

# 文字情報基盤事業報告

2016年3月

独立行政法人 情報処理推進機構

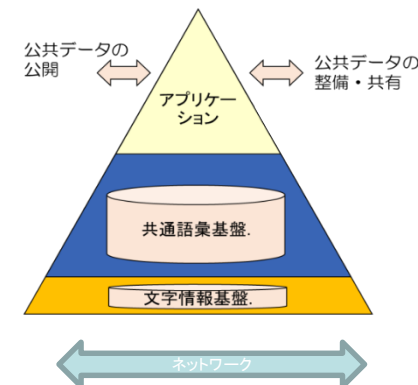
# 閣議決定「世界最先端IT国家創造宣言」

平成25年6月14日決定、平成26年6月24日改定、平成27年6月30日改定



## 世界最先端IT国家創造宣言 工程表より「文字情報基盤」に係る記述

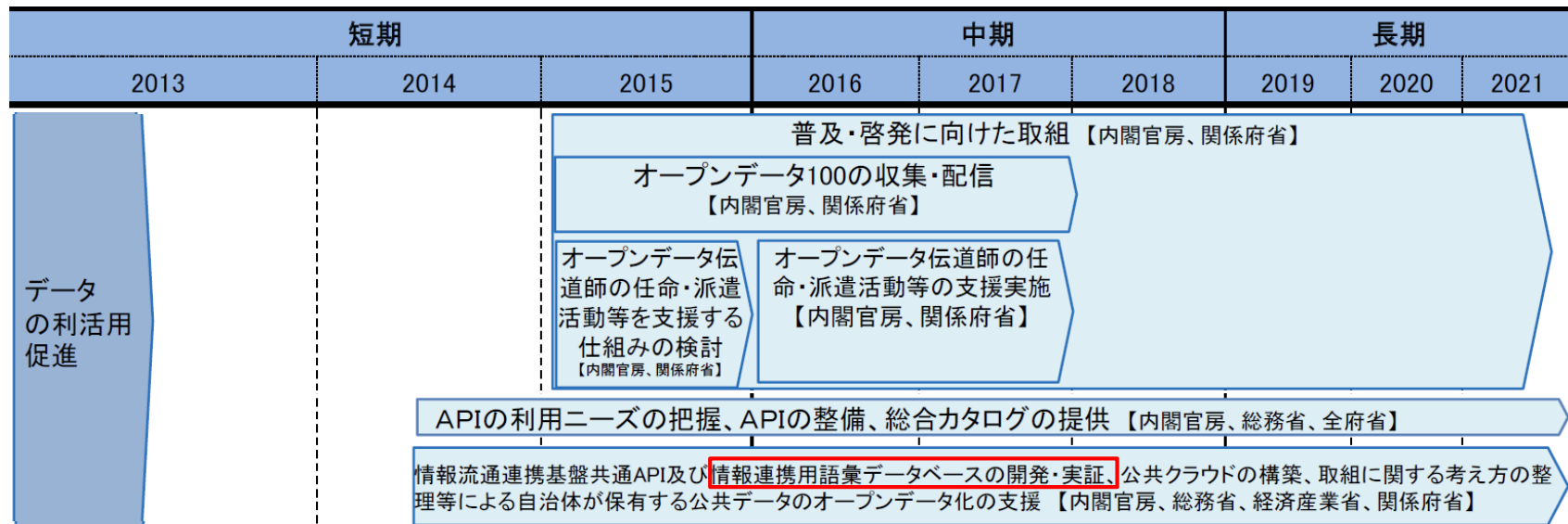
- 2016～2018 アクションプランを踏まえ、府省間連携を推進し、利便性の高い公共サービスを実現【内閣官房及び関係省庁】
  - 業務目的に応じて適切な範囲の文字（常用漢字範囲、JISX0208範囲、JISX0213範囲、**文字情報基盤範囲**）の導入を、同導入ガイドに沿って推進する。【内閣官房、**経済産業省**及び関係府省庁】



