
社団法人情報処理学会情報規格調査会SC2
CIO補佐官技術タスクフォース
電子政府推進員
川口市
内閣官房
総務省
法務省
文化庁
経済産業省

オブザーバ参加組織

国会図書館
警察庁
法務省
国税庁
厚生労働省
独立行政法人国立公文書館
日本年金機構
社団法人日本情報システム・ユーザー協会

22年度事業の概要

- 文字を巡る課題と解決案の整理 (ロードマップの整備)
- 電子行政を推進するための フォントと文字情報一覧表の整備
 - 行政情報の適正で効率的な管理や手続きを行うために統一された文字コードに基づくフォントの整備や利活用のための文字情報基盤を整備。行政機関のみならず、この基盤を民間との連携も視野に入れることで社会的コストを大幅に削減可能。



府省、自治体内
の他のシステム
の文字体系との
連携も可能

- 電子申請で、正しい氏名が送信できる
- 他システムに送信した文字が、相手のシステムでもどの文字が正確にわかる
- システム更新時に、文字変換が容易になる
- フォントをダウンロード可能にする事で、新規外字作成コストが削減される

■ 文字管理体系の検討

82A6

芦

芦田さんを一括して検索

芦 芦 芦 芦

1 2 3 4

異体字による氏名が一括して検索可能になる
名寄せが簡単になる

- 本事業では、行政機関として情報処理をするために必要となる文字情報基盤を整備。本事業の実施に当たっては、民間フォント・ビジネスとの棲み分けという観点から、国として取り組む必要性の高い国民生活に必要な最低限の文字について、基盤としてインデックスを整備するもの。
 - 戸籍、住民票に必要な文字を対象
 - 主要サイズの明朝体のみを整備
 - 文字の美しさは公用文として必要十分なレベル
 - インデックスで調べた後の各機関が外字を作成するのは大変であるので、そのフォントも提供

- 近年の経済社会における活動では、電子的手段による情報交換が行われる機会が増加傾向にあり、個人から企業・行政機関に、企業等から行政機関等に情報転送が行われる場面も多い。このため、行政機関のみによる独自の文字体系を持つだけでは経済社会全体としての効率について十分な改善が見込めないことから、本事業の成果物については、個人・企業で活用をさまたげない。

