
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

平成 24 年 6 月

自治体クラウドの円滑なデータ移行等
に関する研究会

目 次

はじめに.....	1
第1章 中間標準レイアウトの活用について.....	2
1 データ移行費用の実態.....	2
1.1 データ移行費用の背景.....	2
1.2 自治体の抱える課題.....	3
1.3 解決への方向性.....	4
2 中間標準レイアウト仕様の作成.....	5
2.1 中間標準レイアウトとは.....	5
2.2 対象業務.....	5
2.3 作成手順.....	6
2.4 協議事業者等の参画状況.....	8
2.5 中間標準レイアウト仕様の構成.....	9
2.6 中間標準レイアウトの有効性.....	10
3 データ移行費用の削減方法.....	13
3.1 データ移行費用の内訳.....	13
3.2 データ移行費用の削減の流れ.....	14
3.3 データ移行費用の削減方法.....	15
3.4 コスト削減効果の試算.....	18
4 データ移行費用の抑制効果.....	20
5 中間標準レイアウトの具体的活用方法（例）.....	21
6 データ移行以外への活用方法.....	22
7 データ移行時の著作権.....	24
8 今後のフォローアップ.....	25
第2章 外字の課題と調査結果の活用について.....	26
1 外字の実態とその課題.....	26
1.1 外字の主な課題.....	26
1.2 システム移行時の課題と現状の対処.....	29
2 外字実態調査の報告.....	31
2.1 調査対象の外字情報.....	31
2.2 外字の同定.....	31
2.3 文字情報基盤漢字の概要.....	32
2.4 包摂基準の作成.....	33
2.5 同定作業.....	34
3 今回の同定結果の活用.....	37
3.1 外字発生の抑制.....	37
3.2 外字データの移行における活用.....	38
3.3 外字の整理に活用.....	39
3.4 文字環境の見える化に活用.....	41
3.5 システム間連携やシステム移行時の基準文字に活用.....	43
3.6 プリンタ・大量印刷機における活用.....	44
4 今後のフォローアップ.....	45
4.1 常用漢字が変更になった場合の対応.....	45
4.2 その他変更対応.....	45
用語集.....	46

はじめに

近年、クラウドサービスへの関心はますます高まってきている。先の東日本大震災より、自治体においても今後の情報システムの在り方そのものの見直しが迫られてきており、そのためクラウドサービスの活用について、その重要性もさらに増してきている。

総務省では、自治体クラウドの全国的展開に向けた取組を推進するとともに、住民サービス向上のための電子自治体の確立に向けて、全省を挙げて取り組むために「自治体クラウド推進本部」を設置した。そして、自治体における取組状況の把握、課題の抽出・検討及び必要な助言を行うことを目的として、有識者懇談会（以下「懇談会」という。）を設置し、検討課題について具体的な施策を検討、推進してきたところである。

懇談会での検討結果を踏まえ、総務省では平成 23 年度に、クラウドサービス間の相互運用性の確保のため、以下の2つの事項について取り組むこととされた。

一つ目は、新旧システム移行費用の低減を図るため、複数ベンダのシステムデータ構造への変換が容易な中間標準レイアウト作成のための標準化の範囲、データ項目の判定方法等についてのルールを調査・検討し、当該レイアウトを作成することである。

二つ目は、ユーザごとにカスタマイズを想定しないクラウドサービスにおいては、市区町村が独自に作成・保有する外字の標準化が不可欠となるため、市区町村独自の外字の実態調査を行い、同定作業を行うことである。

上記事項を実施するために、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移行等に関する研究会」（以下「本研究会」という。）が開催され、自治体クラウドへの円滑なデータ移行を実現するための中間標準レイアウトの作成方針及び自治体における外字の実態調査を適切に実施するための方策等を検討し、各市区町村におけるクラウド導入の際の参考となる報告書を取りまとめたところである。

第1章 中間標準レイアウトの活用について

1 データ移行費用の実態

1.1 データ移行費用の背景

(1) データの表現形式の不一致

自治体において現在利用されている業務システムでは、業界として各システムが保有するデータの表現形式の統一がなされていない。このため、システム移行時において、新旧システムのベンダが異なる場合には、データ項目の調整等（データの内容及びコードの確認等）多くの作業量とコストが生じている。

(2) 従来のデータ移行

システム移行時において、新旧システムでベンダが異なる場合には、移行対象データ項目等の不一致を解決させるために、移行用のソフトウェアツールを作成する必要がある。移行用のソフトウェアツールは、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」と、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」の2段階構成となることが一般的である。

さらに、これらツールでは対応できない例外的なデータ項目が残る場合には、手作業で対応する必要がある。

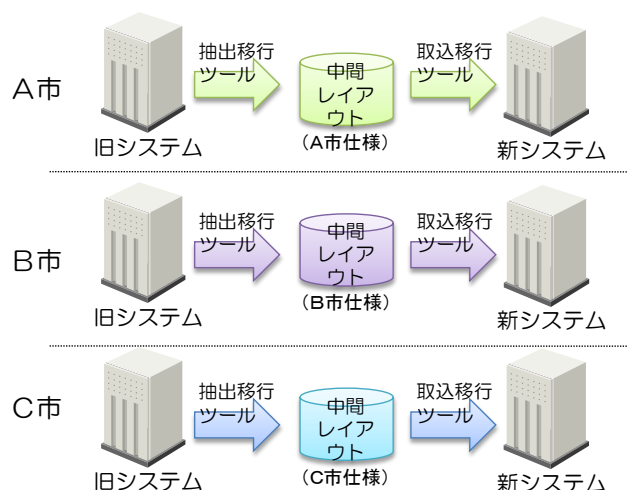


図 1-1 従来のデータ移行作業イメージ

ここで用いられる中間レイアウトは、自治体ごと、ベンダごとに内容が異なるものであり、汎用性がないため、データ移行のたびに自治体とベンダの双方にとって複雑な作業が生じている。

1.2 自治体の抱える課題

(1) 高額な見積り

従来のデータ移行では、移行先のシステムについて不明点が多かったことから、データ抽出作業及びデータ取込作業ともにリスクを含んだ多額の費用見積りになる傾向がある。

また、移行ツールは汎用性がないため、データ移行ごとに開発され、繰り返し利用することができず、割高な費用見積りとなってしまっていることがある。

(2) 移行費用の精査が困難

自治体はベンダに対してデータ移行費用の内訳の提示を要求しても、ベンダから詳細な費用の内訳を受領することが難しいという実態が多い。また、各団体でデータ項目等の仕様が異なるため、費用の内訳を受領できたとしても、他団体と費用を比較することが困難である。

このため、自治体においてはデータ移行費用の内容精査が困難となっている。

(3) 自由なサービスの選択に制限

データ移行にかかる膨大な作業量とコストは、異なるベンダ間でデータ移行を行う場合に発生するものであり、同一ベンダによるデータ移行では、当然リスク分が削られるため、より安価な費用で対応することができる。

このように膨大な作業量とコストを避けるため、事実上、システム更新時にベンダを変更できない、いわゆるベンダロックインが発生する原因ともなっている。

1.3 解決への方向性

(1) 中間標準レイアウトの適用

データ移行費用の低減を図るため、データ項目やその表現形式等を統一した、全国の自治体が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作成する。

中間標準レイアウトを移行ファイル形式に適用することにより、新旧システムの両ベンダは事前にデータ移行に必要な準備を行えることなどから、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を期待することができる。

(2) 新しいデータ移行

データ移行ツールを作成する際に、中間標準レイアウトを標準として定めることで、旧システムからデータを抽出する「抽出移行ツール」と、新システムへデータを取り込む「取込移行ツール」の開発をそれぞれ共通化することができる。