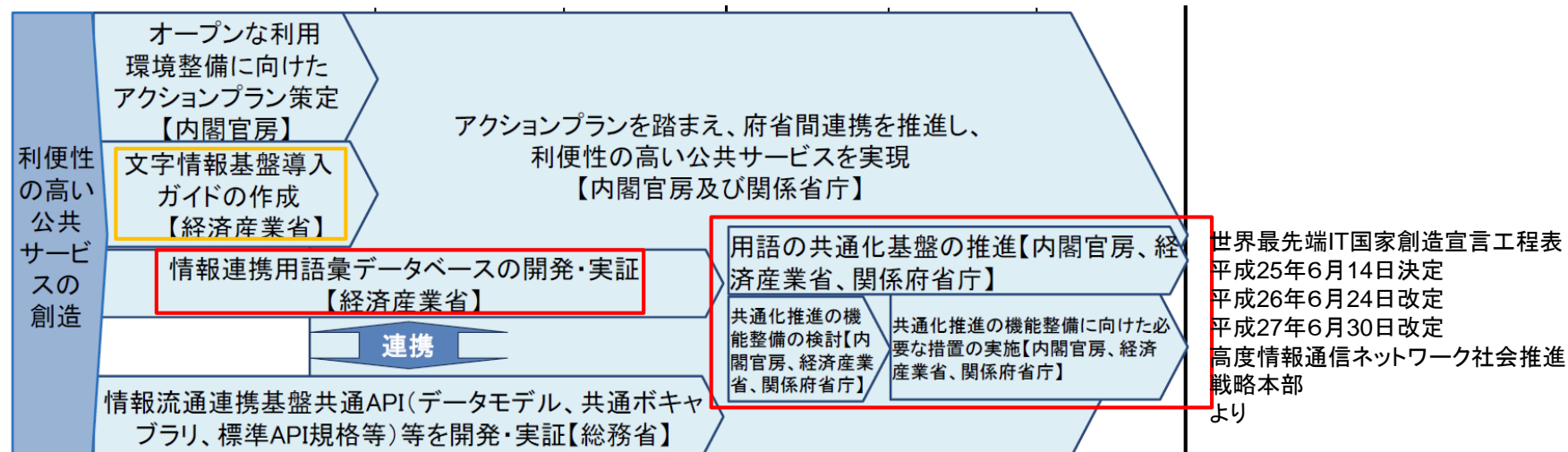
## 世界最先端IT国家創造宣言 工程表より「共通語彙基盤」に係る記述

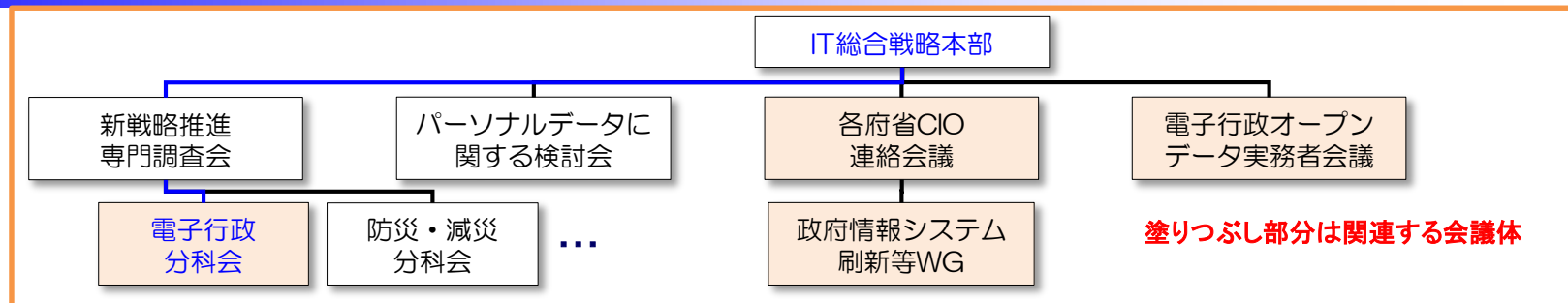
- 2016～2018 **用語の共通化基盤**の推進【内閣官房、**経済産業省**、関係府省庁】
  - 2016～電子行政分野におけるデータ等の共通化・相互運用性確保について、企画立案・維持管理を継続的に行うことができるようにするための機能（日本版NIST）の整備について検討し、必要な措置を行う。【内閣官房、**経済産業省**及び関係府省庁】
  - 2016共通化推進の**機能整備の検討**【内閣官房、**経済産業省**、関係府省庁】
  - 2017～2018共通化推進の機能整備に向けた**必要な措置の実施**【内閣官房、**経済産業省**、関係府省庁】
- 2014～2021情報流通連携基盤共通API及び**情報連携用語彙データベースの開発・実証**、公共クラウドの構築、取組に関する考え方の整理等による自治体が保有する公共データの**オープンデータ化の支援**【内閣官房、総務省、**経済産業省**、関係府省】

## 1. IT利活用の深化により未来に向けて成長する社会



## 4. ITを利活用した公共サービスがワンストップで受けられる社会





協力依頼



報告

## 情報共有基盤推進委員会

委員長 須藤修（東京大学情報学環長）  
 有識者 武田英明（共通語彙基盤WG座長）、林史典（文字情報基盤WG座長）  
 関係団体 JISA（ソフト業界）、JEITA（ハード業界）  
 オブザーバ 内閣官房（IT総合戦略室、社会保障改革担当室）、総務省（行政管理局）

