

2-4 円滑なデータ移行に向けた方策

1. 本講義の学習目標

- 他事業者間でのシステム移行時のデータ移行における主な課題(コスト要因)として、以下について認識できる。
 - ①データレイアウトの多様性
 - ②対応文字コード(特に外字)の多様性
- 上記①の対応策の一つとして「中間標準レイアウト」が示されている。ここでは中間標準レイアウトの概要と有効性、活用方法について理解できる。
- 上記②の対応策の一つとして「文字情報基盤」が整備され、さらに「文字情報基盤」の自治体業務・システムにおける活用性について総務省で調査されている。ここでは文字情報基盤の概要と自治体業務・システムにおける有効性、活用方法について理解できる。

2. 本講義の構成

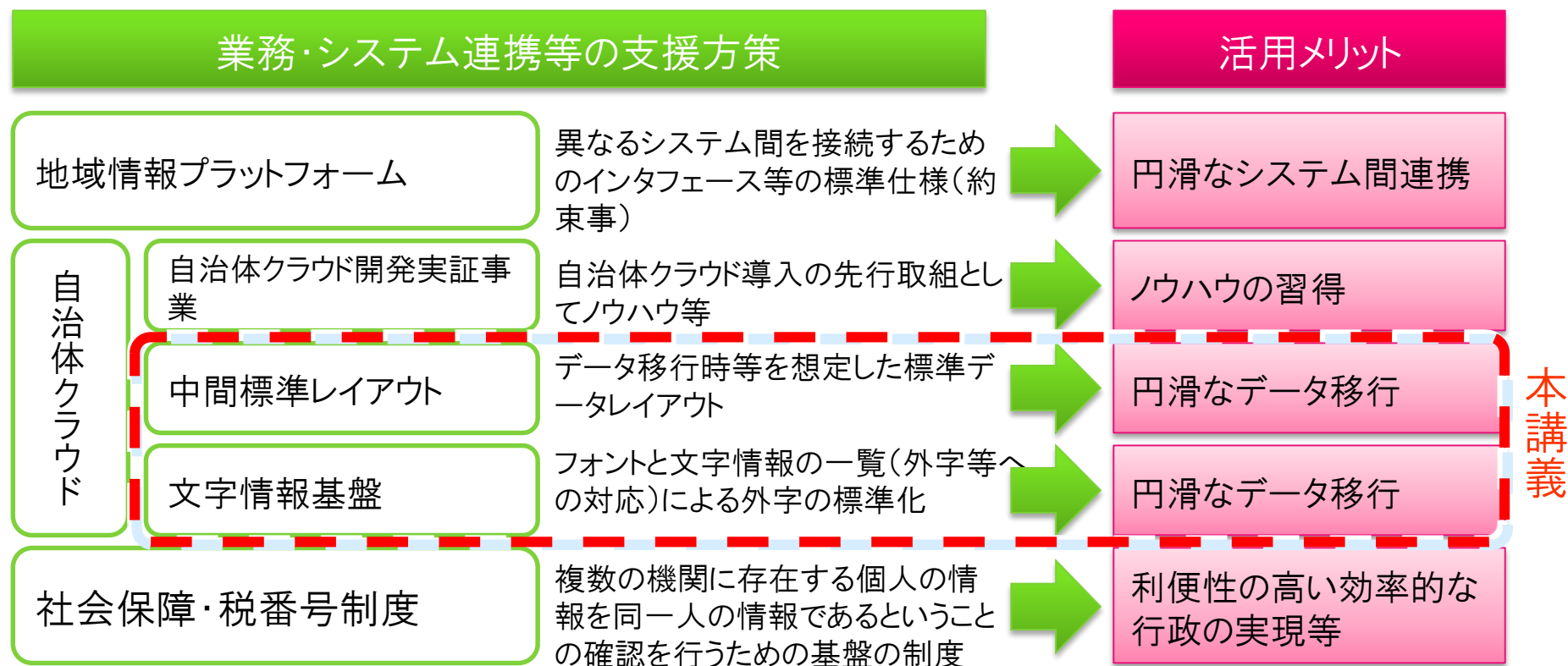
- 業務・システム連携等の支援方策
- 中間標準レイアウト
 - 背景・データ移行における課題
 - 中間標準レイアウトの考え方・概要
 - 中間標準レイアウトの有効性
 - 中間標準レイアウトの活用方法
- 文字情報基盤（外字実態調査を含む）
 - 自治体業務で使用する文字における課題
 - 文字情報基盤の概要
 - 外字実態調査の概要
 - 文字情報基盤の活用方法