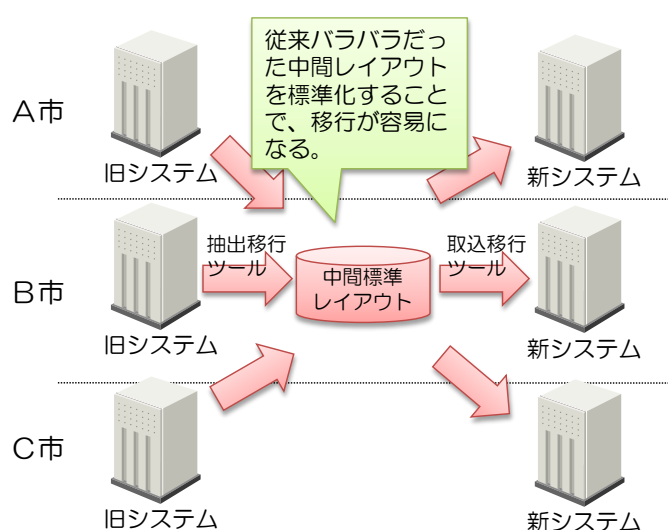


図 1-2 中間標準レイアウトによるデータ移行作業イメージ

ここで用いられる中間標準レイアウトは、データ移行時に必要とされるデータ項目の名称やデータ型などの移行レイアウト仕様を標準的に定めたものであり、自治体やベンダが汎用的に利用できるため、自治体とベンダの双方にとって従来のデータ移行手順よりも作業効率が向上する。

(3) 見積りへの反映

ベンダは、中間標準レイアウトという要件が確定されたデータ仕様に沿って移行ツールを開発することができるため、移行ツール開発上の不確定要素が減ることとなり、データ移行費用へ加算していたリスク分の費用を削減することが可能となる。

2 中間標準レイアウト仕様の作成

2.1 中間標準レイアウトとは

中間標準レイアウトとは、自治体の業務システムにおいて、旧システムから取り出したデータを新システムで使用するために用いる、標準化されたデータの表現形式である。

自治体が現在利用しているシステムから異なるベンダのシステムに更新しようとする際、業界としてのデータの表現形式の標準化が確立されていない状況下においては、データ移行に際しての新旧システムの仕様の確認、調整などの作業が発生し、コストがかかってしまうため、事実上、ベンダやサービスの自由な選択が制限されてしまうことがある。

ベンダ各社が中間標準レイアウトに対応するようになれば、異なるベンダの業務システムへのデータ移行作業を簡略化させるなど、円滑なデータ移行の実現が期待される。

2.2 対象業務

以下の 22 業務を対象として、中間標準レイアウト仕様を作成した。

表 1-1 作成対象業務一覧

No.		No.	
1	住民基本台帳（※）	12	国民健康保険
2	印鑑登録	13	国民年金
3	住登外管理	14	介護保険
4	戸籍	15	後期高齢者医療
5	就学	16	健康管理
6	選挙人名簿管理	17	子ども手当
7	固定資産税	18	生活保護
8	個人住民税	19	障害者福祉
9	法人住民税	20	財務会計
10	軽自動車税	21	人事給与
11	収滞納管理	22	文書管理

（※）平成 24 年 7 月に施行が予定されている「住民基本台帳法の一部を改正する法律」による外国人住民関係の改正に対応している。

2.3 作成手順

(1) 手順

中間標準レイアウト仕様の作成及び中間標準レイアウトの有効性に関する調査研究は、以下の手順で実施した。

作業ステップ	概要	中間標準レイアウト 作成請負事業者	協議 事業者	研究会
STEP1 事前準備	中間標準レイアウト仕様（案）の各ドキュメントの様式や作成ルールを準備する。	各ドキュメントの様式／作成ルール		
STEP2 原案の作成	中間標準レイアウト作成事業者の業務パッケージ等から検討を行い、中間標準レイアウト仕様（案）の原案を作成する。	中間標準レイアウト仕様（案）の原案		
STEP3 協議事業者からの意見収集	自治体の業務パッケージで共通に使用できる仕様とするため、協議事業者に中間標準レイアウト仕様（案）の原案を公開し、これに対する意見を収集する。		確認	
STEP4 有効性に関する調査研究	上記のSTEP2とSTEP3と並行して、中間標準レイアウトの有効性に関する調査を実施する。	有効性に関する調査研究		
STEP5 有識者からの意見収集	中間標準レイアウト仕様（案）及び有効性に関する調査研究について、本研究会にて説明し、有識者から意見を収集する。			確認
STEP6 表形式の仕様の作成	協議事業者及び本研究会（有識者）からいただいた意見の取り込みを実施し、表形式の中間標準レイアウト仕様を確定する。	中間標準レイアウト仕様〈表形式〉		
STEP7 XML形式の仕様の作成	確定した表形式の中間標準レイアウト仕様に基づき、XML形式の中間標準レイアウト仕様を作成する。	中間標準レイアウト仕様〈XML形式〉		
STEP8 報告書作成	上記STEP1からSTEP7までの成果をとりまとめ、中間標準レイアウト仕様及び有効性に関する調査研究報告書を作成する。	中間標準レイアウト仕様〈表形式〉 〈XML形式〉		有効性に関する調査研究報告書

図 1-3 中間標準レイアウトの作成手順

(2) 中間標準レイアウト仕様の項目選定の方針

中間標準レイアウト仕様のデータ項目は、システム間連携を目的とした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 で定義されている項目を包含し、業務を行う上で共通に必要なデータ項目を採用することとした。

(3) 対象項目における標準化内容

① 項目名称

データの内容としては他ベンダと共通であるが、項目名称が異なる場合は、最も一般的な項目名称を採用する。

② データ型

データ型がベンダ間で異なる場合は、最も共通的に使用され则认为られるデータ型を採用する。

③ 桁数

桁数が他ベンダと異なる場合は、原則として最大桁数を採用する。

④ その他の属性情報

最も共通的に使用され则认为られる形式を採用する。

(4) その他

中間標準レイアウト仕様は、表形式に加えて、XML 形式でも定義する。

2.4 協議事業者等の参画状況

(1) 協議事業者

前述「2. 3作成手順」STEP3で記した協議事業者として、総務省が募集した24社が参加した。

(2) 対応業務システム別の協議事業者

業務システム別の各協議事業者の参画状況は以下のとおりである。

表 1-2 協議事業者の参画状況

No.	業務システム	参画事業者数	意見数	No.	業務システム	参画事業者数	意見数
1	住民基本台帳	19社	35件	13	国民年金	15社	8件
2	印鑑登録	18社	17件	14	介護保険	12社	15件
3	住登外管理	18社	14件	15	後期高齢者医療	11社	4件
4	戸籍	6社	6件	16	健康管理	10社	22件
5	就学	13社	16件	17	子ども手当	11社	6件
6	選挙人名簿管理	16社	24件	18	生活保護	7社	35件
7	固定資産税	16社	107件	19	障害者福祉	11社	8件
8	個人住民税	16社	27件	20	財務会計	11社	19件
9	法人住民税	16社	18件	21	人事給与	9社	26件
10	軽自動車税	16社	41件	22	文書管理	8社	36件
11	収滞納管理	16社	48件	—	業務共通	—	5件
12	国民健康保険	16社	72件				
						合計	609件

(3) 本研究会からの意見反映

中間標準レイアウト仕様（案）に対する上記(2)の意見等に加え、本研究会構成員からの意見等も反映した。

2.5 中間標準レイアウト仕様の構成

中間標準レイアウト仕様は、以下の表形式（5種類）とXML形式（2種類）のドキュメントを、対象業務システムごとに作成した。

<表形式のドキュメント>

(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を一覧で示す

<XML形式のドキュメント>

(1) 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

(2) 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

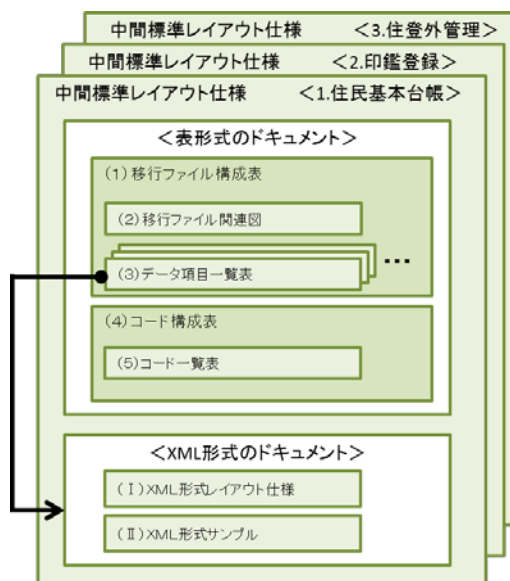


図 1-4 中間標準レイアウト仕様の構成

2.6 中間標準レイアウトの有効性

(1) 作成した中間標準レイアウトに関する分析

自治体業務システム間でのデータ移行への活用を想定した際の中間標準レイアウトの有効性を整理するため、中間標準レイアウトのデータ項目に着目した定量的な分析を実施した。