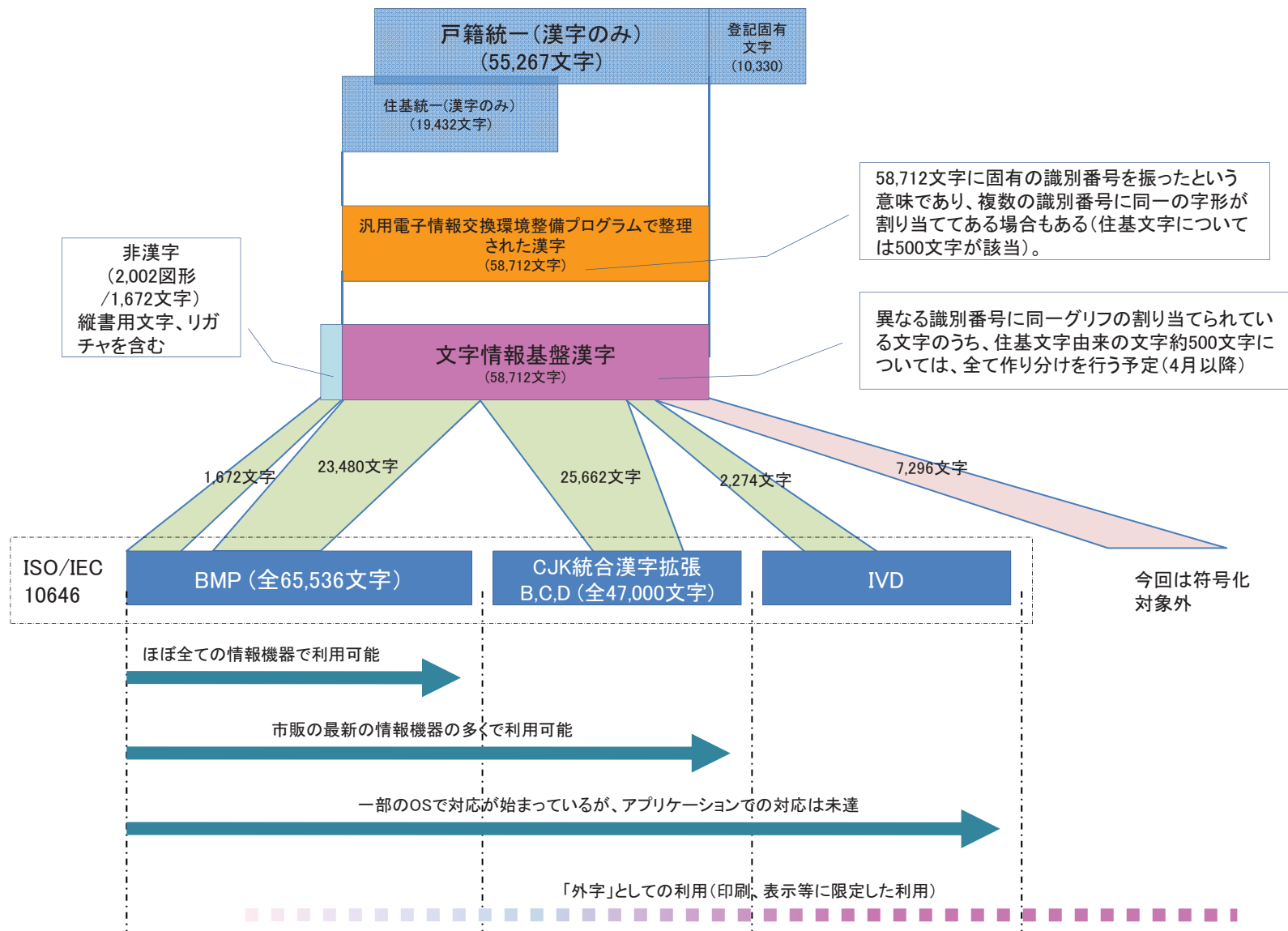
- 基準となる文字セットと検索環境がインデックスとして整備されることにより、国内の情報流通を促進し、それに伴いフォント・ビジネス及び周辺ビジネスの活性化が期待できる。

22年度成果物の状況

- ・ 調査報告書
- ・ フォント(検証版)
- ・ 文字情報一覧表(検証版)
 - － ダウンロードによる提供
 - － 検索システム(試行版)の提供

公開サイト: <http://ossipedia.ipa.go.jp/ipamjfont/>




IPAmj明朝フォント(検証版)



IPAmj明朝フォント(検証版)

・ サンプル

邊 MJ026200	邊 MJ026199	邊 MJ026201	邊 MJ026202	邊 MJ026203
邊 MJ026204	邊 MJ026205	邊 MJ060240	邊 MJ026207	邊 MJ058870

-  JIS例示字形に近い字形
-  IVDに登録済の字形
-  符号化未完了の字形

文字情報一覧表(検証版)

- 成果物の文字情報一覧表を文字図形入りHTML形式で検証版として公開

- 一覧表の項目
 - 文字図形
 - MJ文字図形名
 - 各種コード
 - 辞書情報
- 等

字形 link	MJ文字図形名	UCS	X0213	X0212	住基ネット統一文字コード	戸籍統一文字番号	登記統一文字番号	部首	内画数	読み	漢字施策	汎用電子整理番号	平成明朝	大漢和	康熙字典番号	新大字典番号	新大字典ページ	新大字典文字の種類	JIS X0213 包根連番	JIS X0213 包根区分	UCS 符号化	IVS 符号化	IVS
財	MJ025001	8CA1	1-26-66		8CA1	415830	00415830	154	3	10	ザイ・サイ・たから・わずか	常用漢字	25001	JA2666	36664	1204.110	16257	2209	正字	0		○	
貢	MJ025002	8CA2	1-25-55		8CA2	415840	00415840	154	3	10	コウ・タ・みづみづ	常用漢字	25002	JA2555	36665	1205.010	16258	2209	正字	0		○	
貧	MJ025003	8CA3				415850	00415850	154	3	10	トク		25003	JP8CA3	36666	1205.020	16259	2210	正字			○	
賤	MJ025004	8CA4	2-89-11	63-09	8CA4	415860	00415860	154	3	10	イ・シ・ジ		25004	JD8911	36667	1205.030	16260	2210	正字	0		○	
賤	MJ025005	8CA5				415940	00415940	154	4	11	コウ・ゴウ		25005	JP8CA5	36673	1205.060						○	
賤	MJ025006	8CA6				415960	00415960	154	4	11	ガン		25006	JP8CA6	36675	1205.080						○	
貧	MJ025007	8CA7	1-41-47		8CA7	415980	00415980	154	4	11	ヒン・ビン・まずしい	常用漢字	25007	JA4147	36677	1205.100	16262	2210	正字	0	○	○	8CA7_E0102
貧	MJ025008	8CA7	1-41-47		BBC2			154	4	11	ヒン・ビン・まずしい		25008	FT1666	36677	1205.100	16262	2210	正字	0		○	8CA7_E0103
貨	MJ025009	8CAB	1-18-63		8CAB	415990	00415990	154	4	11	カ・たから	常用漢字	25009	JA1863	36678#	1205.120	16263	2211	正字	0	○	○	8CAB_E0102
化																		正					

