政策動向と外字の関係と課題

一般社団法人 文字情報技術促進協議会 理事
袴田　博之

本資料は日本電気株式会社の公式発表・見解を表すものではありません

© NEC Corporation 2022

行政デジタル化に関する法令・政策

法令・政策	目的・位置づけ	制定・改正時期
官民データ活用推進基本法	データ流通環境の整備や行政手続のオンライン利用の原則化など、官民データの活用に資する各種施策の推進が政府の取組として義務付け	2016年12月
世界最先端デジタル国家創造宣言・官民データ活用推進基本計画	全ての国民がデジタル技術とデータ利活用の恩恵を享受するとともに、安全で安心な暮らしや豊かさを実感できるデジタル社会の実現に向けた、政府全体のデジタル政策を取りまとめたもの	2017年5月 2019年6月改定 2020年12月改定
デジタル社会の実現に向けた改革の基本方針	デジタル社会の将来像、IT 基本法の見直しの考え方、デジタル庁設置の考え方等について、政府の方針を示す	2020年12月
デジタル・ガバメント推進方針	国民・事業者の利便性向上に重点を置き、行政の在り方そのものをデジタル前提で見直すデジタル・ガバメントの実現を目指す	2017年5月
デジタル・ガバメント実行計画	データ環境の整備や行政手続のオンライン化などを実現するために制定	2018年1月初版 2018年7月改定 2019年12月改定 2020年12月改定
デジタル社会の実現に向けた重点計画	目指すべきデジタル社会の実現に向けて、政府が迅速かつ重点的に実施すべき施策を明記し、各府省庁が構造改革や個別の施策に取り組み、それを世界に発信・提言する際の羅針盤となるもの	2021 年 12 月