(2) 中間標準レイアウトでのデータ項目の適合率【指標】

① 【指標】の内容

本指標は、中間標準レイアウトのデータ項目が、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での移行対象データ項目に対して、どの程度適合しているかを測ることを目的とした指標である。

本指標に関する評価値を試算することで、対象業務システムの中間標準レイアウトが、自治体業務システムのデータ移行時に対象となるデータ項目をどの程度カバーしているか把握することができ、業務システムのデータ移行における中間標準レイアウトの有効性を確認することができる。

② 評価値の試算方法

本指標の評価値は、中間標準レイアウトが自治体業務システムのパッケージ製品の仕様をベースに作成されていることを前提として、パッケージ製品が固有で保有するデータ項目を除いた当該業務で移行対象となるデータ項目数（下図内（2））に対する、中間標準レイアウトのデータ項目数（下図内（3））の割合で試算する。

業務システムのパッケージ製品が保有するデータ項目（下図内（1））には、システムの制御用や作業用のデータ項目等のパッケージ製品固有のデータ項目が含まれているため、それらを除いたデータ項目が、当該業務での移行対象となるデータ項目（下図内（2））となる。そして、その移行対象となるデータ項目の中から、標準化したデータ項目が中間標準レイアウトのデータ項目（下図内（3））となる。

本指標の評価値を試算するために用いるデータ項目数の関係イメージ及び評価値の試算式を以下に示す。

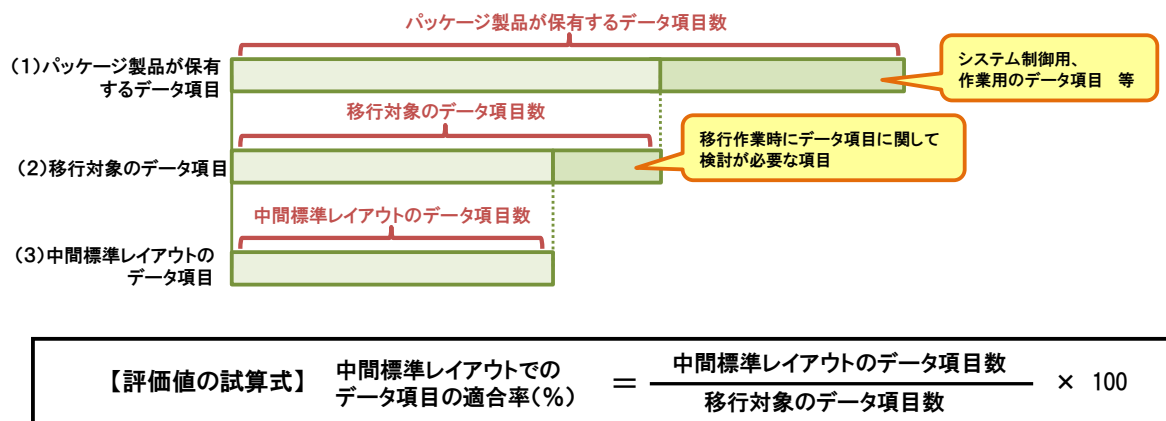


図 1-5 【指標】の試算に用いるデータ項目数の関係イメージ及び評価値の試算式

- ③ 各対象業務システムにおける評価値の試算結果
 対象業務の半分以上で、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仕様のデータ項目は、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 において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。

表 1-3 各対象業務システムにおける評価値の試算結果

No.	対象業務システム	【指標】の評価値 (a/b)	評価値の試算に用いる数値	
			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)
1	住民基本台帳	95%	268	281
2	印鑑登録	66%	23	35
3	住登外管理	100% (※1)	151	118
4	戸籍	100% (※2)	3,227	2,300
5	就学	97%	95	98
6	選挙人名簿管理	97%	114	118
7	固定資産税	91%	821	898
8	個人住民税	98%	536	547
9	法人住民税	100% (※1)	190	173
10	軽自動車税	84%	108	128
11	収滞納管理	95%	756	800
12	国民健康保険	82%	903	1,100
13	国民年金	93%	185	200
14	介護保険	61%	1,196	1,971
15	後期高齢者医療	73%	574	790
16	健康管理	33%	435	1,316
17	子ども手当	100% (※1)	116	79
18	生活保護	85%	2,012	2,380
19	障害者福祉	90%	1,374	1,529
20	財務会計	83%	216	260
21	人事給与	59%	928	1,582
22	文書管理	63%	177	279

(※) 【指標】の評価値が100%以上になったところは、100%と示す。

(※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。

(※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。

(3) 中間標準レイアウトの活用による効果について

中間標準レイアウトは、システムのオープン化及び自治体クラウドへの移行において、有効性を発揮することが期待される。

まず、大きな効果の一つとして、旧システムから新システムへのデータ移行時におけ

るコストの削減が挙げられる。特にデータ移行費用については前述の「1 データ移行費用の実態」でも述べているとおり、システム移行の際の大きな課題の一つであり、この課題の解決によってベンダロックインを防ぎ、円滑なデータ移行の実現を図ることができる。

データ移行費用の削減方法として想定される、以下の事項の内容については、後述の「3 データ移行費用の削減方法」、「4 データ移行費用の抑制効果」、「5 中間標準レイアウトの具体的活用方法（例）」において述べる。

- 移行ツール開発費の低減化
- 移行ツールの複数団体での再利用効果
- 移行ツール開発費における複数団体による割り勘効果
- データ移行費用の抑制効果
- 次々期システム調達時における活用

次に、データの表現形式を標準化することにより、データ連携を容易にする効果も挙げられる。他業務システム間の連携はもとより、異なるベンダのシステムとの連携が容易になることで、自治体業務の利便性向上を見込むことができる。