文字情報一覧表(HTML版)

http://ossipedia.ipa.go.jp/ipamjfont/moji_list.html

文字情報閲覧システム(試行版)

- 文字情報一覧表を簡易に閲覧できる閲覧システム(試行版)を公開中
 - 検索機能や一覧情報を今後拡充予定



IPA文字情報検索システム(試行版)
http://ossipedia.ipa.go.jp/ipamjfont_glyph/

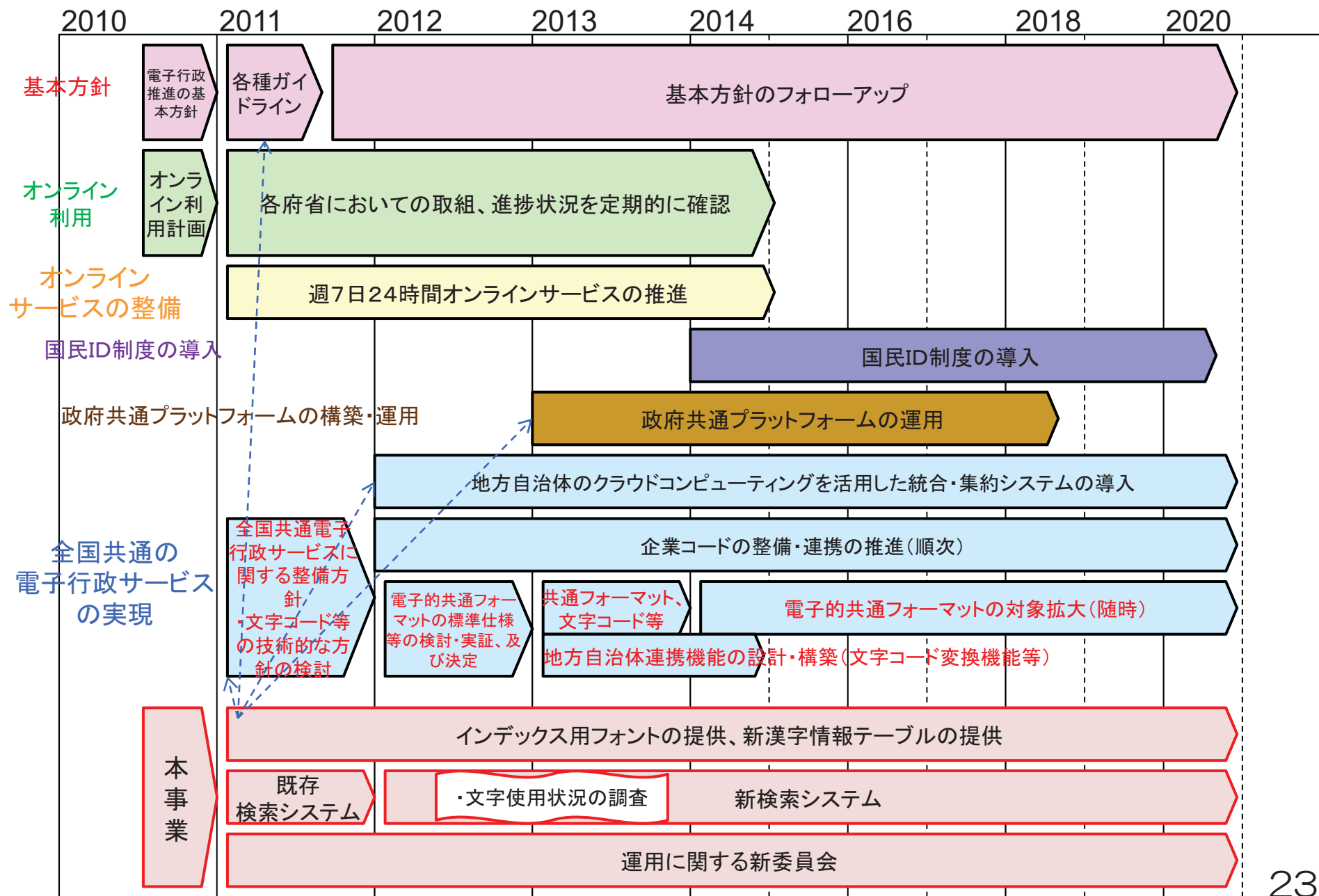


文字情報基盤23年度事業の概要

文字情報基盤

- 政策的運営体制
 - 文字情報基盤推進委員会
 - 内閣官房IT室、経済産業省、IPAが共同で設置
 - 有識者、省庁、自治体関係者、WG主査により構成
- 実務的運営体制
 - IPA
 - 文字、文字情報、運用ガイド等の整備作業
 - 実証実験など普及活動の実施
 - 専門家を集めたWGを設置して議論

【参考】新たな情報通信技術戦略 工程表(文字関連主要事項)



電子行政推進に関する基本方針に係る提言

〔 高度情報通信ネットワーク社会推進戦略本部(IT戦略本部)
電子行政に関するタスクフォース 2011年7月 〕



第4 重要施策の推進

2. 国民ID制度、企業コード等

(1) 社会保障・税に関わる番号制度と国民ID制度

.....

(2) 企業コード

.....

(3) 文字情報基盤の活用

行政情報の適正な管理や行政機関間の効率的な情報連携を実現するため、行政機関が利用できるフォントや文字情報等の文字情報基盤を活用し、文字コードやいわゆる外字等に係る問題の解決を図っていくこととし、氏名等の正確性と社会全体での利便性・効率性を考慮した仕組みを検討していく。

23年度事業実施状況(1/5)



- 作字の状況 (1/3) -

- 追加文字(約500文字)の作字