© NEC Corporation 2022

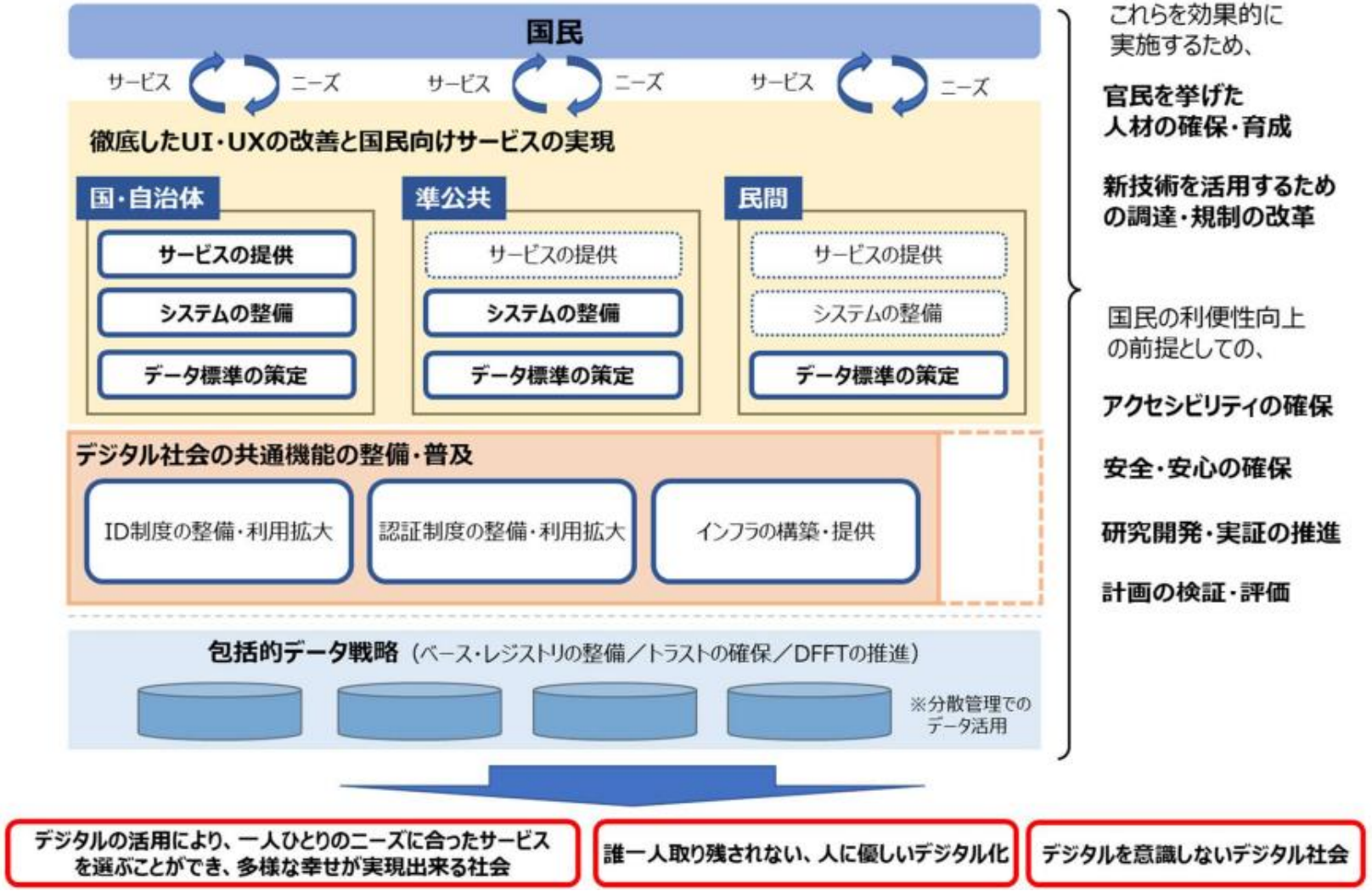
法律と計画の関係

2021年12月24日 閣議決定

デジタル社会形成基本法	デジタル社会の実現に向けた重点計画
情報通信技術を活用した行政の推進等に関する法律	情報システム整備計画
官民データ活用推進基本法	官民データ活用推進基本計画