このようなシステム間のデータ連携への活用方法の内容については、後述の「6 データ移行以外への活用方法」において述べる。

3 データ移行費用の削減方法

3.1 データ移行費用の内訳

データ移行費用は、移行ツール開発費と移行作業費に分割することができる。

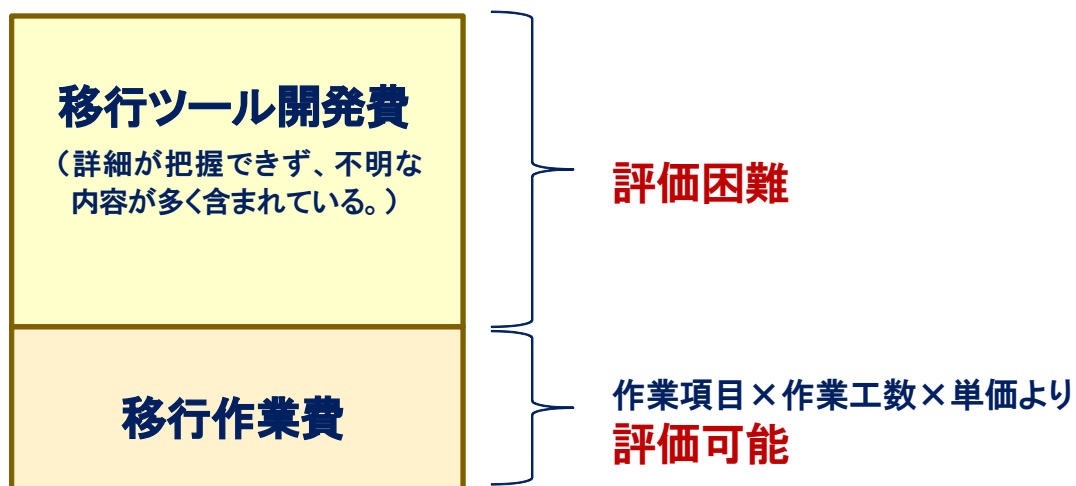


図 1-6 データ移行費用の実態イメージ

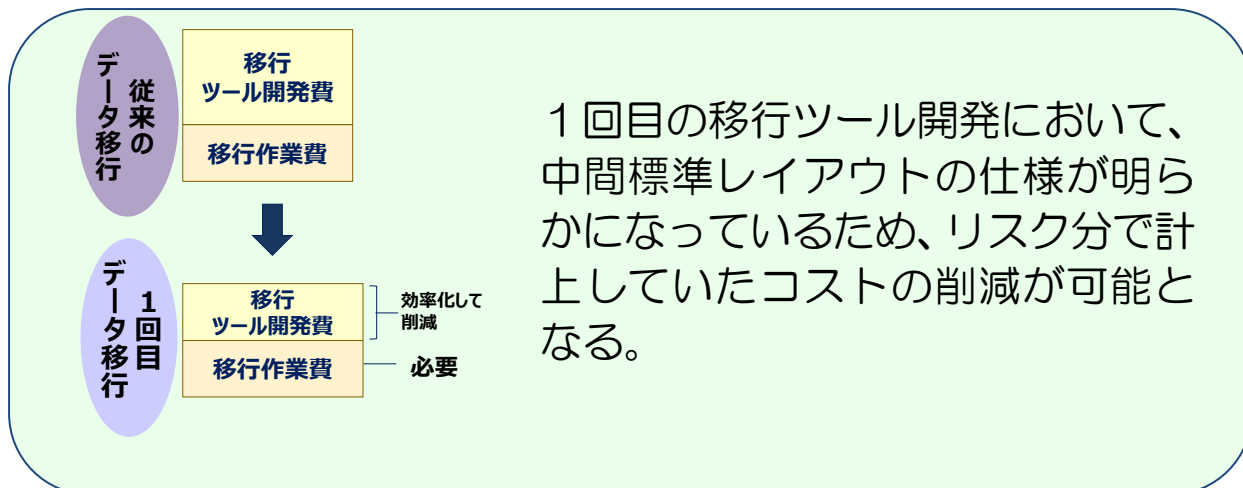
移行ツール開発費は、前述のような不確定なリスク要素を含んだ高額な費用となる傾向があり、ベンダから提示された見積り額の妥当性を評価することが困難である。

一方、移行作業費は、データ項目の調整作業や移行ツールで対応できないデータの手作業等にかかる費用であり、作業項目と作業工数はある程度可視化できるため、ベンダから提示された見積り額の妥当性を評価することが可能である。

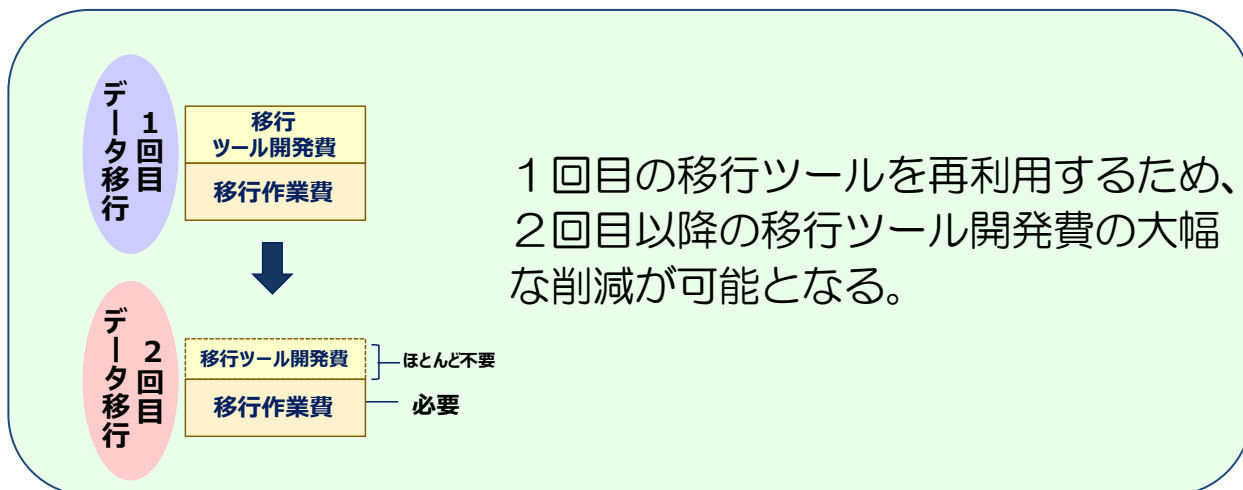
3.2 データ移行費用の削減の流れ

自治体業務システムへのクラウドサービスの導入が進展すると、一度開発した移行ツールの再利用や複数団体での活用が普及・促進することが予想される。データ移行の際に中間標準レイアウトを適用すると、以下のとおり、データ移行費用が削減していくと想定される。

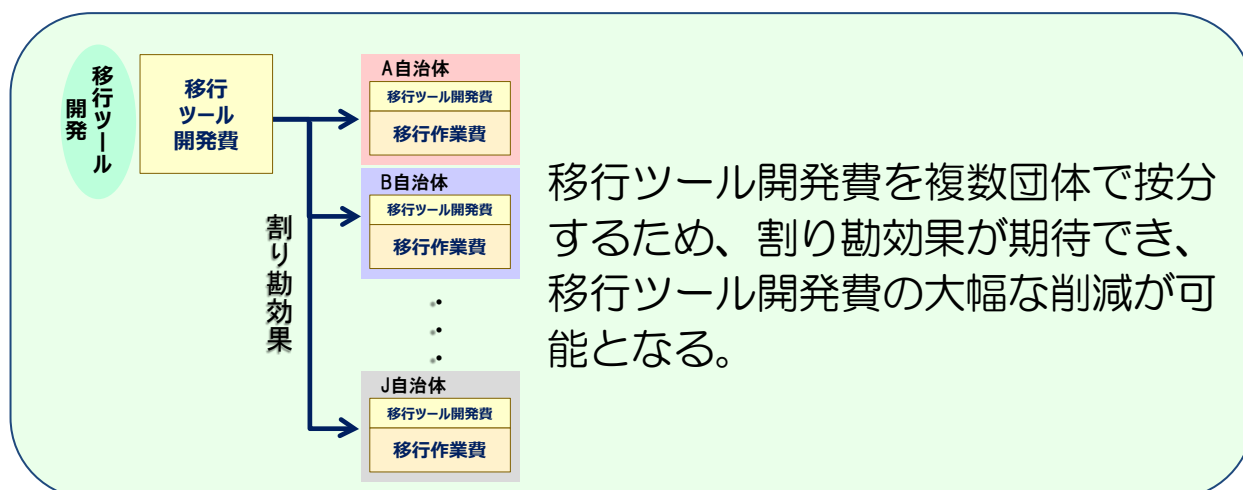
(1) 移行ツール開発費の削減（1回目）



(2) 移行ツールの再利用効果（2回目以降）



(3) 複数団体による割り勘効果（複数回）



3.3 データ移行費用の削減方法

中間標準レイアウトの活用によるデータ移行費用の削減方法について、各段階の具体的な内容を以下に記述する。

(1) 移行ツール開発費の削減（1回目）

① 問題点・課題

従来はデータ移行時の中間レイアウトに業界標準がないため、ベンダは、移行ツール開発においてデータ項目の不一致を解決させるために多くの作業工数が必要であり、移行ツール開発費が高額となる傾向がある。

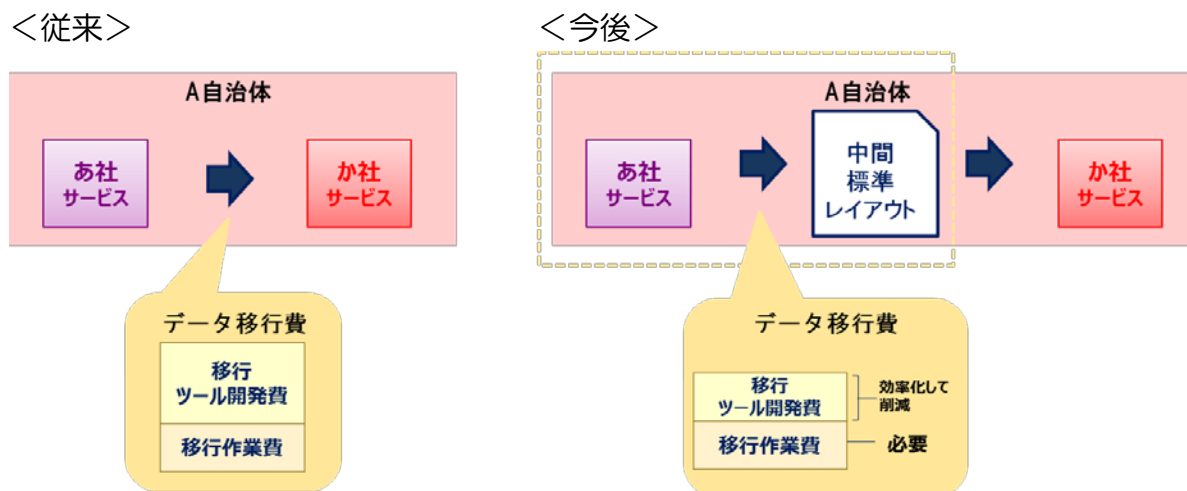
また、データ移行先システムのデータ形式が不明もしくは既存システムのデータ形式と不一致であるため、リスク分を含めた移行ツール開発費が請求される傾向がある。