 - 住基ネット統一文字では字形を区別しているが、IPAmj明朝フォント検証版では同一の字形となっている文字(約500文字)の作り分けを行う。
 - IPAmj明朝フォント検証版は汎用電子情報環境整備プログラムの成果物^{注)}に対して同定を行った。
 - 汎用電子情報環境整備プログラムは、平成明朝体のデザインポリシーに従ってデザイン統合をしている。このため、住民基本台帳ネットワーク統一文字(以下、住基文字)の複数の字形に対し、IPAmj明朝の字形が同一となっている場合が存在する(「MJ文字図形名」は異なっている)。

注) 汎用電子情報交換環境整備プログラムの成果報告書別冊 文字対応作業委員会編「文字一覧表」

23年度事業実施状況(2/5)

- 作字の状況 (2/3) -

追加文字の整理

- 住基文字:平成明朝=n:1 (n=2 or 3)になっているケースを抽出した。
 - ・ 平成明朝3重複(n=3) 合計:17組(作字対象文字数 17×2=34文字)
 - ・ 平成明朝2重複(n=2) 合計:393組(作字対象文字数 393文字)
- 上記の結果、作字対象は427文字
- 平成明朝3重複(n=3)の例

新規作字	IPAmj明朝	平成明朝	住基文字	UCS
	 MJ000027	 JD0119S	 342E	342E
	 MJ000028	 JD0119S	 BB0D	342E
	 MJ000029	 JD0119S	 BB0E	342E

23年度事業実施状況(3/5)

- 作字の状況 (3/3) -

- 平成明朝2重複 (n=2)

新規作字	IPAmj明朝	平成明朝	住基文字	UCS
				4ED7
				4ED7

スケジュール

	5月	6月	7月	8月	9月	
MJ文字情報一覧表の整理	→					
確認・方針確定作業		→				
作り分け作業			→			
SVG生成				→		
新規符号化					→	
バグ修正					→	
フォント検証作業					→	
MJ文字情報一覧表修正				→		