3. 本講義の範囲



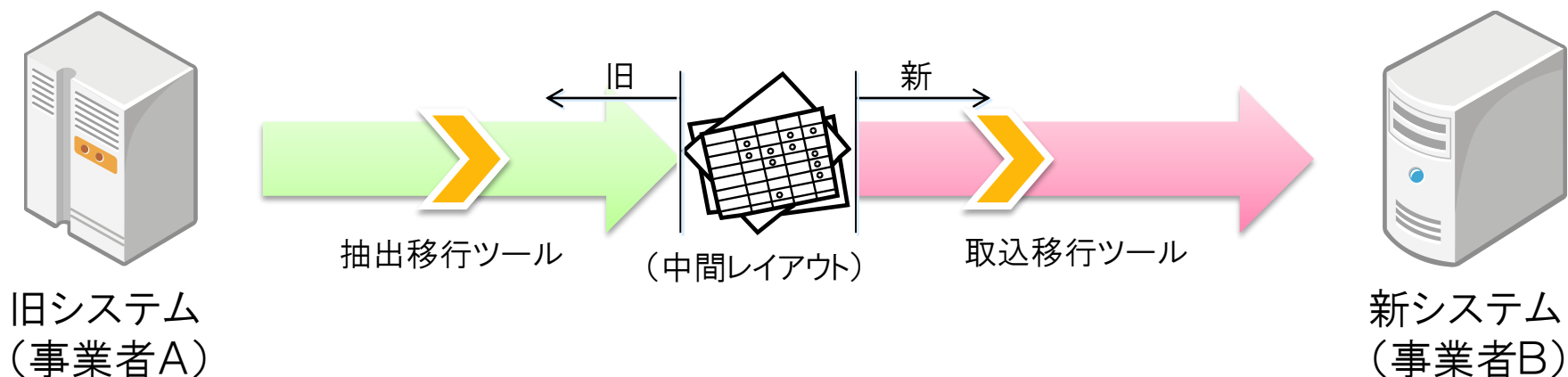
4. 業務・システム連携等の支援方策

- 業務・システムの連携等を支援する方策として、以下が整備されている。⇒有効に活用し円滑に移行



5. 中間標準レイアウト ～背景～

■ システム移行(データ移行)におけるコスト



新システムと旧システムで事業者が異なる場合、移行用のソフトウェアツールを作成する必要がある。

事業者A・・・旧システム用データから中間レイアウトへ変換する「抽出移行ツール」

事業者B・・・中間レイアウトから新システム用データへ変換する「取込移行ツール」



移行ツールは移行毎に開発が必要／中間レイアウトは各自治体仕様

5. 中間標準レイアウト ～データ移行における課題～

■ 移行時の課題

移行ツール部分の開発費用：

- 中間レイアウトの解釈や不明確要素に対するリスク ⇒ リスク費用として見積り
- 入念なテストの必要性
- 移行ツールは自治体毎の仕様・設計となるため他事例との比較が困難
- 上記により開発費用部分の妥当性評価が困難

移行ツール開発費

〔設計・開発であり開発費用の妥当性評価が困難〕

移行作業費

〔手作業による移行（入力等）、確認作業、作業項目、作業量はイメージしやすい〕

移行ツール開発費用の高額化

事実上の
ベンダロックイン

5. 中間標準レイアウト ～中間標準レイアウトの考え方～

■ 中間標準レイアウトの考え方

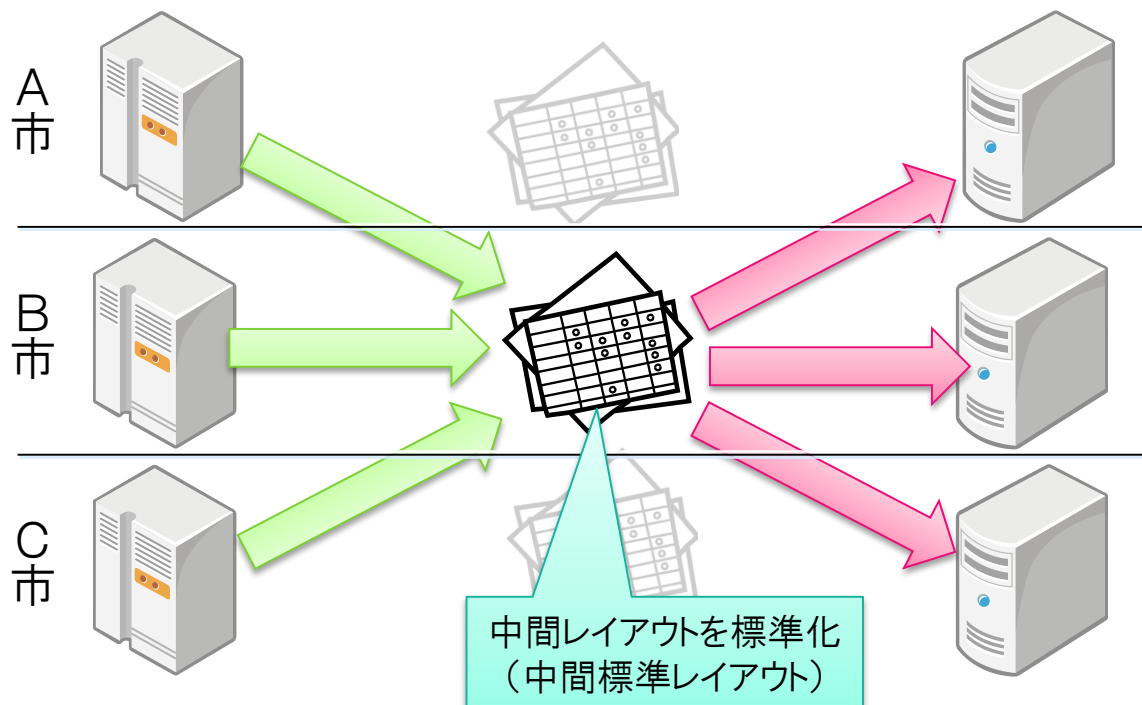
- 自治体毎に異なっていた中間レイアウトのデータ項目やその表現形式等を統一
- データ移行時に共通的に利用できる中間標準レイアウトを定義



- 事業者A、Bとも移行ツールの共通化が可能
- 自治体におけるシステム導入経費の削減と移行作業の軽減に寄与

旧システム
(事業者A)

新システム
(事業者B)



5. 中間標準レイアウト ～中間標準レイアウトの概要①～

■ 中間標準レイアウト仕様の概要

- 既存の標準を反映

地域情報プラットフォーム標準仕様

戸籍、後期高齢者医療等の標準仕様

- コンソーシアム3社による原案作成
- 協議事業者(24社)からの意見を反映
- 自治体職員及び有識者により構成された「自治体クラウドの円滑なデータ等に関する研究会」からの意見を反映
- 以下の22業務システムについて作成