© NEC Corporation 2022
\Orchestrating a brighter world NEC

デジタル庁のデータ戦略



これらを効果的に
実施するため、

官民を挙げた
人材の確保・育成

新技術を活用するため
の調達・規制の改革

国民の利便性向上
の前提としての、

アクセシビリティの確保

安全・安心の確保

研究開発・実証の推進

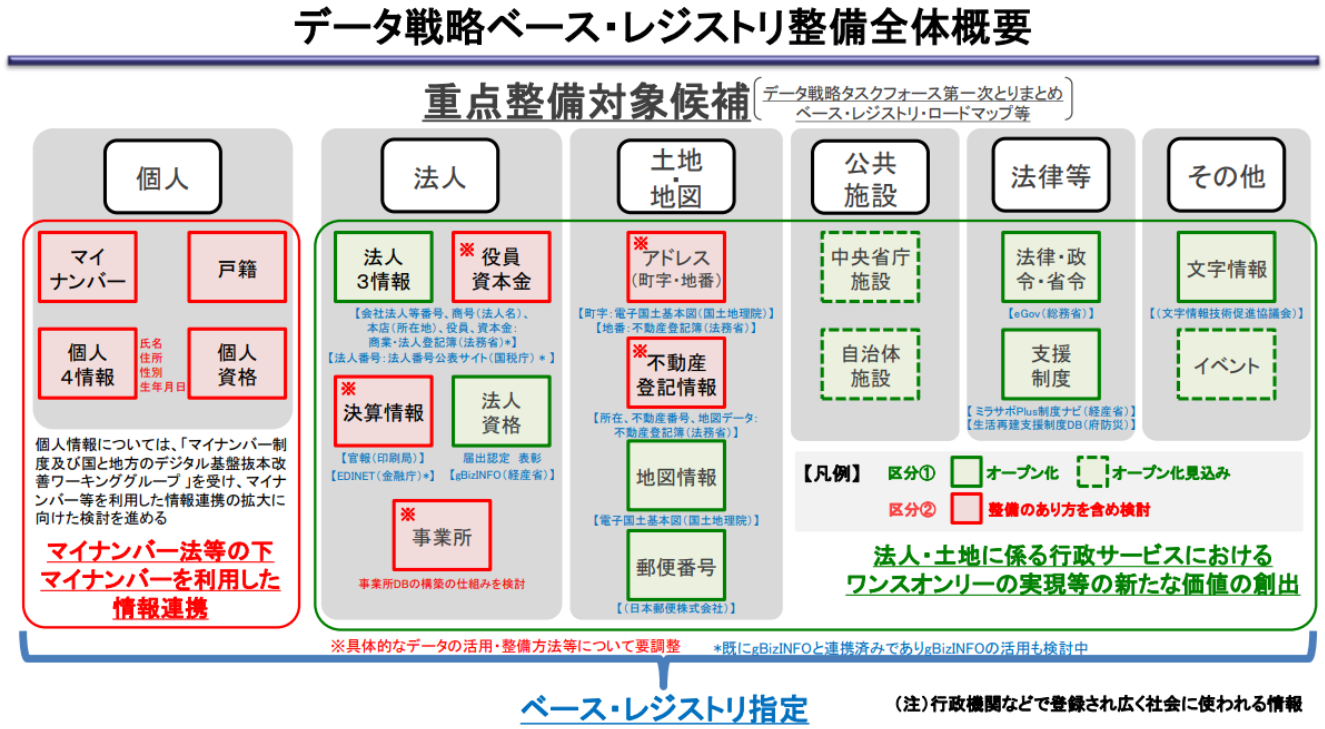
計画の検証・評価

出典：デジタル庁「デジタル社会の実現に向けた重点計画（ 概 要 ）」
https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/5ecac8cc-50f1-4168-b989-2bcaabffe870/40b7eb9b/20220607_policies_priority_outline_03.pdf

© NEC Corporation 2022

ベース・レジストリ

ベース・レジストリでは語彙の正規化は前提→文字の正規化も大前提
文字情報基盤はベース・レジストリの整備対象だが、民間に委ねられている



■ベース・レジストリ(注)の指定に際しては、以下の区分により指定を行った

区分①:即効性の観点から、早期にベース・レジストリとしての利活用を実現するものとして指定するデータ

区分②:今後ベース・レジストリとして整備のあり方を含め検討するものとして指定するデータ

⇒区分②については、マスターデータベースが不在、共通キーの不在、台帳間連携にあたって技術的、制度的課題が存在していることから、まずは目指すべき姿から検討

© NEC Corporation 2022
\Orchestrating a brighter world NEC

地方自治体標準化対象業務

地方自治体の事務

Ⅰ．地域情報プラットフォーム／中間標準レイアウトで示されている事務

1．住民基本台帳	10．国民健康保険	16．生活保護
4．選挙人名簿管理	11．国民年金	19．健康管理
5．固定資産税	12．障害者福祉	20．就学
6．個人住民税	13．後期高齢者医療	23．児童扶養手当
7．法人住民税	14．介護保険	子ども子育て支援
8．軽自動車税	15．児童手当	**合計：17業務**

標準仕様書作成済み

21．戸籍

※戸籍の附票は、「21.戸籍」の業務ユニットに含まれている

法律に基づかない事務

2．印鑑登録
17．乳幼児医療
18．ひとり親医療
30．住登外管理

自治体の内部管理事務

9．収滞納管理
50．財務会計
51．庶務事務
52．人事給与
53．文書管理

Ⅱ．Ⅰ以外の事務

Ⅰ以外の事務は、システム間の連携が大きくない。

インフラの点検・維持管理

等

番号は地域情報プラットフォームの業務ユニット番号に準拠

出典:デジタル庁「地方公共団体の基幹業務システムの統一・標準化のために検討すべき点について」2021年12月(2022年4月加筆)
https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/c58162cb-92e5-4a43-9ad5-095b7c45100c/034f5707/20220428_policies_local_governments_outline_01.pdf

© NEC Corporation 2022
\Orchestrating a brighter world NEC

地方自治体標準化の文字要件(1)