23年度事業実施状況(4/5)



- 符号化の状況 -

- ・ 汎用電子フォローアップ
- ・ CJK統合漢字拡張D
 - 2011年3月にISO/IEC 10646第2版出版
 - 日本提案として107字を収録
- ・ CJK統合漢字拡張E
 - 2011年11月のIRGミーティングから審議再開
 - 日本から425字を提案済み
- ・ IVDコレクション
 - 約2800組のIVS登録を今年度の汎用電子の活動で準備中
- ・ 今後の符号化
 - UCSへの国際提案
 - IVDへの登録

2011年9月リリース版のフォントに追加符号化

追加符号化の対応時期は来年度以降を想定

2012年4月リリース版のフォントにIVS追加を想定

23年度事業実施状況(5/5)

- 文字情報一覧表の整備



・ 辞書情報の追加

- 新大字典(講談社)見出し番号

- ・ UCSとの対応表は完成(約19000見出し)
- ・ 戸籍統一文字との対応表の作成中(約1850見出し程度。汎用電子およびJSC2と協力した、国際提案の典拠として利用(今年度中))

- 大字源(角川書店)見出し番号

- 日本語漢字辞典(新潮社)

23年度事業実施計画(1/7)

- 文字情報一覧表の整備



- ・ さらなる追加情報の検討

- 異体字情報

- ・ 異体字関係にある字を容易に識別できる情報
 - ・ 正確さよりも、検索の利便性を重視
 - ・ 各種辞典類情報の収集整理

- 読み

- ・ 各種辞典類情報の収集整理

作業方針については、今後設置するワーキンググループで精査する

23年度事業実施計画(2/7)



- 実証実験1 -

- WWWによる一般向けデモ
 - 人名漢字を豊富に使える環境の体験
 - 技術的問題点の洗い出し

- 主要ブラウザメーカーの参加
 - 仕様の検討
 - (必要に応じ)ブラウザの修正と修正版ダウンロードサイトの運営
- デモサイトの構築

23年度事業実施計画(3/7)

- 実証実験1 -

・デモサイトの構築

- ・ 対応ブラウザの配布
- ・ 一般市民の参加



- ・ 多様な文字の閲覧体験
- ・ 文字入力の体験
- ・ 模擬的な「名刺」、「証明書」等、豊富な人名漢字を生かしたイメージを作成して出力

23年度事業実施計画(4/7)

- 実証実験2 -

- 自治体の現場を対象とした実験
 - 実装例の創出
 - 問題点の洗い出し(技術面、運用面)

- ベンダーと自治体がペアで実施
 - 想定される実験:
 - システム連携
 - 情報の送受信
 - 文字入力現場での試用

23年度事業実施計画(5/7)



- 運用ガイドの検討 -

- 運用現場向けガイド
 - 実装のための技術ガイド
 - 実証実験(Webにおける運用)の結果を反映
 - 実務運用ガイド
 - 実証実験(自治体における運用)の結果を反映
 - 文字集合間のマッピング指針の策定
- 文字同定ハンドブック
 - 文字集合間のマッピング指針に基づく個別同定事例の収集整理
 - IPAmj明朝個別グリフのメンテナンス

23年度事業実施計画(6/7)