1 住民基本台帳*
2 印鑑登録
3 住登外管理
4 戸籍
5 就学
6 選挙人名簿管理
7 固定資産税
8 個人住民税

9 法人住民税
10 軽自動車税
11 収滞納管理
12 国民健康保険
13 国民年金
14 介護保険
15 後期高齢者医療
16 健康管理

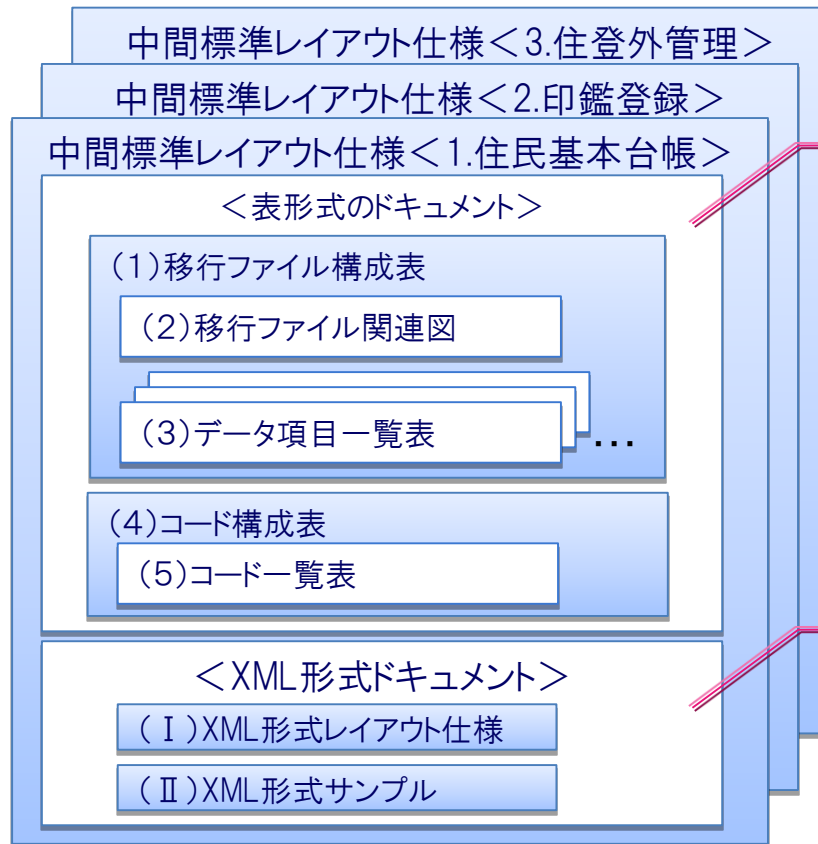
17 子ども手当
18 生活保護
19 障害者福祉
20 財務会計
21 人事給与
22 文書管理

* 外国人住民関係の改正に対応

出典)「自治体クラウドの円滑なデータ移行等に関する研究会とりまとめ(概要)」(自治体クラウドの円滑なデータ移行等に関する研究会)をもとに作成

5. 中間標準レイアウト ～中間標準レイアウトの概要②～

- 中間標準レイアウト仕様として表形式(5種類)、XML形式(2種類)のドキュメントを対象業務ごとに作成。



【表形式の中間標準レイアウト仕様のドキュメント】

- (1)移行ファイル構成表
…データ移行時に移行するファイル構成を一覧で示す
- (2)移行ファイル関連図
…データ移行時に移行するファイル間での関連を示す
- (3)データ項目一覧表
…移行ファイル内のデータ項目を一覧で示す
- (4)コード構成表
…移行ファイル内で使用されるコードを一覧で示す
- (5)コード一覧表
…移行ファイル内で使用されるコードのコード値と
その内容の一覧

【XML形式の中間標準レイアウト仕様のドキュメント】

- (I)XML形式レイアウト仕様
…表形式のデータ項目一覧表に対応したXML形式
のレイアウト仕様
- (II)XML形式サンプル
…XML形式レイアウト仕様に対応したXMLサンプル

出典)「自治体クラウドの円滑なデータ移行等に関する研究会とりまとめ(概要)」(自治体クラウドの円滑なデータ移行等に関する研究会)をもとに作成

5. 中間標準レイアウト ～中間標準レイアウトの有効性①～

■ 中間標準レイアウトの有効性(評価方法)

中間標準レイアウトでのデータ項目の適合率【指標】

中間標準レイアウトがデータ移行の対象となるデータ項目をどの程度カバーしているかの指標

【評価値の計算式】

$$\text{中間標準レイアウトでのデータ項目の適合率(\%)} = \frac{\text{中間標準レイアウトのデータ項目数}}{\text{移行対象のデータ項目数}} \times 100$$

■ 移行対象のデータ項目数:

主に自治体業務システムのパッケージ開発事業者の提供する業務システムパッケージ製品が保有するデータ項目のうち、システム制御や作業用のデータ項目等を除いたもの

5. 中間標準レイアウト ～中間標準レイアウトの有効性②～

■ 中間標準レイアウトの有効性(評価結果)

No.	対象業務システム	【指標】 評価値 (a/b)	試算に用いた数値		No.	対象業務システム	【指標】 評価値 (a/b)	試算に用いた数値	
			中間標準レイアウト のデータ項目数 (a)	移行対象の データ項目数 (b)				中間標準レイアウト のデータ項目数 (a)	移行対象の データ項目数 (b)
1	住民基本台帳	95%	268	281	12	国民健康保険	82%	903	1100
2	印鑑登録	66%	23	35	13	国民年金	93%	185	200
3	住登外管理	100% _(※1)	151	118	14	介護保険	61%	1196	1971
4	戸籍	100% _(※2)	3227	2300	15	後期高齢者医療	73%	574	790
5	就学	97%	95	98	16	健康管理	33%	435	1316
6	選挙人名簿管理	97%	114	118	17	子ども手当	100% _(※1)	116	79
7	固定資産税	91%	821	898	18	生活保護	85%	2012	2380
8	個人住民税	98%	536	547	19	障害者福祉	90%	1374	1529
9	法人住民税	100% _(※1)	190	173	20	財務会計	83%	216	260
10	軽自動車税	84%	108	128	21	人事給与	59%	928	1582
11	収滞納管理	95%	756	800	22	文書管理	63%	177	279

※1 住登外管理、法人住民税、子ども手当については、「口座」「送付先」等に関する情報を住登外管理側で保持するのか、それとも各業務側で保持するかといった調整により追加した結果、中間標準レイアウトのデータ項目数が増加した。(評価値が100%以上になったところは、100%と示す。※2も同様)

※2 戸籍については、法務省の標準仕様をスリム化して実現したパッケージ製品と、法務省の標準仕様にあわせて作成した今回の中間標準レイアウトを比較したため、後者のデータ項目数が多い結果となっている。