② 対策

自治体は新たに業務システムを調達する際、データ移行の要件として中間標準レイアウトを採用する旨を調達仕様書へ明記する。

③ 得られる効果

中間標準レイアウトの仕様が明らかになっているため、移行ツール開発の作業効率が向上し、開発工数が減ることでデータ移行費用が削減される。



(2) 移行ツールの再利用効果（2回目以降）

① 問題点・課題

異なるベンダ間における新旧システム間のデータ移行の際に用いられる従来のデータ移行ツールは、自治体ごと、ベンダごとに開発され、汎用性がないため、繰り返し利用することができなかった。このため、データ移行のたびに移行ツール開発費が必要になっていた。

② 対策

自治体は新たに業務システムを調達する際、データ移行の要件として中間標準レイアウトを採用する旨を調達仕様書へ明記する。

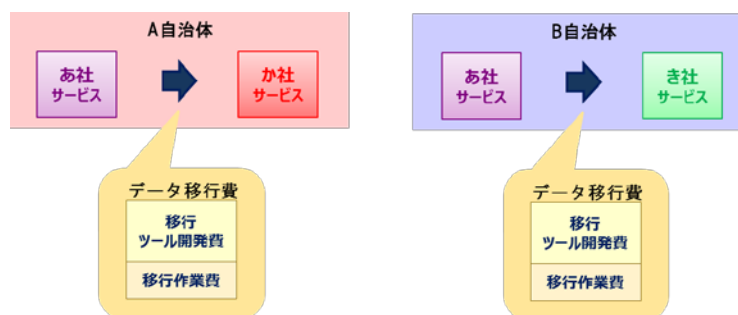
③ 得られる効果

全国の自治体がデータ移行の際の移行ファイルに中間標準レイアウトを採用することで、ベンダ各社は中間標準レイアウトへの移行を経験し、この実施ノウハウを社内に蓄積することができる。

これによりベンダは、1回目の移行ツールを再利用することで、2回目以降の移行ツール開発を大幅に効率化することができ、自治体はデータ移行費用を大幅に削減することが可能となる。

移行ツールの開発に必要となる作業は、各自治体の独自カスタマイズに対応するための最低限の作業だけとなる。

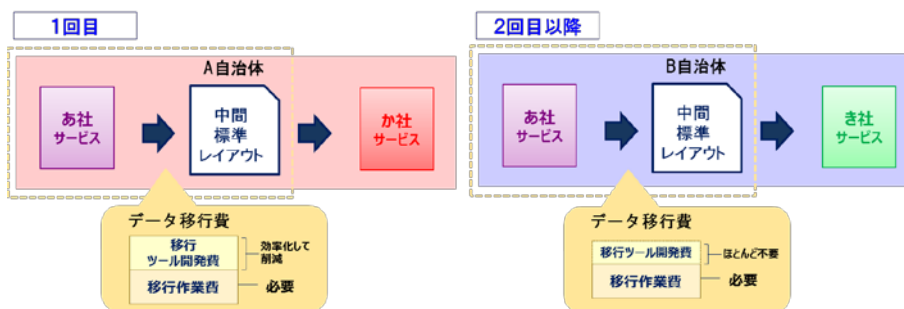
<従来>



繰り返し効果なし

移行先フォーマットがベンダごとに異なるため、毎回移行ツール開発費が必要になる。

<今後>



繰り返し効果あり

移行先フォーマット（中間標準レイアウト）が同一のため移行ツール開発費は2回目以降からほとんど不要になる。

図 1-8 移行ツール開発費の削減イメージ

(3) 複数団体による割り勘効果（複数回）

① 問題点・課題

複数団体によるシステムの共同利用を想定したクラウド環境においては、ベンダは中間標準レイアウトに対応した移行ツールの開発において、新たに設計、製造、試験という負担が生じる。

自治体としては中間標準レイアウトを適用する1回目のデータ移行となった場合、後発で追随する他自治体よりも、移行ツール開発費の負担が割高になってしまう可能性がある。

② 対策

複数団体へ同一業務システムを納入しているベンダは、今後発生するデータ移行の全体作業を見込み、これを利用団体の数による割り勘効果を反映したデータ移行費の提示が可能となる。

具体的には1回目の移行ツール開発費を、複数の利用自治体の数により按分した金額にすることで、利用団体による移行ツール開発費の負担差をなくすることができる。

③ 得られる効果

割り勘効果により、自治体1団体あたりが負担する移行ツール開発費が削減される。

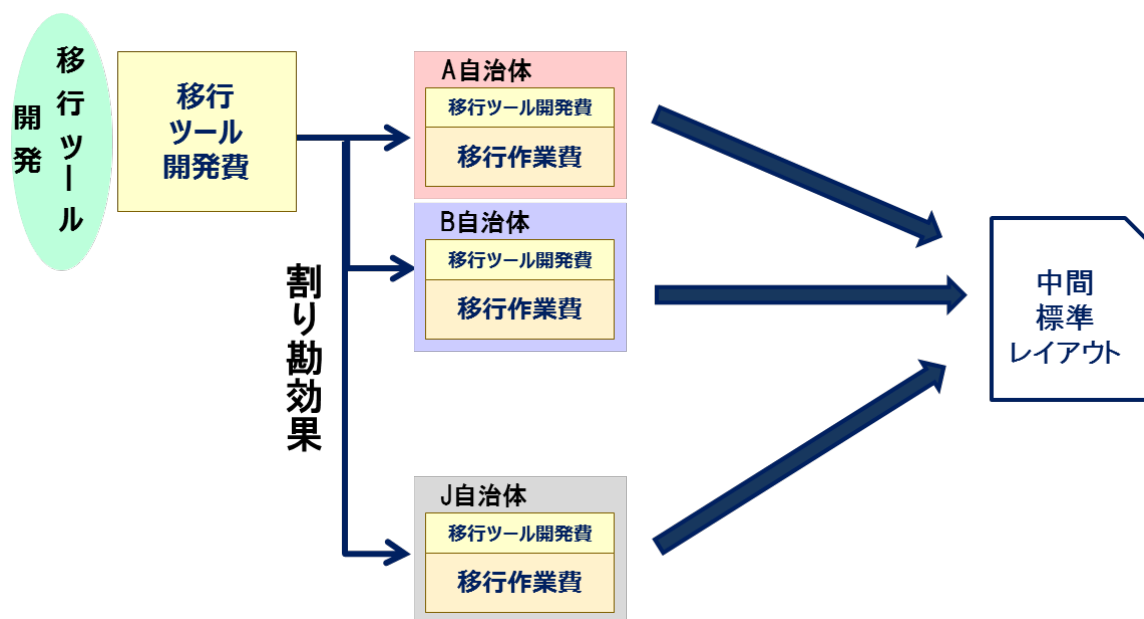


図 1-9 複数団体利用の割り勘効果によるコスト削減イメージ

3.4 コスト削減効果の試算

中間標準レイアウトを自治体業務システム間でのデータ移行に活用した際のデータ移行に関わるコスト削減効果は、中間標準レイアウトを活用するタイミングで異なるため、次期システムリプレース時、次々期システムリプレース時におけるコスト削減効果について分析を行うこととする。