データ要件・連携要件標準仕様書で一括定義

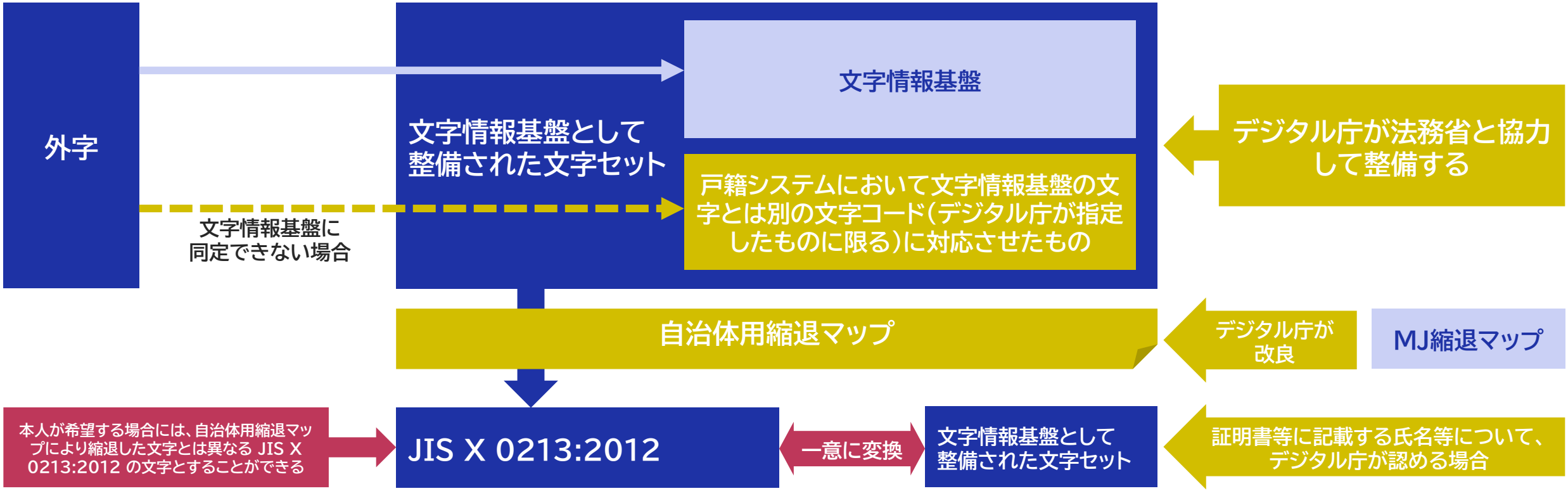
標準化対象 業務システム	項目	文字セット	文字コード	符号化方式 (システム間連携)	符号化方式 (システム内)
戸籍システム及び 戸籍附票システム	氏名、本籍、筆頭者及び住所/方書(左記の情報を基に記録される他の項目も含む。)	**「文字情報基盤として整備された文字セット」** ※従来の文字セットを、文字情報基盤として整備された文字セットと対応させて保持することは、経過措置として、当分の間、可能とする	JIS X 0221:2020	UTF-8	UTF-8/UTF-16
住民記録システム及び 印鑑登録システム	氏名等(氏名／旧氏／通称、世帯主の氏名、住所／方書、本籍及び筆頭者(左記の情報を基に記録される他の項目も含む。)				
上記以外の 標準準拠システム	氏名等	JIS X 0213:2012 ※住民に対して発行する**証明書等に記載する氏名等**について、文字情報基盤として整備された文字を表記する必要があるとデジタル庁が認める場合においては、**氏名等に係る文字情報基盤として整備された文字とJIS X 0213:2012 へ縮退された文字を一意に変換して表示**すること ※氏名等について、文字情報基盤として整備された文字から JIS X 0213:2012 の文字への縮退は、デジタル庁が MJ 縮退マップを改良して作成した自治体用縮退マップを用いて行う。ただし、縮退した文字について、**本人が希望する場合には、自治体用縮退マップにより縮退した文字とは異なる JIS X 0213:2012 の文字とすることができる。**			
	氏名等以外	JIS X 0213:2012			

デジタル庁「地方公共団体情報システム データ要件・連携要件 標準仕様書【第1.0版】」に基づき筆者が整理。
https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/c58162cb-92e5-4a43-9ad5-095b7c45100c/dcaad375/20220831_local_governments_02.pdf

© NEC Corporation 2022

地方自治体標準化の文字要件(2)