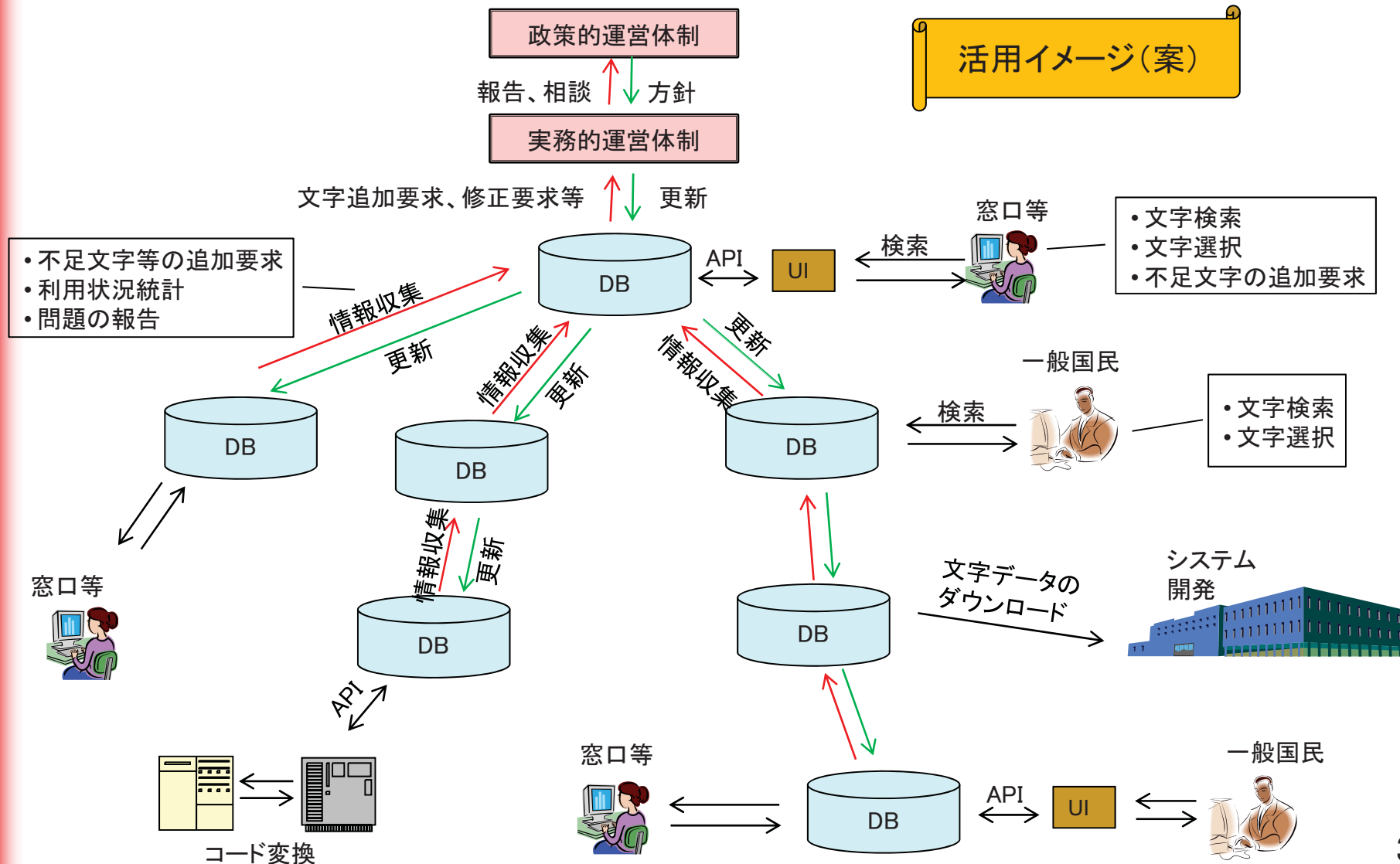


- データベースの検討 -

- 想定利用者
 - 自治体の現場(→検索、同定の補助等)
 - システム開発現場(→変換表、類似関係情報等の提供)
 - 一般国民(→文字の見える化)
- 目的
 - 情報の配布
 - 文字図形
 - 文字コード
 - 文字情報(読み、類似文字情報、辞書情報等)
 - 情報の収集
 - バグ情報
 - 文字追加要求
 - 統計情報
 - 不足文字の発生状況
 - 文字活用状況

23年度事業実施計画(7/7)

－ データベースの検討 －



ワーキンググループの設置

- ・ 課題に応じ、ワーキンググループを柔軟に設置して検討
 - － 文字情報検討WG
 - ・ 文字同定
 - ・ 符号化
 - ・ 文字情報一覧表内容の検討
 - － 運用検討WG
 - ・ 運用ガイドラインの検討
 - ・ ライセンス等の検討
 - ・ 文字情報データベースへの要求要件の検討
 - － 技術検討WG
 - Phase1 : webによる実証実験プロジェクト
 - Phase2 : 自治体現場における実証実験
 - Phase3 : 実験結果を踏まえた技術検討

検討課題

1. 文字情報基盤運用に係る基本方針の策定
例)
 - 文字数の調整に関する方針
 - 個人識別に関する方針
2. 文字情報基盤活用に関する方針策定
例)
 - 府省、自治体への普及に係るロードマップ
 - 社会的活用に関する方針