対象業務システムの半分以上で、評価値が80%を超える結果となった

なお、中間標準レイアウトのデータ項目は、地域情報プラットフォーム標準仕様の自治体業務アプリケーションユニット標準仕様V2.3における該当する業務ユニットのデータ項目をすべて採用している。

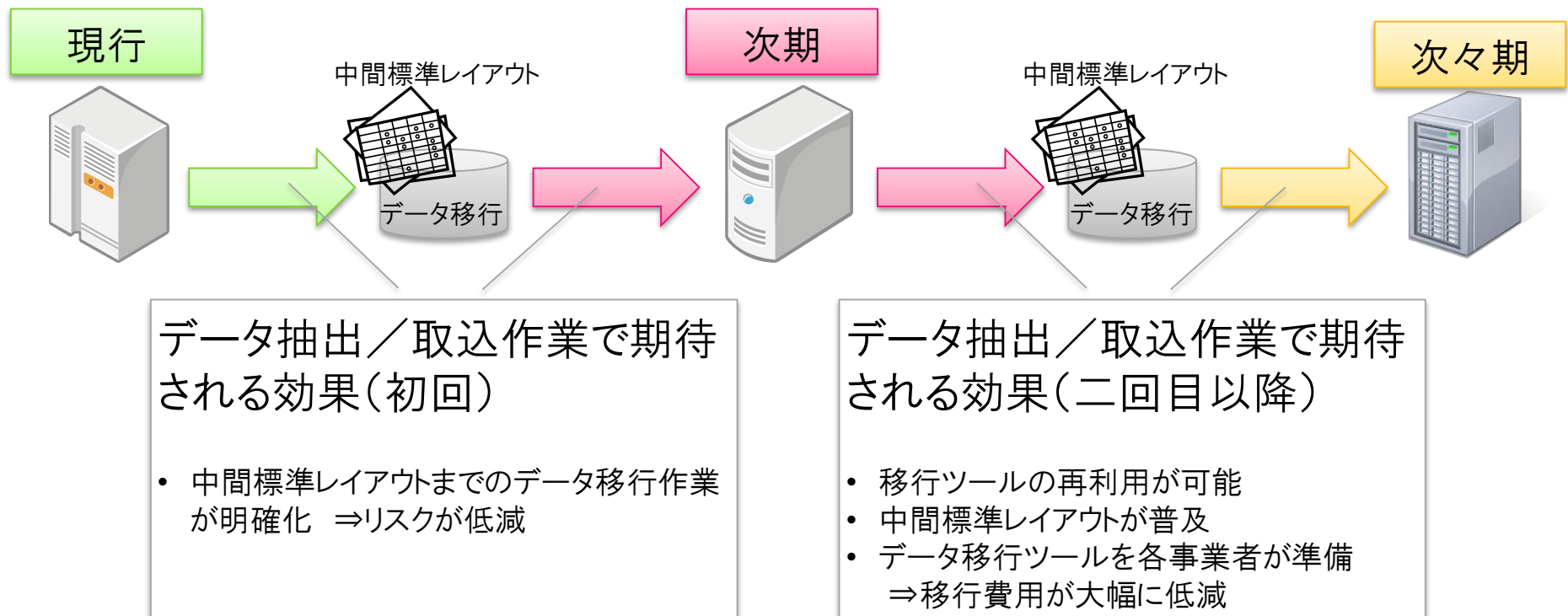
出典)「自治体クラウドの円滑なデータ移行等に関する研究会とりまとめ(概要)」(自治体クラウドの円滑なデータ移行等に関する研究会)をもとに作成

2-4 円滑なデータ移行に向けた方策

5. 中間標準レイアウト ～中間標準レイアウトの有効性③～

■ 移行費用削減効果 ～考え方～

- 中間標準レイアウトの普及が進めば、さらに大きな効果が期待できる
- 次期システム更新時(初回)、次々期システム更新時(二回目以降)におけるコスト削減効果について分析



5. 中間標準レイアウト ～中間標準レイアウトの有効性④～

■ 移行費用削減効果の試算結果（次期システムリプレイス時）

従来のデータ移行では、データ形式が不明な部分が多いため、リスク分にあたる移行ファイル数、データ項目数における安全係数を1.5と仮定。
【従来のデータ移行の見積り時に用いる移行ファイル数、データ項目数】
・移行ファイル数：10ファイル×1.5=15ファイル
・データ項目数：200項目×1.5=300項目

	従来のデータ移行	中間標準レイアウトを活用したデータ移行 (次期システムリプレイス時)	
		中間標準レイアウト適用可能	中間標準レイアウト適用不可
①データ抽出作業	ファイル:15,データ項目:300 ※計算式1 3.00人月	ファイル:10,データ項目:185 ※計算式2 1.91人月	データ項目:15 ※計算式3 0.09人月
②③打ち合わせ調整作業	データ項目:300 ※計算式4 0.15人月	作業なし	データ項目:15 ※計算式5 0.01人月
④データ変換ツールの作成作業 (①と同等の作業量)	3.00人月	1.91人月	0.09人月
小計	6.15人月	3.82人月	0.19人月
合計	6.15人月		4.01人月

中間標準レイアウトの適用によるコスト削減効果

2.14人月

約35%削減

計算式1:100+10×20+200×20×5%+100+100=700step(※20=300項目/15ファイル 1ファイルあたりの平均データ項目数)
700step×15ファイル/3,500step=3.00人月
計算式2:100+10×18.5+200×18.5×5%+100+100=670step(※18.5=185項目/10ファイル 1ファイルあたりの平均データ項目数)
670step×10ファイル/3,500step=1.91人月
計算式3:10×15+200×15×5%=300step(データ項目数のみの追加なので、項目数により変動する数値のみを作業の増分とする。)
300step/3,500step=0.09人月
計算式4:300/100=3人日
3人日/20日=0.15人月(1ヵ月を20日とした場合)
計算式5:15/100=0.15人日
0.15人日/20日=0.01人月(1ヵ月を20日とした場合)