- 各委員会、WGの設置主体は経済産業省、事務局はIPA
- 各WGの下には、具体的な検討を行なうSWGを設置。

### 共通語彙基盤WG

<実施内容>

- 情報連携用語彙基盤DBの構築
- APIカタログの整備
- 情報連携ツールの整備

【主査】武田英明（NII教授）

【関係省庁】（課長級）

IT室、社保室、内閣府防災、  
総務省、国交省、国会図書館

コア語彙  
検討運用  
SWG

コア語彙  
検討技術  
SWG

### 文字情報基盤WG

<実施内容>

- 導入ガイドの作成
- 文字情報DBの構築
- フォント整備
- 標準化の実施

【主査】林 史典（聖徳大学教授）

【関係省庁】（課長級）

IT室、総務省、法務、省

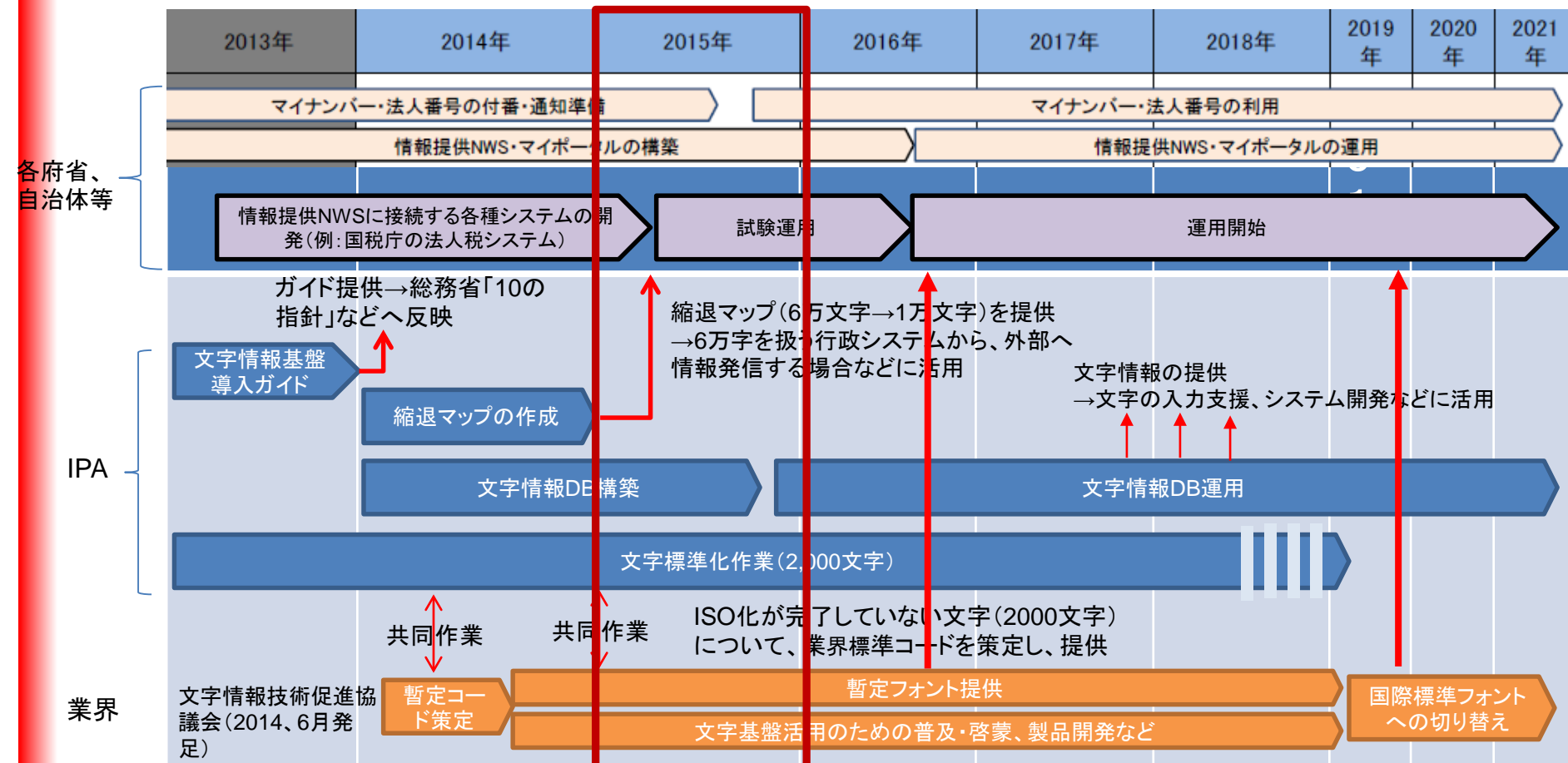
文字情報  
検討SWG

文字情報基  
盤データ  
ベース検討  
SWG

文字情報  
基盤運用  
検討SWG

経済産業省からIPAが事務局業務を受託

# 文字情報基盤導入ロードマップ（政府工程表との対比）



**IPA®**



- 縮退マップ

- メンバー

- 笹原 宏之      早稲田大学
    - 高田 智和      国立国語研究所
    - 守岡 知彦      京都大学
    - 安岡 孝一      京都大学