外字を排除
JIS X 0213への縮退設計がカギ

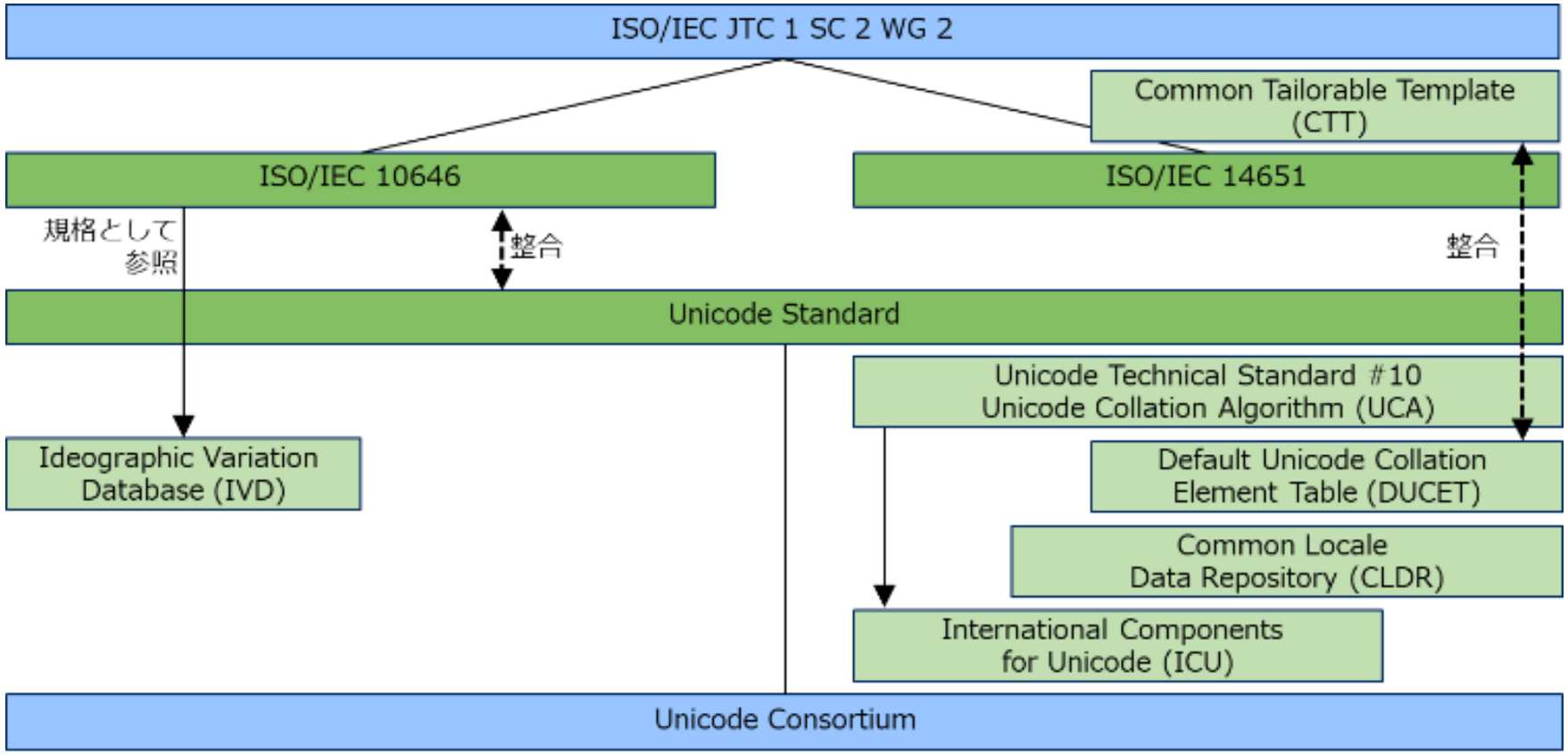


デジタル庁「地方公共団体情報システム データ要件・連携要件 標準仕様書【第1.0版】」に基づき筆者が整理。
https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/c58162cb-92e5-4a43-9ad5-095b7c45100c/dcaad375/20220831_local_governments_02.pdf

© NEC Corporation 2022
\Orchestrating a brighter world NEC

文字に関する規格・標準

- 文字の形に文字コードを割り当てる：国際規格と業界標準が協調
- 文字/文字列の処理：業界標準が先導、リファレンス実装（ICU）採用が主流に



© NEC Corporation 2022
\Orchestrating a brighter world NEC

ISO/IEC-10646とUnicodeの関係

- OSSはサポートする文字規格をUnicodeバージョンで示すことが多い
- ISO/IEC 10646:2017以降＝Unicode 10以上となる

	ISO/IEC 10646 (UCS)	Unicode	補足
1991		1	
1992		1.0.1	
1993	10646-1:1993 (1st edition)	1.1	
1996	10646-1:1993/Amd 1,2,3,4	2	サロゲートペア
1997	10646-1:1993/Amd 6,7,8,9		
1998	10646-1:1993/ Amd 5,11,12,10,13,16,19,20	2.1	
1999	10646-1:1993/Amd 23,21,17,18,14,22,24,25,26, 27,29,30		
2000	10646-1:1993/Amd 15,28,31		
	10646-1:2000 (2nd edition)	3	
2001	10646-2:2001	3.1	
2002	10646-1:2000/Amd 1	3.2	JIS X 0213
2003	10646:2003 (1st edition)	4	
2005	10646:2003/Amd 1	4.1	
2006	10646:2003/Amd 2	5	