自治体の業務システムにおいてクラウド化が進んだ場合や、中間標準レイアウト向けの汎用ツールを事業者が使用する場合は、提示した作業工数よりも、さらに作業量が低減されと考えられる。

※ 本試算結果は、サンプルとした業務システムにおけるシミュレーションの結果であるため、移行対象とする業務システムの状況等によって得られる効果は変動することが考えられる。

出典)「中間標準レイアウトの有効性に関する調査研究報告書」(総務省)をもとに作成

5. 中間標準レイアウト ～中間標準レイアウトの有効性⑤～

■ 移行費用削減効果の試算結果（次々期システムリプレイス時）

	前回のデータ移行	中間標準レイアウトを活用したデータ移行 （次々期システムリプレイス時）	
		中間標準レイアウト適用可能	中間標準レイアウト適用不可
①データ抽出作業	※次期システムリプレイス時における中間標準レイアウトを活用したデータ移行での作業工数	ファイル:10,データ項目:185 ※計算式1 0.19人月	データ項目:15 ※計算式2 0.09人月
②③打ち合わせ調整作業		作業なし	データ項目:15 ※計算式3 0.01人月
④データ変換ツールの作成作業(①と同等の作業量)		0.19人月	0.09人月
小計		0.38人月	0.19人月
合計	4.01人月	0.57人月	

中間標準レイアウトの普及によるコスト削減効果

3.44人月

約86%削減

計算式1: $100 + 10 \times 18.5 + 200 \times 18.5 \times 5\% + 100 + 100 = 670\text{step}$ (※18.5=185項目/10ファイル 1ファイルあたりの平均データ項目数)
 $670\text{step} \times 10\text{ファイル} / 3,500\text{step} = 1.91\text{人月}$
 $1.91\text{人月} / 10 = 0.19\text{人月}$ (10サイトで按分した場合)
 計算式2: $10 \times 15 + 200 \times 15 \times 5\% = 300\text{step}$ (データ項目数のみの追加なので、項目数により変動する数値のみを作業の増分とする。)
 $300\text{step} / 3,500\text{step} = 0.09\text{人月}$
 計算式3: $15 / 100 = 0.15\text{人日}$
 $0.15\text{人日} / 20\text{日} = 0.01\text{人月}$ (1ヵ月を20日とした場合)

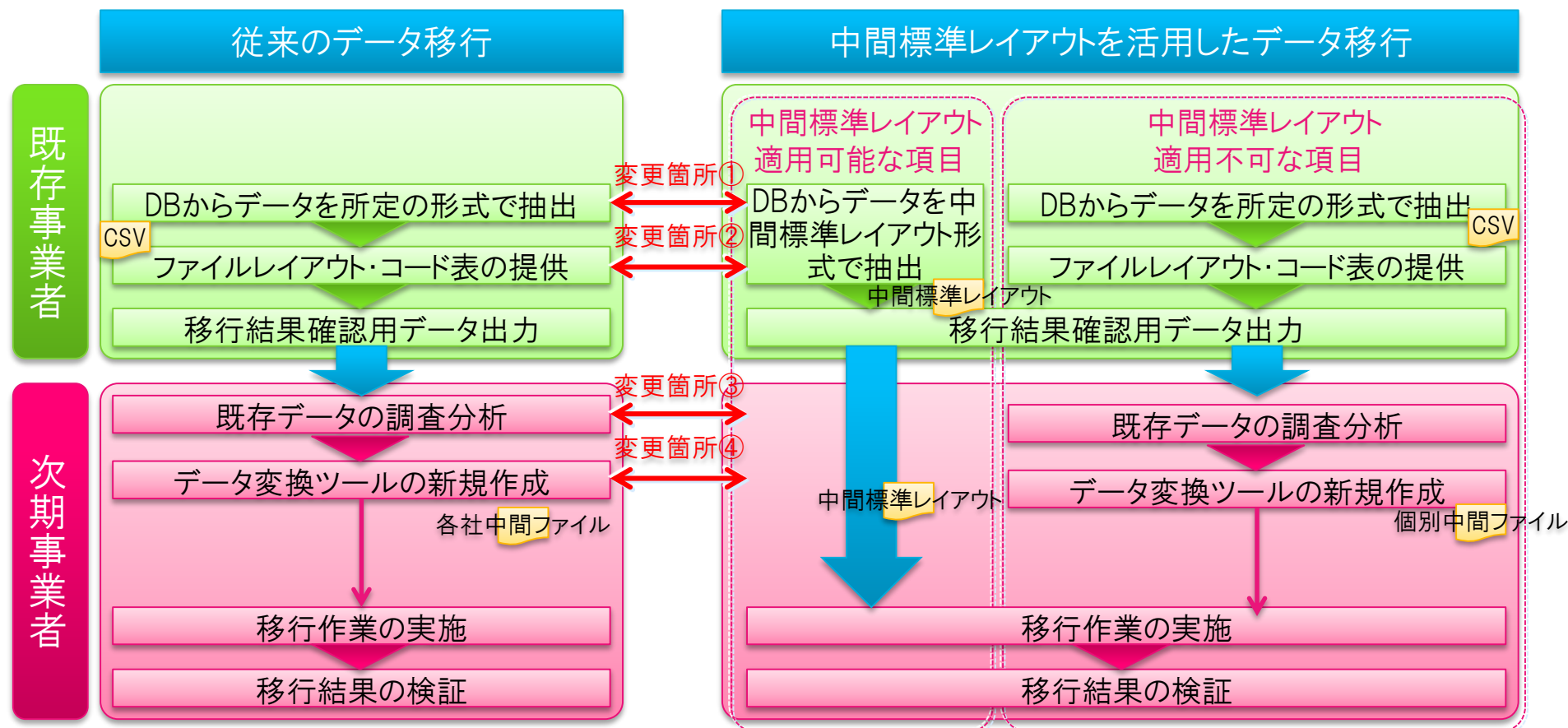
一度開発した中間標準レイアウト適用可能なデータ項目に対するデータ移行ツールを汎用的に使用することを想定し、当該データ移行ツールの作成工数について使用が想定される団体数で按分。
 本試算では、10団体でのデータ移行で使用すると仮定。