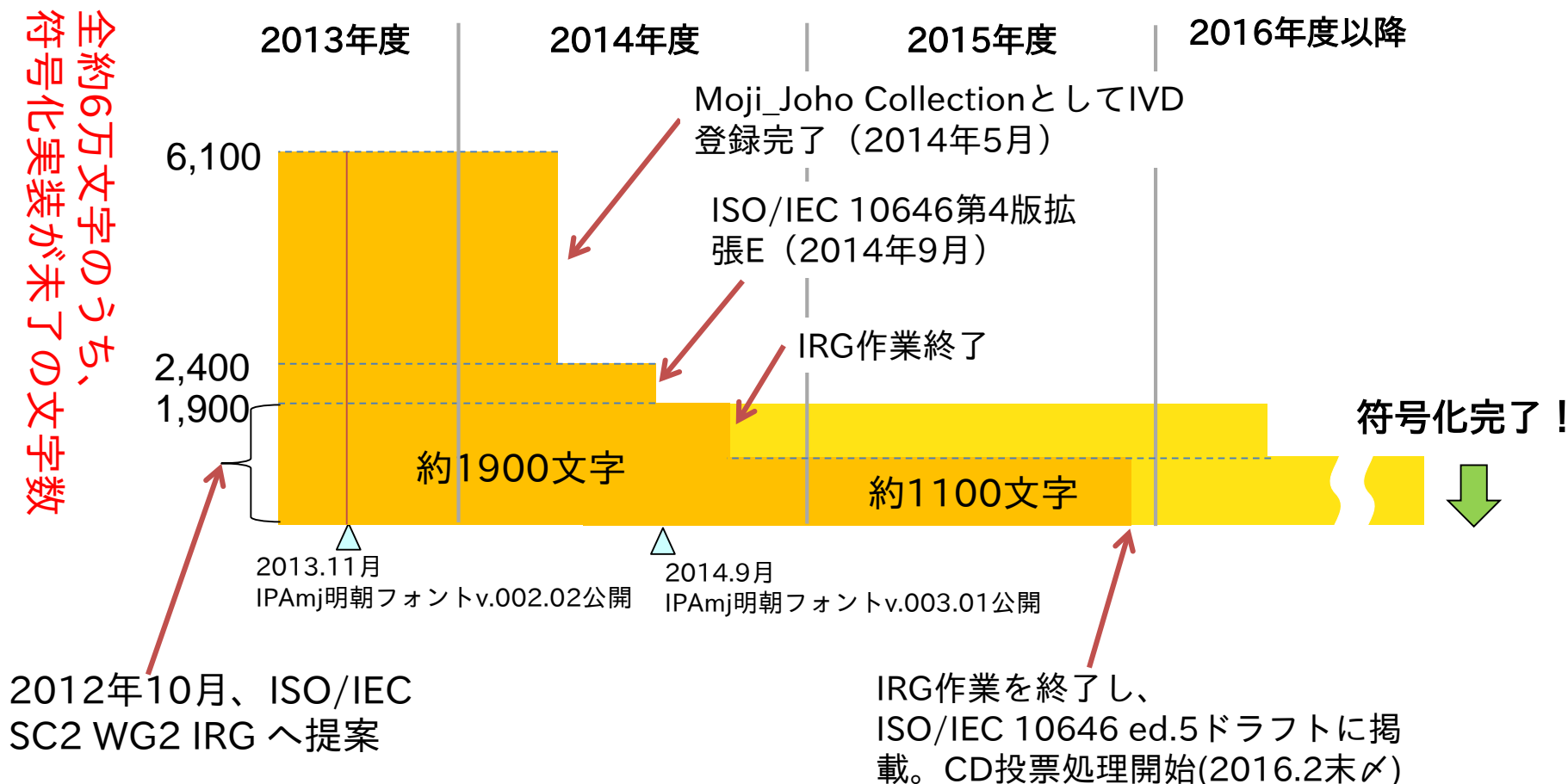
- 変体仮名

- メンバー

- 高田 智和      国立国語研究所
    - 矢田 勉      大阪大学

(50音順)

# 標準化の状況(漢字)





- 2015年3月：情報規格調査会SC2専門委員会で原案を決定
- 2015年6月30日～8月21日：IPAからパブコメ実施
- 2015年7月22日：ISO/IEC JTC1/SC2/WG2へRFC(Request for Comment)文書を提出
- 2015年7月27日～31日：UTC(Unicode Consortium Technical Committee)会議

符号化アーキテクチャ案をSVSモデルから独立符号化モデルへ変更

- 2015年10月19日～23日：ISO/IEC JTC1/SC2会議@島根県松江市

同一字母複数音価の文字を統合することに変更

- 2015年12月16日：リバイズ版を提出
- 2016年1月25日～28日：UTC会議

コンセンサス

# 提案中の変体仮名（抜粋）

Proposed Glyph Reference	RevisedCharacterName	Phonetic Value	family registration character #	academic use character#	Source ID
あ	HENTAIGANA LETTER A-1 •Derived From 5B89 安	あ (U+3042)	900010	010010010	JMJ-090001
あ	HENTAIGANA LETTER A-2 •Derived From 611B 愛	あ (U+3042)	900030	010030010	JMJ-090003
あ	HENTAIGANA LETTER A-3 •Derived From 963F 阿	あ (U+3042)	900020	010040010	JMJ-090004
あ	HENTAIGANA LETTER A-WO •Derived From 60E1 惡	あ;を (U+3042;U+3092)		010020010	JMJ-090002
い	HENTAIGANA LETTER I-1 •Derived From 4EE5 以	い (U+3044)	900040	020010010	JMJ-090005
い	HENTAIGANA LETTER I-2 •Derived From 4F0A 伊	い (U+3044)	900050	020020010	JMJ-090006
い	HENTAIGANA LETTER I-3 •Derived From 610F 意	い (U+3044)	900060	020030010	JMJ-090007
い	HENTAIGANA LETTER I-4 •Derived From 79FB 移	い (U+3044)		020050010	JMJ-090008
う	HENTAIGANA LETTER U-1 •Derived From 5B87 宇	う (U+3046)		030020010	JMJ-090009
う	HENTAIGANA LETTER U-2 •Derived From 5B87 宇	う (U+3046)	900070	030020020	JMJ-090010
う	HENTAIGANA LETTER U-3 •Derived From 6182 憂	う (U+3046)		030030010	JMJ-090011

- 「MJ縮退マップ」の正式版公開
  - 特徴：複数の縮退候補を根拠と共に示す
  - 縮退候補の追加
    - 法務省告示582号(外国人登録関連)別表第二を縮退根拠に追加
  - 表示する縮退根拠情報の追加
    - 「戸籍統一文字情報の親字・正字情報」を何段階辿ったかを示す「ホップ数」(本資料p.5-6参照)の表示を追加
    - 法務省告示582号別表第四を用いた縮退について、「別表第四の一」「別表第四の二」の区別の表示を追加
- 縮退先を一つに絞った縮退変換表の公開
  - 特徴：MJ縮退マップを元に、縮退先を一つ以下に絞り込んだもの。

- JSON形式のテキスト（図形は含まない）
  - 意見募集の結果、表形式の提供を望む意見が寄せられた、システム開発者等現場からはコンピュータ処理で使いやすいという意見が多かった
    - 既に多くのベンダーが検証版を利用
- 簡易検索システムでの縮退情報の提示
  - 目視に適した情報提供への要望に対応