	ISO/IEC 10646 (UCS)	Unicode	補足
2008	10646:2003/Amd 3,4,5	5.1	IVS
2009	10646:2003/Amd 6	5.2	
2010	10646:2003/Amd 7	6	
2011	10646:2011 (2nd edition)		
2012	10646:2012 (3rd edition)	6.1, 6.2	
2013	10646:2012/Amd 1	6.3	
2014	10646:2014 (4th edition)	7	
2015	10646:2014/Amd 1	8	
2016	10646:2014/Amd 2	9	
2017	10646:2017 (5th edition)	10	文字情報基盤 登録完了
2018	10646:2017/Amd 1:2019	11	
2019	10646:2017/Amd 2:2019	12	
2020	10646:2020 (6th edition)	13	
2021		14	
2022		15 (予定)	

© NEC Corporation 2022

課題の例(1)

◆ 文字検索でVS付き文字と基底文字とは区別できない

- UCAで”default ignorable”と規定。ICUでも基底文字が同じであれば区別しない実装
  - OSS含めOS/言語処理系はICUに従っている
    - 唯一の例外:Microsoft SQL Server:VS Sensitiveの指定で区別可能
  - 当協議会で、VSを区別するブランチ(テーラリング)を検討中だが、実現時期は未定
- 区別できない前提で設計する必要がある

© NEC Corporation 2022
\Orchestrating a brighter world NEC

課題の例(2)

◆ 文字合成への対応

- JIS X 0213では「か＋゜」を定義しているが、合成形で符号化していない
  - この表現が行政処理で必要かどうかは別の問題
- 文字情報基盤の変体仮名は、濁点付き・半濁点付き合成形を符号化していない
  - 合成処理が前提となっている
- 表示・印刷の確認が必要
  - 外字で「濁点付き・半濁点付き変体仮名」が存在する場合に、変体仮名＋合成用濁点・半濁点に同定すべきと主張しているわけではない
  - システム設計・検証では、文字合成も考慮しておくのが無難ということ

© NEC Corporation 2022
\Orchestrating a brighter world NEC

課題の例(3)

◆ JIS X 0213と文字情報基盤の規準の違い

- JIS X 0213は、「使える字形が限られる」のではなく、「字形の細かい差異を捨象(包摂)している」という前提
  - 従って、MJ縮退マップによるJIS X 0213への縮退も「別の字形で代替」しているのではない。解像度を落としているようなもの
- JIS X 0213で「氏名等」を表現する場合は、この前提を理解する必要がある
  - データ要件・連携要件標準仕様書の規定はあくまで例外措置と理解すべき
  - この前提が業務の運用設計に即しているかどうかは別の問題

© NEC Corporation 2022
\Orchestrating a brighter world NEC

参考：文字情報技術促進協議会 体制

◆ 文字情報基盤の維持体制

- 協議会活動とは独立した組織として「文字情報基盤委員会」を組織化
- 関係各省庁からもオブザーバーに参加いただき、公共性を担保する
- 国内の要望（各省庁に対する窓口）と国際規格（情報規格調査会との連携）との整合を図る
  - 文字情報基盤は行政で利用している文字集合を整理統合したもの
    →今後も各省庁からの要望に応じて検討するというのが基本的な姿勢

© NEC Corporation 2022

参考：文字情報技術促進協議会 計画

- ◆ IPAフォントライセンスv1.0に基づく、文字情報基盤成果物の公開は継続
  - ■ バグフィックス版を公開予定だが時期は未定
- ◆ 以下を会員限定で配布予定（時期・利用条件は未定）
  - ■ IPAmj明朝 いわゆる半角英数の固定幅化版
  - ■ IPAmjPUP v006.01対応版
- ◆ 以下を会員限定でライセンス予定
  - ■ IPAmj明朝のWOFF化ファイルのサービス利用・提供
    - IPAフォントライセンスv1.0では許諾対象外というのが協議会の見解
    - （1） 商用サービス提供、（2） 商用/非商用によらず、サービス構築・運用の受託、が対象
    - 定款を改正して利用規模に応じて会費（口数）を徴収する方向で検討中

© NEC Corporation 2022

\Orchestrating a brighter world

NECは、安全・安心・公平・効率という社会価値を創造し、
誰もが人間性を十分に発揮できる持続可能な社会の実現を目指します。

\Orchestrating a brighter world

NEC