※ 本試算結果は、サンプルとした業務システムにおけるシミュレーションの結果であるため、移行対象とする業務システムの状況等によって得られる効果は変動することが考えられる。

出典)「中間標準レイアウトの有効性に関する調査研究報告書」(総務省)をもとに作成

5. 中間標準レイアウト ～中間標準レイアウトの活用方法①～

■ 想定される移行作業のフロー



出典)「中間標準レイアウトの有効性に関する調査研究報告書」(総務省)をもとに作成

5. 中間標準レイアウト ～中間標準レイアウトの活用方法②～

■ 中間標準レイアウトの活用方法

- 中間標準レイアウトを使用した場合、人口規模が同程度の自治体ではデータ移行費用も同程度と想定される。 ⇒ 先行事例の実績額から移行費用の評価が可能
- 契約期間満了時に業務システムが保有する実データを、中間標準レイアウトに対応した形式(中間標準レイアウト対応不可分は個別レイアウト)でデータ提供する旨を 調達仕様書に明記
- 移行データのデータベースを含む納入成果物に関する 著作権は自治体に帰属する旨を調達仕様書等に明記

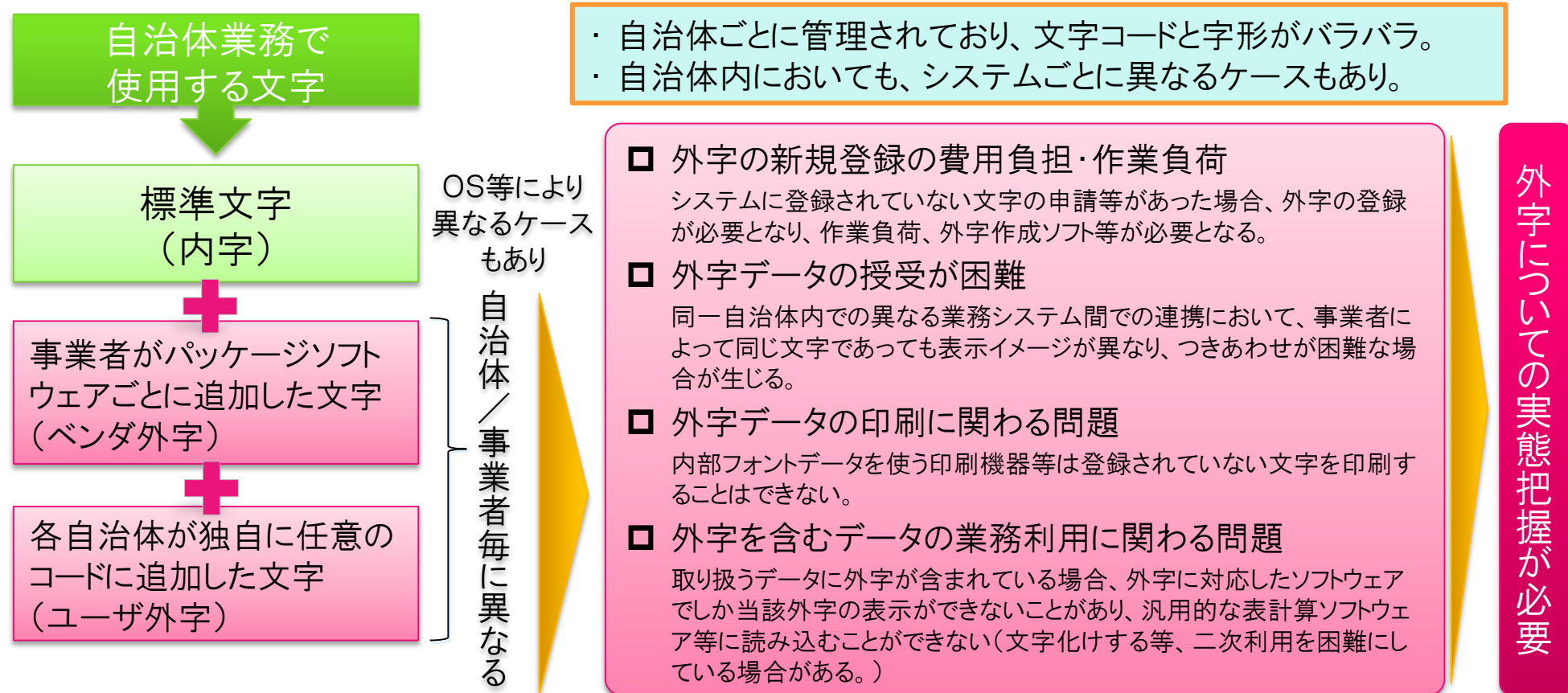
※「移行データのデータベースは著作権法におけるデータベースの著作物に該当する」として旧事業者が著作権の移転に係る追加的な費用を請求する懸念があるが、自治体の情報の選択又は体系的な構成に創作性を認めることは一般に困難であり、個々のデータそれ自体は一般的に創作性は認められず、その場合、著作権法第12条の2第1項にはあたらないとされる。

⇒ 調達仕様書に明記し無用な混乱を回避しておく方が良い。

- その他に以下の用途(業務連携)に活用可能
 - 異なる業務システム間連携のための標準データ形式として活用
 - 他自治体とのデータ交換時の標準データ形式として活用