ホームへ戻る 検索結果へ戻る

MJ030418



MJ文字情報一覧表Ver.004.02

コピペフィールド	5
戸籍統一文字番号	001970
住基ネット統一コード	
入管正字コード	
入管外字コード	
漢字属性	
JIS X 0213	
包摂区分	
包摂番号	
JIS水準	---
JIS X 0212	
対応するUCS	U+200D1
対応する互換漢字	

MJ縮退マップVer.0.1(検証版)

JIS包摂・UCS統合

法務省戸籍法関連通達・通知

隠	隠	隠
2-01-15	1-17-03	1-80-12
戸籍統一文字情報	戸籍統一文字情報	戸籍統一文字情報
観字・正字	観字・正字	観字・正字

法務省告示582号別表四

読み・字形による関連字

読み・字形による関連字

参考情報

## 文字関連マップ

[illegible]

# 文字検索結果への縮退情報の表示

## IPA MJ文字情報検索システム(簡易版)

詳細

[ホームへ戻る](#)

[検索結果へ戻る](#)

**MJ053396**

齋

### MJ文字情報一覧表Ver.004.03

コピーフィールド	齋
戸籍統一文字番号	490050 (親字・正字を探す)
住基ネット統一コード	
入管正字コード	
入管外字コード	
漢字施策	
JIS X 0213	
包摂区分	
包摂連番	
JIS水準	---
JIS X 0212	
対応するUCS	U+29413 (関連字を探す)
対応する互換漢字	
UCS対応カテゴリー	A
実装したUCS	U+29413
実装したMoji_Joho IVS	
対応するHanyo-	

### MJ縮退マップVer.1.0.0

JIS包摂・UCS統合

法務省戸籍法関連通達・通知

法務省告示582号別表第四 その一

法務省告示582号別表第四 その二

辞書類等による関連字

齋 齋 齋

1-80-77 1-80-78 1-83-78

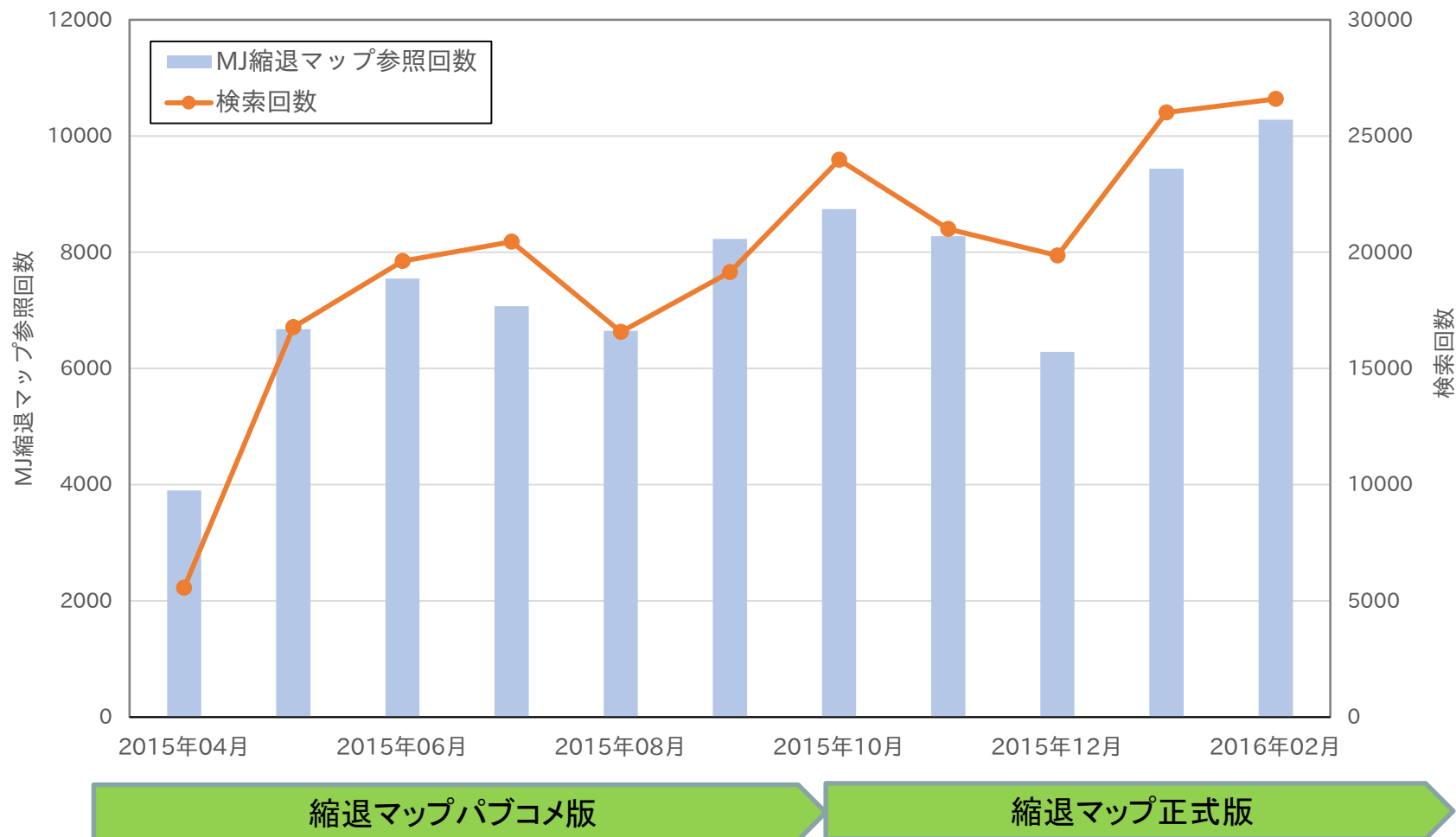
読み・字形による類推

参考情報

## 活用状況

# 簡易検索システム利用状況

MJ文字情報検索簡易検索システム 利用状況







The screenshot shows the homepage of the National Tax Authority's Corporate Number Disclosure Site. The page is in Japanese and features a search bar at the top. Below the search bar, there are two main search sections:

- 法人番号 (Corporate Number) Search:** A search bar with the text "法人番号" and a search button. Below the search bar, there is a text input field with the example number "1234567890123" and a search button. To the right of the search bar, there is a note: "複数の法人番号でまとめて検索したい場合は、「法人番号でまとめて検索する」をクリックしてください。" (If you want to search for multiple corporate numbers at once, click on "Search by corporate number").
- 法人の商号及び所在地 (Corporate Name and Location) Search:** A search bar with the text "法人の商号及び所在地" and a search button. Below the search bar, there is a text input field with the example name "株式会社" and a search button. To the right of the search bar, there is a note: "商号又は名称については、通称や略称ではなく、正式名称で入力してください。" (For the corporate name or name, please enter the formal name, not a nickname or abbreviation).

- 法人の商号及び所在地の文字を登記統一文字(7万)からJIS X 0213(1万)に変換するのに利用
- 縮退マップにより変換できないものは”\_”に置き換え
- 約430万の法人のうち、変換できなかった法人は約1700(仮名、記号等や中国等の法人を含む)



The screenshot shows a web browser window with the URL [www.houjin-bangou.nta.go.jp/download/](http://www.houjin-bangou.nta.go.jp/download/). The page content is in Japanese and discusses the sorting of data by corporate number and the use of image IDs for digital signatures. It includes a section for 'Text Handling' (文字の取扱い) and a section for 'Digital Signature' (デジタル署名). A specific URL for external character confirmation is provided: <http://www.houjin-bangou.nta.go.jp/image?imageid=>. An example is given for an image ID of '00000001'. A link at the bottom points to 'About the PGP public key of the Japanese Tax Authority's Corporate Number Disclosure Site'.

番号(昇順)に加え、履歴の古い順でソートしています。  
全件データは、法人番号(昇順)でソートし、一連番号を設定しています。

## 文字の取扱い

法人の商号及び所在地には、JIS第一・第二水準以外の文字が使用されている場合があります。このサイトでは、利用者のシステムで二次活用が容易にできるよう、該当する文字をJIS第一～第四水準の文字及びJIS第一・第二水準の文字に縮退し、データの提供を行っています。

なお、該当文字の登記記録等における表記を確認したい場合、該当する法人を検索し、検索結果に表示される「外字」ボタンをクリックすることにより、該当文字の登記記録等における表記が画面に表示されます。

また、ダウンロードファイルで提供するイメージIDを利用し、該当文字の登記記録等の表記を確認することもできます。ダウンロードファイルから取得したイメージIDを次のURLの後に続けて入力し、アクセスしてください。

**外字確認用URL** <http://www.houjin-bangou.nta.go.jp/image?imageid=>

入力例：イメージIDが「00000001」の場合  
<http://www.houjin-bangou.nta.go.jp/image?imageid=00000001>

## デジタル署名

ダウンロードファイルには、OpenPGPを用いてデジタル署名を付与しております。署名の検証方法等については、「国税庁法人番号公表サイトのPGP公開鍵について」を参照してください。

➡ [国税庁法人番号公表サイトのPGP公開鍵について](#)

⬆ [ページの先頭へ戻る](#)

19

# 法人番号公表サイト表示例1(2/2)

社会保険・税番号制度  
国税庁 法人番号公表サイト

法人番号で わかる。つながる。ひろがる。

閲覧支援ツール(国税庁ホームページ) > サイトマップ

文字サイズ + 大きく 元に戻す - 小さく

ホーム 法人番号とは 基本3情報ダウンロード Web-API 各種手続 よくある質問

ホーム > [redacted] の情報

■ 有限会社[redacted] 情報

■ 最新情報

×

 閉じる

有限会社[redacted]

本店又は主たる事務所の所在地  
[redacted]

最終更新年月日  
平成27年10月30日

## IPA MJ文字情報検索システム(簡易版)

詳細

[ホームへ戻る](#)

[検索結果へ戻る](#)

**MJ025760**

辻

### MJ文字情報一覧表Ver.004.03

コピペフィールド	辻
戸籍統一文字番号	437660 (親字・正字を探す)
住基ネット統一コード	J+8FBB
入管正字コード	
入管外字コード	
漢字施策	
JIS X 0213	1-36-52
包摂区分	2
包摂連番	
JIS水準	第1水準
JIS X 0212	
対応するUCS	U+8FBB (関連字を探す)
対応する互換漢字	
UCS対応力カテゴリー	A
宝樹のUCS	

### MJ縮退マップVer.1.0.0

JIS包摂・UCS統合

辻

1-36-52

法務省戸籍法関連通達・通知

辻

1-36-52

戸籍統一文字情報 親字・正字

法務省告示582号別表第四 その一

法務省告示582号別表第四 その二

# 法人番号公表サイト表示例2(1/2)

社会保障・税番号制度  
国税庁 法人番号公表サイト

法人番号で わかる。つながる。ひろがる。

閲覧支援ツール(国税庁ホームページ) > サイトマップ

文字サイズ [+ 大きく](#) [元に戻す](#) [- 小さく](#)