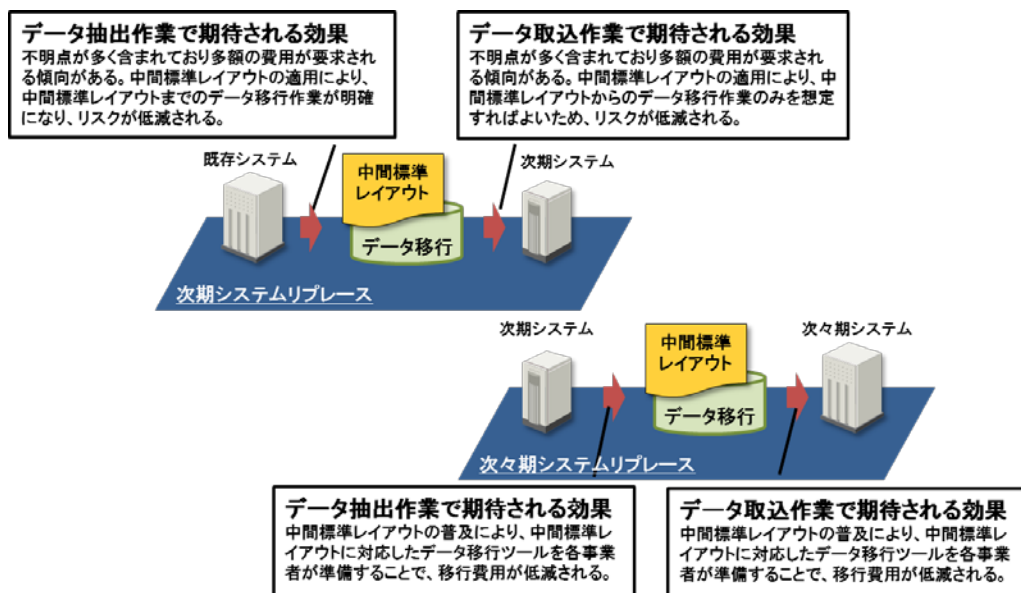


図1-10 中間標準レイアウトの活用によるデータ移行で期待される効果

中間標準レイアウトの適用による次期システムリプレース時、次々期システムリプレース時におけるコスト削減効果については、「中間標準レイアウトの有効性に関する調査研究報告書」において、以下のとおりデータ移行費用の削減効果が試算されている。

表 1-4 次期システムリプレース時におけるコスト削減効果の試算結果

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で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ース時)	
		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	<div style="border: 2px dashed red; padding: 5px; display: inline-block;">4.01人月</div>	

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

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ると考えられる。

表 1-5 次々期システムリプレース時におけるコスト削減効果の試算結果

	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ース時)	
		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ー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	

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

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。

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)

ここでの試算は、データ抽出作業と打ち合わせ調整作業、データ変換ツールの作成作業という移行に必要な全体作業工数を対象としており、次期システム調達時には約 35% 削減 (6.15 人月→4.01 人月)、次々期システム調達時には約 86%削減 (4.01 人月→0.57 人月) と試算されている。

なお、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。

4 データ移行費用の抑制効果

中間標準レイアウトの活用により、人口規模が同程度等の類似の自治体間で、複数ベンダのデータ移行費用を比較することが可能となる。

(1) 問題点・課題

異なるベンダ間における新旧システム間のデータ移行では、自治体業務システムの仕様が自治体ごとに異なるため、自治体間でデータ移行費用の比較を行うことが困難であった。このように、比較が困難である上に、その適切な金額が不明であったため、ベンダから提示されたデータ移行費用に対する交渉が困難であった事例も見受けられた。

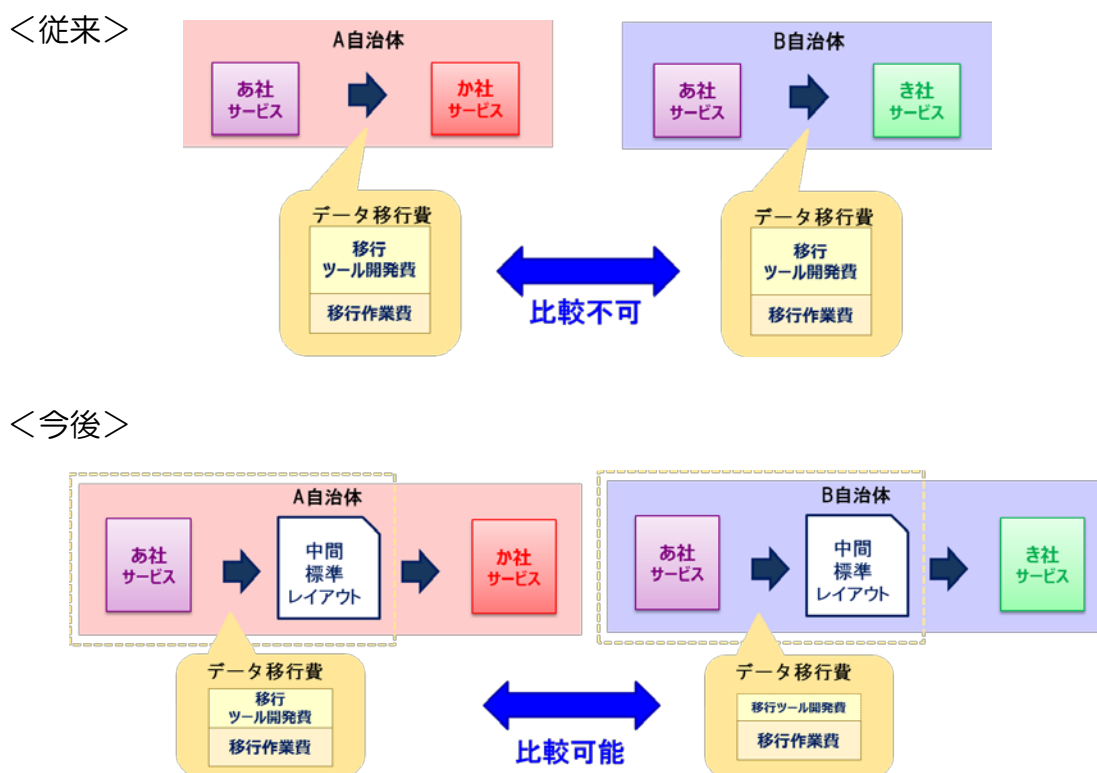
(2) 対策

人口規模が同程度等の類似の自治体において、中間標準レイアウトを適用した際のデータ移行費用は、同程度になることが想定される。

自治体は、先行して中間標準レイアウトを使用してデータ移行を実施した、類似の自治体の移行費用の実績額を参考とし、ベンダへ提示して交渉を行う。

(3) 得られる効果

自治体から類似条件下で行ったデータ移行の先行実績額を提示されることにより、ベンダはどの自治体向けのデータ移行費用の見積りにおいても、中間標準レイアウトの適用を前提とした、透明化された見積り提示が可能となる。これにより、データ移行費用の抑制効果が期待される。



(※) A自治体及びB自治体は人口規模が同程度であることを想定している。

図 1-11 データ移行費用抑制効果イメージ

5 中間標準レイアウトの具体的活用方法（例）

(1) 問題点・課題

従来は、システム調達に係る契約締結時に、契約期間満了後のデータの抽出及び引き渡しについては、契約内容に盛り込まないケースが多く見られた。この場合、契約期間満了後、新たにデータ抽出に係る契約を結ぶ必要が生じることとなる。

(2) 対策

自治体は、新たに業務システムを調達（次期調達）する際、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でデータ提供する旨を調達仕様書へ明記する。

(3) 得られる効果

これにより、次期システムの構築から数年後に次々期システムへ更新する際、仕様が明らかな中間標準レイアウトのデータ形式でデータを提供すればよく、このデータ抽出費は次期システムの調達費用に既に含まれているため、データ移行費用を抑えることができる。

また、次々期システム調達時において、移行データを中間標準レイアウトのデータ形式で入手することで、データ移行費用に極端な差異はなくなるものと考えられる。

このため、次々期システム調達時に既存システムと同一ベンダが優位となるベンダロックインを回避し、オープンな競争原理が働くことが想定される。

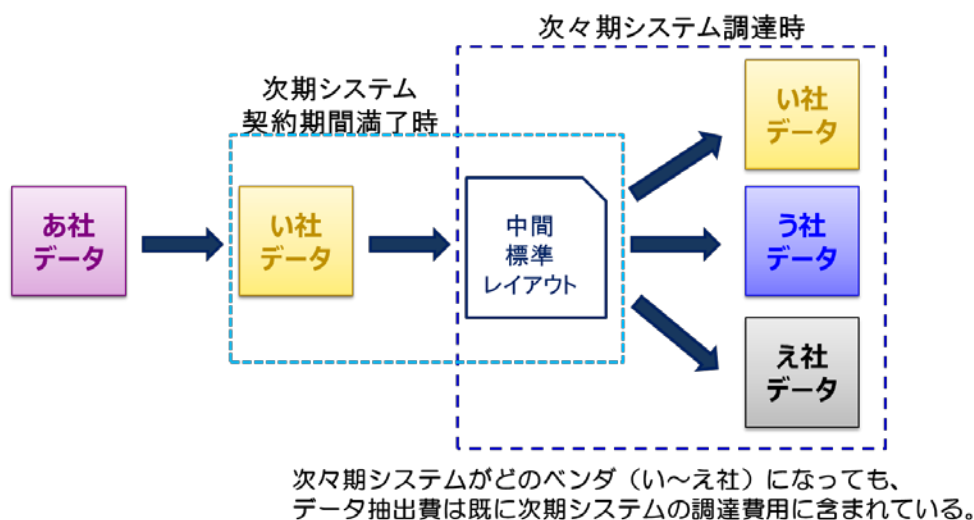


図 1-12 次期システムの契約期間満了後のデータ提供イメージ

6 データ移行以外への活用方法

(1) 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

① 問題点・課題

同一自治体内の様々な業務システム間でデータ交換を行うためには、データ交換用プログラムを開発することが一般的である。このデータ交換プログラムの開発は、自治体ごと、ベンダごとに仕様が異なるため、データ交換用プログラム開発ベンダにとって非常に効率の悪い作業となっている。