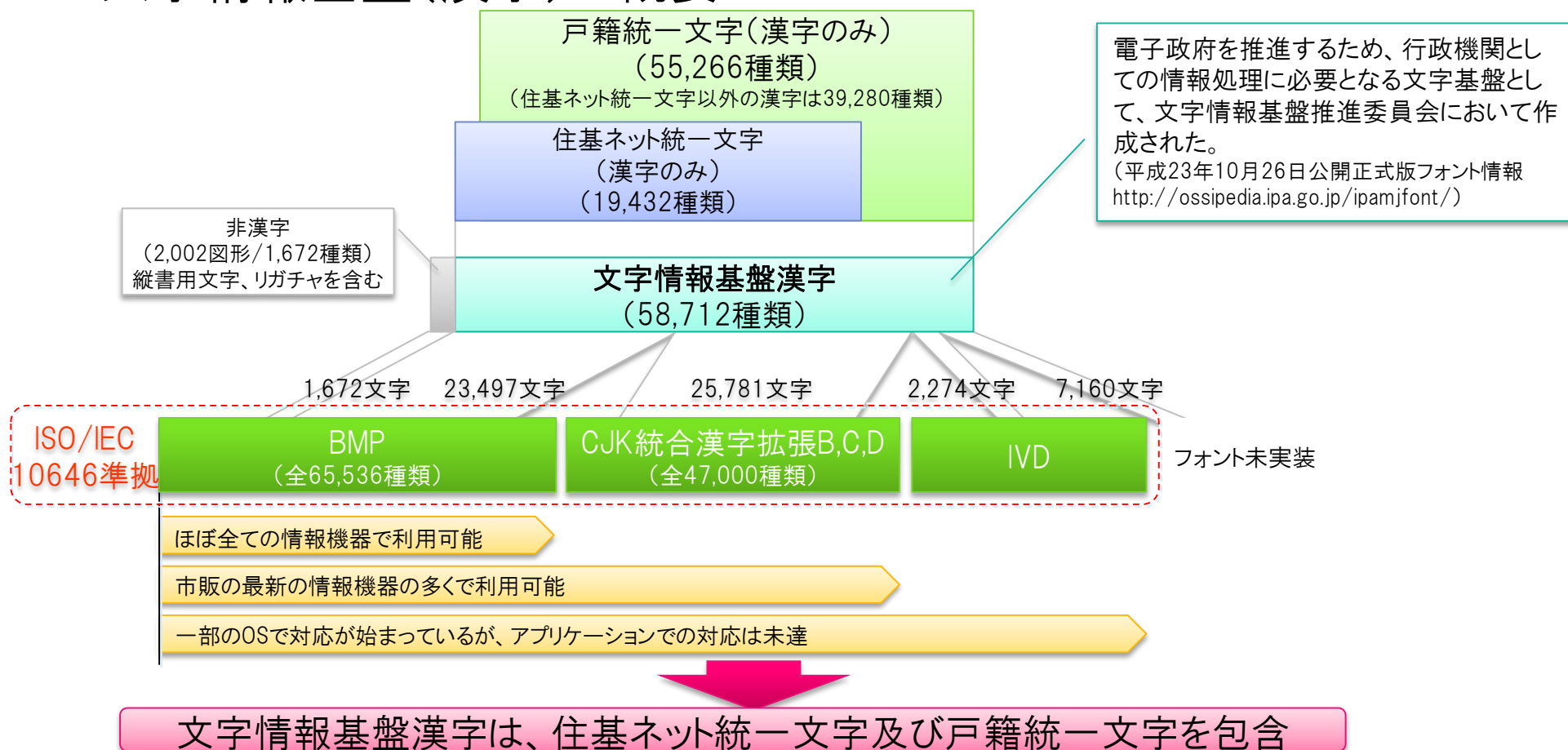
6. 文字情報基盤 ～自治体業務で使用する文字における課題～

■ 自治体業務で使用する文字における課題



6. 文字情報基盤 ～文字情報基盤の概要①～

■ 文字情報基盤(漢字)の概要



出典)「自治体クラウドの円滑なデータ移行等に関する研究会とりまとめ」(自治体クラウドの円滑なデータ移行等に関する研究会)をもとに作成

6. 文字情報基盤 ～文字情報基盤の概要②～

【参考】一般の文字(漢字)分類及び文字数

分類	字数	備考
常用漢字	1,945字	新常用漢字 : 2,136字
表外漢字	1,022字	印刷用文字
人名漢字 常用漢字 漢字の表(一) 漢字の表(二)	2,930字 1,945字 776字 209字	常用漢字の異体字でないもの 常用漢字の異体字
JIS (情報交換用文字) 第一水準 第二水準 第三水準 第四水準	10,038 字 2,965字 3,388字 1,249字 2,436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 準1級 2級	— 約6,000字 約3,000字 約1,945字	他に人名用漢字が対象となる

6. 文字情報基盤 ～外字実態調査の概要①～

■ 外字実態調査(外字の収集)

■ 調査対象

総務省から全国の市区町村に対して外字情報の提出を任意に依頼し、1,386団体(全市区町村の約8割)から提供を受けた外字(116万余)を調査対象とした。

<文字数>

提供市区町村数 1,386団体 / 総文字数 1,166,536文字

【内訳】

収集した外字の種類	概要	文字数
① 住基ネット残存外字	住基ネットに標準で搭載されていなかった文字で、市区町村が独自に登録、使用している文字	1,064,660文字
② 既存住基システム外字	既存住基システム外字のうち住基ネットで使用していない文字	69,142文字
③ その他システム外字	その他システムの外字のうち住基ネット及び既存住基システムで使用していない文字	32,734文字

これらの外字について前述「文字情報基盤漢字(58,712種類)」と
字形レベルの同定を行った。(＝文字情報基盤の対応状況を確認)

出典)「市区町村が使用する外字の実態調査」報告書(総務省)をもとに作成

6. 文字情報基盤 ～外字実態調査の概要②～

【参考】包摂基準に基づく分類

字形一致	すべての文字構成要素の配置・画数・形状が一致した文字	字形レベルで文字情報基盤漢字と一致している文字と判断
デザイン差	すべての文字構成要素の配置・画数が同じであるが、文字構成要素の形状の一部が、デザイン差の基準の範囲内で異なっている文字	
類似文字	文字情報基盤漢字に類似字形が存在するが、字形一致もしくはデザイン差の基準に当てはまらない文字	文字情報基盤漢字に置き換えて利用すべきか、そのまま利用すべきか判断が必要
同定不可能文字	文字情報基盤漢字に類似字形が存在しない文字、変体仮名及び記号等	