[ホーム](#) | [法人番号とは](#) | [基本3情報ダウンロード](#) | [Web-API](#) | [各種手続](#) | [よくある質問](#)

[ホーム](#) > 上海 [redacted] の情報

## 上海 [redacted] 有限公司の情報

### 最新情報

法人番号  
[redacted]

商号又は名称  
上海 [redacted] 有限公司

本店又は主たる事務所の所在地  
中華人民共和国上海市嘉定区迎\_\_中路 [redacted] [> 外字](#)

国内における主たる事務所等の所在地  
[redacted]

最終更新年月日  
平成27年11月17日

# 法人番号公表サイト表示例2(2/2)

社会保障・税番号制度  
国税庁 法人番号公表サイト

法人番号でわかる。つながる。ひろがる。

閲覧支援ツール(国税庁ホームページ) > サイトマップ

文字サイズ + 大きく 元に戻す - 小さく

ホーム | 法人番号とは | 基本3情報ダウンロード | Web-API | 各種手続 | よくある質問

ホーム > 上海 [redacted] の情報

上海 [redacted] 有限公司の情報

最新情報

✕ 閉じる

中華人民共和国上海市  
嘉定区迎园中路 [redacted]  
[redacted]

国内における主たる事務所等の所在地

[redacted] 24

最終更新年月日  
平成27年11月17日

## IPA MJ文字情報検索システム(簡易版)

[詳細](#)[ホームへ戻る](#)**MJ008835**

園

### MJ文字情報一覧表Ver.004.03

コピペフィールド	園
戸籍統一文字番号	053850 (親字・正字を探す)
住基ネット統一コード	J+56ED
入管正字コード	56ED
入管外字コード	
漢字施策	
JIS X 0213	
包摂区分	
包摂連番	
JIS水準	---
JIS X 0212	22-94
対応するUCS	U+56ED (関連字を探す)
対応する互換漢字	
UCS対応カテゴリー	A
実装したUCS	U+56ED
実装したMoji_Joho IVS	
対応するHanyo-Denshi IVS	
実装したSVS	
MJ文字図形バージョン	1.0
登記統一文字番号	00053850
部首・内画数	口部4画
読み	ガン・タン・ダン・エン
総画数	7
諸橋大漢和	4714
日本語漢字辞典	
新大字典	2470
大字源	

### MJ縮退マップVer.1.0.0

JIS包摂・UCS統合
法務省戸籍法関連通達・通知
法務省告示582号別表第四 その一
法務省告示582号別表第四 その二
辞書類等による関連字
読み・字形による類推
参考情報



# 文字情報DBの構築

2016年5月一般公開予定

日付	事象
2014年7月	「新文字情報基盤データベース」のシステム開発スタート
2015年2月	データの整備完了 → 縮退マップ等へ活用
2015年7月	システムの仮運用スタート
2015年10月 ～2016年2月	外部専門家によるシステム検証
2016年5月	システム一般公開（予定）

	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 ・ 戸籍法上の人名漢字 ・ 総画数 ・ 康熙部首・部首内画数 ・ 音読み ・ 訓読み ・ JIS X 0213水準
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

# 文字検索機能（サンプル画面）

## トップページ

文字情報基盤データベース

IPA Better Life with IT 情報処理推進機構

ホーム 基本検索 SPARQLクエリ検索 変体仮名一覧 管理画面

文字を入力してクイック検索

漢字を検索する

漢 かん U+6F22

検索

さらに詳細な検索はこちらへ  
基本検索の詳しい使い方  
変体仮名一覧へ

条件を指定した検索はここから

独立行政法人 情報処理推進機構 - IPA

## 検索結果、文字詳細表示

文字情報基盤データベース

IPA Better Life with IT 情報処理推進機構

ホーム 基本検索 SPARQLクエリ検索 変体仮名一覧 管理画面

MJ015841

漢

コピーフィールド

漢

MJ文字図形名	MJ015841
総画数	13
音読み	タン、カン
訓読み	から
康熙字典	水 通番号:005 内画数:10
バージョン	1.0
更新日時	2011-05-17

検字番号
日本国漢字辞典 SN-06258
大漢和辞典 TH-1806801
新大字典 KS-08094
大字典 KD-05145
大漢語林 TD-06170

平成明瞭グリフ名	JA2033
戸籍統一	203690
戸籍統一URI	URI
住基統一	J+6F22
入管正字	6F22
入管外字	
漢字施策	常用漢字
登記統一	00203690
JISX0213	1-20-33
JIS対応適応包摂標準	
JIS水準	1
JISX0213追加履歴	2000
JISX0213変更履歴	

**IPA®**



# 複合条件での検索（サンプル画面）

基本検索 | 文字情報基盤... x +

172.16.141.25/subtest/search 検索

文字情報基盤データベース IPA Better Life with IT 情報処理推進機構

ホーム 基本検索 SPARQLクエリ検索 変体仮名一覧 管理画面

ホーム / 基本検索 ヘルプ お問い合わせ

### 検索条件

☒ 次の全ての条件に一致する ☐ 次のいずれかの条件に一致する  
検索条件を追加してください。

JIS水準 が次と一致する ☐ 第一 ☐ 第二 ☐ 第三 ☐ 第四 ✕

読み が次と一致する ✕

部首 が次と一致する 部首選択 ✕  
部首をクリックすると削除できます。

単漢字 が次と一致する 漢 ✕

単漢字  
MJ文字図形名  
平成明朝グリフ名  
旧籍統一文字番号  
**住基ネット統一コード**  
大管正字コード  
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  
大漢語林検字番号