また、データ交換プログラムの開発には、相互の業務システム内のデータ形式やデータレイアウト、通信規約など、多様なインタフェース条件を開示して、整合させるためには、ベンダ同士での相互調整が必要となる。ベンダ間の連絡及び調整は各ベンダと契約している自治体の職員によってなされる場合が多い。

中間標準レイアウトがなければ、一方の業務システムのインタフェース条件に合わせて、もう一方の業務システムを改修することになるため、費用負担についても同様に自治体職員の仲介による相互調整が発生する。

② 対策

同一自治体内の様々なシステム間で必要となるデータ交換用プログラムの開発に、中間標準レイアウトを共通的な標準データ形式として適用する。

③ 得られる効果

ベンダは、業務システムに中間標準レイアウトに対応したデータ抽出機能を実装しておくことで、データ交換用プログラムを開発する必要がなくなる。

また、自治体は、職員が双方のベンダとの間の仕様調整を行う必要がなくなる。

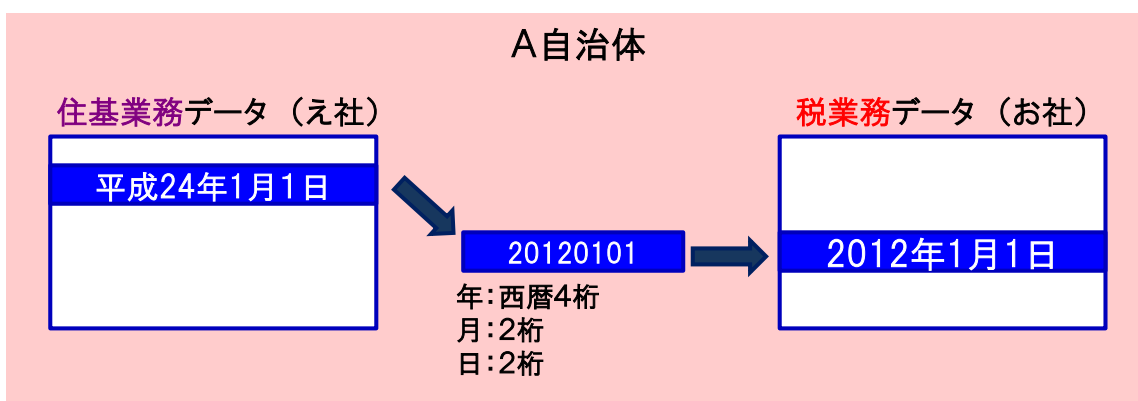


図 1-13 異なる業務システム間連携イメージ

(2) 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

① 問題点・課題

市町村合併等、他自治体とのデータの授受が必要な場合、他自治体の業務システム間でデータ交換用プログラムを開発することが一般的である。このデータ交換プログラムの開発は、自治体ごと、ベンダごとに仕様が異なるため、データ交換用プログラム開発ベンダにとって非常に効率の悪い作業となっている。

また、データ交換プログラムの開発には、相互の業務システム内のデータ形式やデータレイアウト、通信規約など、多様なインタフェース条件を開示して、整合させるためには、ベンダ同士での相互調整が必要となる。ベンダ間の連絡及び調整は各ベンダと契約している自治体の職員によってなされる場合が多い。

中間標準レイアウトがなければ、一方の業務システムのインタフェース条件に合わせて、もう一方の業務システムを改修することになるため、費用負担についても同様に自治体職員の仲介による相互調整が発生する。

② 対策

他自治体の業務システム間で必要となるデータ交換用プログラムの開発に、中間標準レイアウトを共通的な標準データ仕様として適用する。

③ 得られる効果

ベンダは、業務システムに中間標準レイアウトに対応したデータ抽出機能を実装しておくことで、データ交換用プログラムを開発する必要がなくなる。

また、各自治体は、各自治体職員が双方のベンダとの間の仕様調整を行う必要がなくなる。

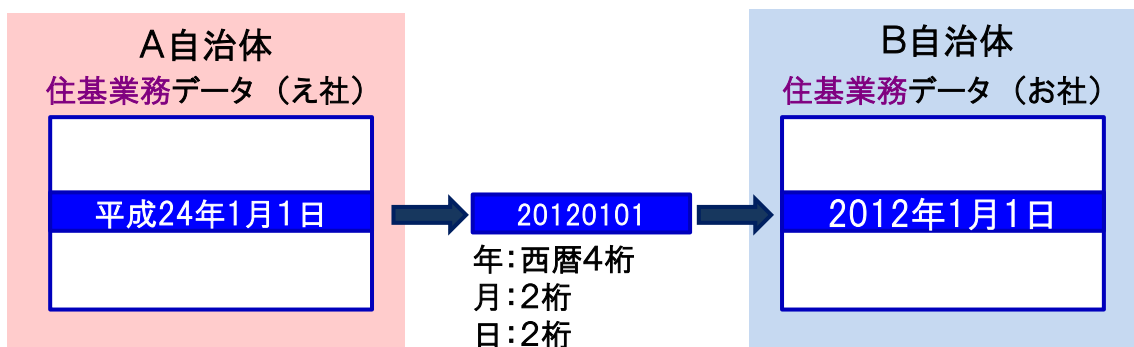


図 1-14 他自治体とのデータ交換イメージ

7 データ移行時の著作権

データ移行時に懸念される著作権に関する解釈について、以下に記述する。

○ データ移行時の著作権の取扱いについて明示

調達仕様書等に、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明記することにより、データ移行時の著作権の取扱いに関する無用の混乱を避けることができる。

① 課題

「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、ベンダが移行データのデータベースの著作権を保有していると主張し、当該著作権の移転に係る追加的な費用を請求されることが懸念されている。

しかし、移行データのデータベースの著作権については、以下の解釈が整理されている。

② 解釈

データベースの著作物として保護されるには、データベースであり、かつ、その情報の選択又は体系的な構成によって創作性を有することが要件とされる（著作権法第12条の2第1項）。しかしながら、自治体業務に必要な情報を扱う場合には、これに必要十分な限度で必然的に情報が選択されるため、体系的構成も類似せざるを得ず、そ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と考えられる。さらに、個々のデータそれ自体は、一般的に創作性は認められないものと考えられる。

（※平成23年6月自治体クラウド推進本部有識者懇談会とりまとめから引用）

【著作権法】

（データベースの著作物）

第12条の2 データベースでその情報の選択又は体系的な構成によつて創作性を有するものは、著作物として保護する。

2 （略）

③ 今後の対応

自治体は、この解釈を元にして、データ移行契約に際しての調達仕様書等に上記の旨を確認的に規定することが考えられる。

また、次期調達以降、システム開発とデータ移行（抽出）とを一体的に契約する際には（21頁参照）、次のような各項を調達仕様書等に規定することが、データ移行時の著作権の取扱いに関する事後的な混乱を回避する観点からも有効である。

【調達仕様書等の記載例】

- 1 本契約履行過程で生じた納入成果物（※）に関し、著作権法第27条及び28条に定める権利を含むすべての著作権は、△△（ベンダ）又は第三者がツール等として従前から著作権を有している場合を除き、○○（自治体名）に帰属するものとする。なお、この場合において△△は著作権者人格権を行使しないものとする。
- 2 ○○は、前項により△△に著作権が留保された著作物につき、納入成果物を利用するために必要な範囲で、複製、翻案又は改変することができるものとし、△△は、かかる利用について著作権者人格権を行使しないものとする。