デザイン差・類似文字の例

同定不可能文字の例

変体仮名

記号

出典)「自治体クラウドの円滑なデータ移行等に関する研究会とりまとめ」(自治体クラウドの円滑なデータ移行等に関する研究会)をもとに作成

6. 文字情報基盤 ～外字実態調査の概要③～

■ 外字の実態調査結果

字形レベルの同定結果をもとに、「字形一致」「デザイン差」「類似文字」「同定不可能文字」の4つに分類

文字種類 包摂基準	文字情報基盤漢字 58,712種類		文字情報基盤漢字に同定できなかった文字 (同定不可能文字)		
	住基ネット統一文字 19,432種類	戸籍統一文字 39,280種類*	不明文字	変体仮名	記号等
字形一致 8.4% 98,030文字 (12,545種類)	74,040文字 (8,856種類)	23,990文字 (3,689種類)			
デザイン差 18.6% 217,313文字 (12,217種類)	166,452文字 (8,252種類)	50,861文字 (3,965種類)			
類似文字 53.1% 619,469文字 (14,222種類)	579,528文字 (10,293種類)	39,941文字 (3,929種類)			
同定不可能文字 19.9% 231,724文字			不明文字 52,294文字	変体仮名 97,791文字	記号等 81,639文字
	小計 820,020文字 (14,852種類)	小計 114,792文字 (6,709種類)			
総合計 1,166,536文字	合計 934,812文字(21,561種類)		合計 52,294文字	合計 97,791文字	合計 81,639文字

*戸籍統一文字39,280種類は、戸籍統一文字の総数55,266種類から住基ネット統一文字に含まれる15,986種類を除いた数字

出典)「自治体クラウドの円滑なデータ移行等に関する研究会とりまとめ」(自治体クラウドの円滑なデータ移行等に関する研究会)をもとに作成

6. 文字情報基盤 ～外字実態調査の概要④～

■ 類似文字の分析

要対応 : 今後、外字として対応が必要なもの

① 氏名に利用できるが、文字情報基盤漢字に存在しない新字体

類似文字のうち、昭和58年3月22日付け 法務省民二第1500号通達、平成2年10月20日付け 法務省民二第5200号通達 及び 平成6年11月16日付け 法務省民二第7005号・第7006号通達(漢和辞典に記載されている正字中の旧字体の構成要素部分に、新字体字形が使われている文字は氏名に利用できる)によって、氏名に利用できる新字体だが、文字情報基盤漢字に存在しない文字。

戸籍や住民票に
利用可能

要対応

② 「誤字俗字・正字一覧表」の誤字及び簡体字

類似文字のうち、平成22年11月30日付け法務省民一第2905号通達による「誤字俗字・正字一覧表」に示された誤字及び簡体字。

正字への置き換
えが進む

③ 「誤字俗字・正字一覧表」の俗字

類似文字のうち、平成22年11月30日付け法務省民一第2905号通達による「誤字俗字・正字一覧表」に示された俗字。

戸籍や住民票に
利用されている
可能性が高い

要対応

6. 文字情報基盤 ～外字実態調査の概要⑤～

■ 同定不可能文字の分析

要対応 : 今後、外字として対応が必要なもの

■ 不明文字

分類	分類の詳細		外字数	文字種類
①	漢和辞典に掲載されている文字	康熙字典	147文字	46種類
		新大字典	98文字	56種類
		その他の漢和辞典	555文字	92種類
②	戸籍関連通達及び法務局通知の文字		829文字	105種類
③	「誤字俗字・正字一覧表」の誤字		5,692文字	2,037種類
④	簡体字	入国管理局で利用される簡体字	2,145文字	795種類
		その他の簡体字	9,069文字	726種類
—	分類不可能だったもの		33,759文字	
合計			52,294文字	3,857種類

■ 変体仮名と記号等

- 変体仮名が多数存在
- 行政用途に利用されている記号等が存在

出典)「市区町村が使用する外字の実態調査」報告書(総務省)をもとに作成

要対応

6. 文字情報基盤 ～文字情報基盤の活用方法～

■ 活用方法

① 文字運用ポリシーの策定に活用

- 文字運用に役立つ定義がなされており、市区町村内の文字運用ポリシーとして流用し活用可能
- 文字運用ポリシーを策定し運用することで、外字を作成する基準が定義され、不要な外字を作成することがなくなり、文字運用負荷が軽減

② 文字情報基盤漢字とサイトの活用

- 文字情報基盤で提供されているIPAmj文字(フォント)をそのまま流用して、市区町村内の各種業務システム用外字として利用することで、外字作成の負荷を軽減
- IPA MJ文字情報検索システムではIPAmj全文字について、文字属性情報が設定されており、簡単に類似文字を検索でき、さらに戸籍統一文字であることも判別できるため、窓口での住民への説明に活用可能

③ 同定結果リスト(及びCSVファイル)の活用

※ 外字情報を提供した自治体に対して、総務省が提供した外字実態調査で実施した外字同定作業の当該自治体文の同定結果リスト及び同定結果CSVファイル

- 外字の整理に活用(以下例)
 - 住基ネット統一文字に一致した外字を削除し、本来の住基ネット統一文字コードに変更
 - 類似文字中の誤字は戸籍での使用文字の確認をして住民票等のデータを修正することで外字解消
- 文字環境の見える化に活用
 - 市区町村内の文字環境の見える化を行う際の参考、今後の文字整理方針の策定
 - 低解像度外字の再作成時の字形参照
- システム間連携やシステム移行時の基準文字に活用

7. 本講義のまとめ

- 移行や連携を円滑化する方策として次の2つが提供されており、積極的な活用が勧められる。
 - 中間標準レイアウト
 - 自治体の主要な22業務システムについて表形式、XML形式で提供されている。
 - 活用することにより将来的に大幅なデータ移行費の削減が期待される。
 - データの標準化により連携の促進が容易になる。
 - 文字情報基盤（外字実態調査を含む）
 - 戸籍統一文字と住基ネット統一文字をカバーした文字情報基盤が文字情報基盤推進委員会より提供されている。
 - 自治体1,386団体の外字を調査し同定したところ、既に27%程度の外字は一致している。
 - 他の外字についての整理や文字運用ポリシーの作成に活用可能である。
 - 文字基盤の連携性を高めることができる。