条件 + コード検索条件 + 文字部品検索条件 + 条件グループ

上記条件に該当するMJ文字図形名の数：0件 件数更新

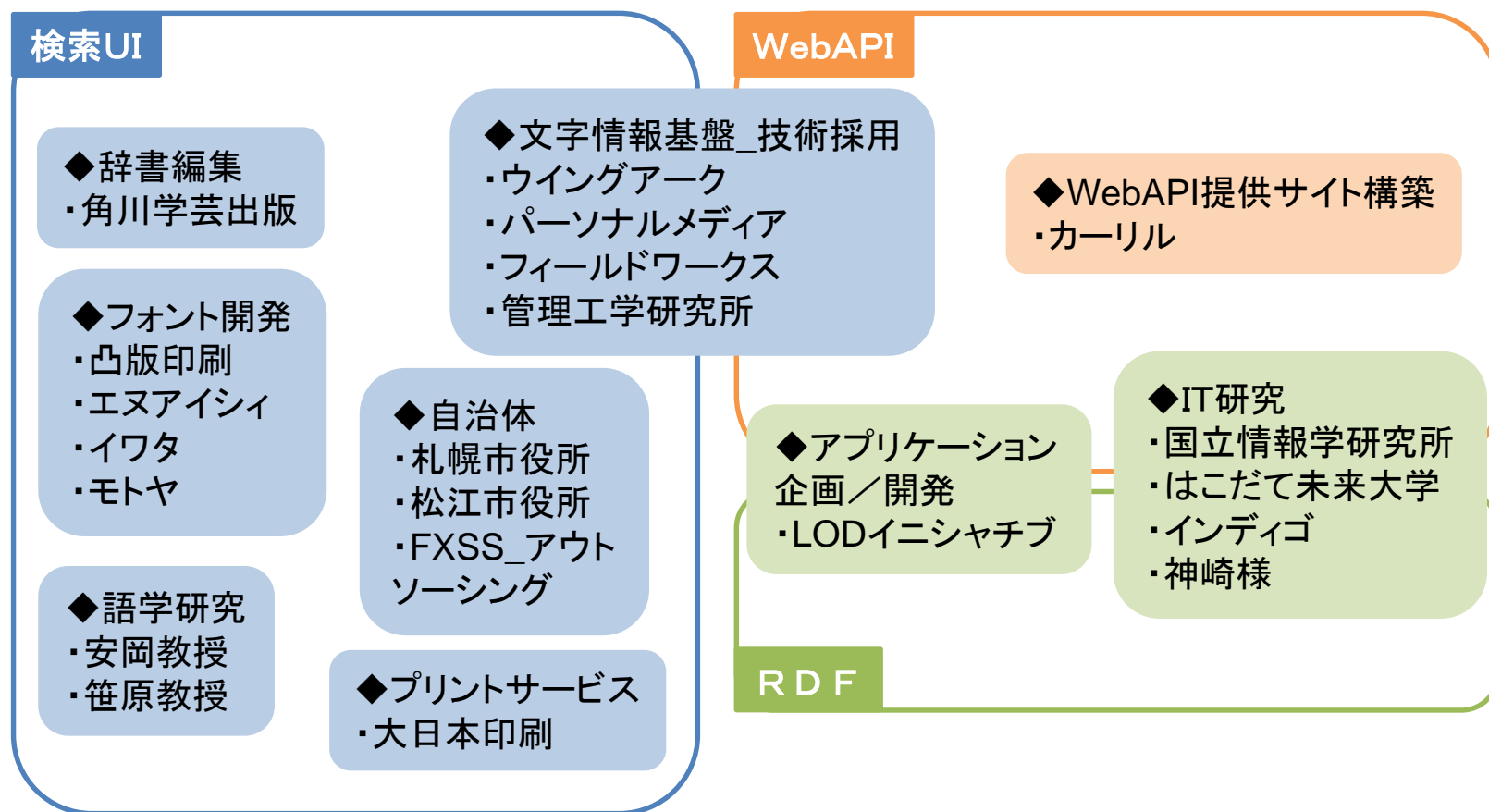
検索

いろいろな条件  
をして検索

独立行政法人 情報処理推進機構 - IPA

# システム検証

(2015年10月～2016年2月)



評価結果については、書面でのアンケートと対面でのヒアリングを実施

- 文字情報へのURIによる直接アクセス機能
  - 自治体間などでの文字情報の共有への有効性を確認
  - 現状はIPA外部の情報資源については法務省サイト以外は仮のURIとなっており、アクセス不能。
    - 多くの外部情報資源との連携を目指す必要
    - 公的情報のオープンデータとしての提供について、政府全体の取り組みも必要
    - 有償情報との連携も検討が必要
- 応答速度等基本的性能
  - 特に問題点は指摘されなかった
- 検索機能
  - 機能への問題は指摘されなかったが、多様な機能の提示法など、UIについての改善指摘があった。
  - 文字部品検索の有効性を指摘する声が多かった
- SPARQLエンドポイントの提供
  - 専門家による試用においては、問題はなかったが、活用できる人がどれだけいるかについての疑問が指摘された。→テンプレート等の提供が必要
- ユーザインタフェース(UI)
  - 簡単な検索については、現在提供中の「簡易検索」の方が使いやすいとの指摘
  - 縮退マップとの連携への要望