（※） 契約期間満了時に提出される移行データのデータベースを含む。

（※※） 第三者の著作権の利用許諾などは別に規定が必要。

8 今後のフォローアップ

今後、自治体業務システムへのクラウドサービスの導入が本格化していく段階において、自治体業務システム間のデータ移行等における中間標準レイアウト仕様の利活用・普及の観点から、以下のフォローアップが重要と考えられる。

<制度改正等に伴う改定に係る維持管理>

自治体において住民サービスを提供する各種業務システムでは、各種の社会保障制度や税制改正等に大きな影響を受ける。中間標準レイアウト仕様については、今後想定されるこれらの制度改正の内容等をタイムリーに反映することがデータ移行における標準仕様としての利活用、普及に際して重要となる。

そのためには、中間標準レイアウト仕様の制度改正等に伴う改定が必要不可欠であり、受け皿となる活動や主体などについての検討が求められる。また、その際には関係する各種の標準化活動、団体等の連携なども視野に入れていくべきであると考えられる。

なお、中間標準レイアウト仕様の追加、変更時においては、自治体では遡及利用する場合や過去のログ内容を確認する場合もあるため、過去版の中間標準レイアウト仕様についても参照可能とする必要がある。

<実適用を通じての評価と知見の蓄積>

中間標準レイアウト仕様は、自治体の業務システム構築を実施している事業者の業務パッケージ等をベースに作成されている。当該レイアウト仕様の作成にあたっては、複数の事業者や団体の視点でのチェックを行っているが、このレイアウト仕様はあくまでも「ひとつの標準的な仕様」であり、実際の適用における課題などについての具体的な検証までは至っていない部分がある。

今後は、この中間標準レイアウト仕様をもとに、自治体における実際のデータ移行等への適用を行い、実運用における仕様の内容に関しての過不足の補完や、適用時における留意点等の知見・ノウハウなどを蓄積し、自治体での幅広い利活用に資するための取り組みを実施していくべきと考えられる。

第2章 外字の課題と調査結果の活用について

1 外字の実態とその課題

1.1 外字の主な課題

(1) 外字の実態

外字とは特定の文字集合（文字コードなど）に含まれない文字を指す。コンピュータが処理する文字については、各ベンダのパッケージソフトウェアにおいて、そのベンダによる標準の文字コードに即した標準文字の他に、自治体が独自に文字を登録し、任意のコードを割り当てる文字（外字）領域が準備されており、自治体ごとに多くの外字が設定されている。

通常、同じ業務システムにおいてもベンダが異なれば、標準文字（内字）の定義（文字セット）、外字領域（開始文字コード及び領域サイズ）が異なる。標準文字の定義はベンダにより、JIS2004、Shift_JIS、Unicode 等、多くの規格化された文字コードの中から選択して使われていることが多い。それに加え、ベンダ独自の外字セットをシステムの標準文字として含む場合もある。外字としてユーザが登録可能な文字領域も、ベンダにより様々である。

このように、文字の取扱いについては、ベンダごとに大きく異なっているのが現状である。

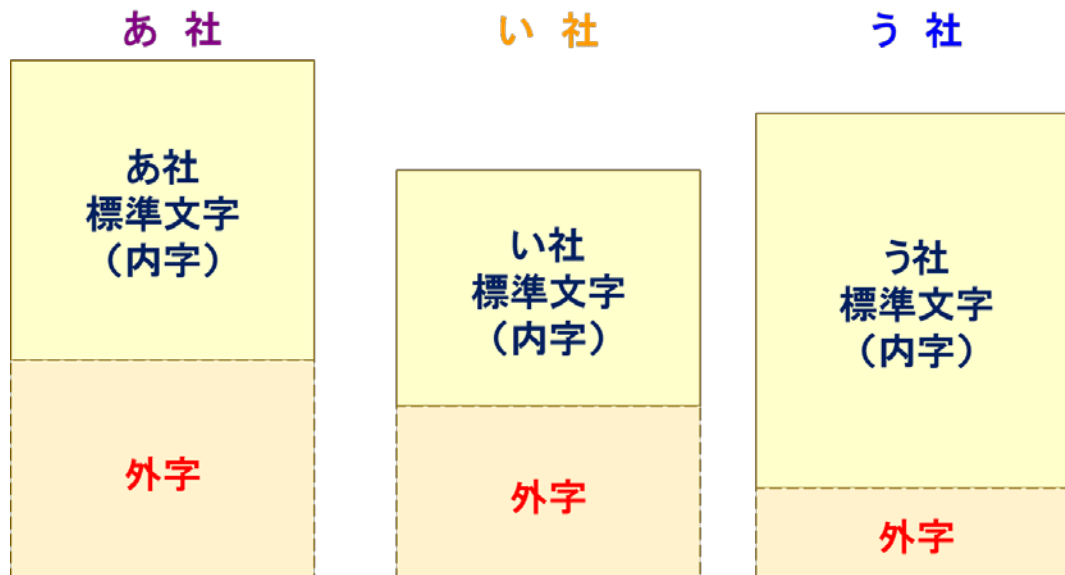


図 2-1 ベンダによる標準文字及び外字領域の違い

(2) 外字の新規登録の費用負担・作業負荷

外字の発生要因として、戸籍業務システムに登録されていない文字の申請、住民基本台帳ネットワークシステム統一文字（以下「住基ネット統一文字」という。）にない文字が記載された申請書が提出された場合などがあげられる。

このように外字の登録が必要となっていることで、外字を作成する職員の作業負荷

のみならず、外字作成ソフトの調達・保守コストが必要となる。

また、外字は画像として登録されていることから漢字で検索することができず、住民からの申請書に記載された文字が既に外字として登録済みであるかどうか調べるために、登録した文字字形とコードを一覧化するなどの作業が必要になる。なお、専用の検索ソフトを使用している場合は、保守作業も必要になる可能性がある。

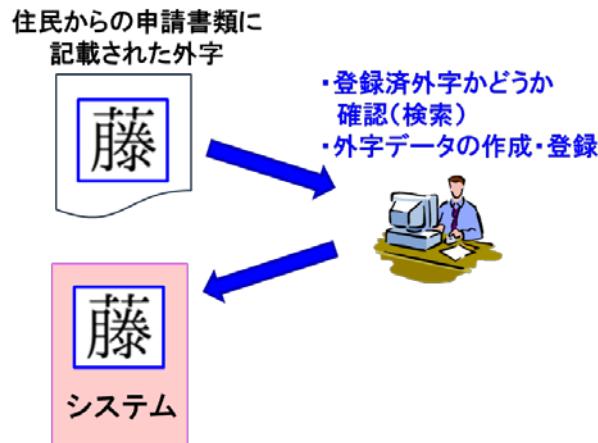


図 2-2 外字の新規登録作業イメージ

その他、業務システムの外字フォント登録領域のサイズによっては、登録できる外字の残数が逼迫することも懸念される。

(3) 外字データの授受が困難

同一自治体内での異なる業務システム間連携においても、連携する業務システムのベンダが異なっていると、連携元システムでは標準文字としている文字が、連携先システムでは外字になっている場合、同じ文字であっても表示イメージが異なり、つきあわせが困難な場合がある。

このようなケースで外字を含むデータを連携するためには、両ベンダの文字コード体系の変換テーブルを作成し、文字の同定作業を行った上で、それらを継続的にメンテナンスしなければならず、作業負担や保守コストが必要となる。

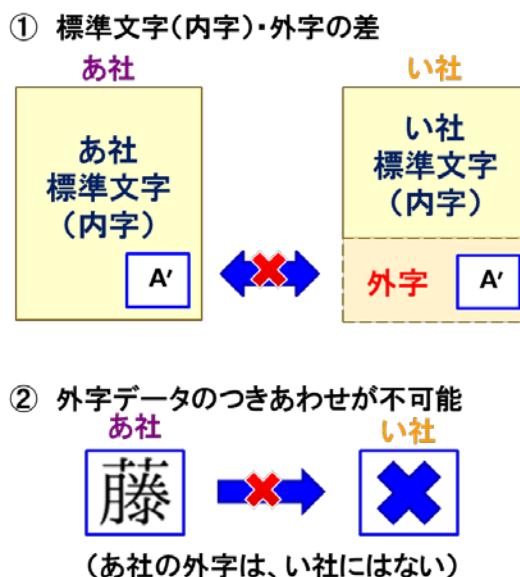


図 2-3 外字データの授受に関する問題

(4) 外字データの印刷に関わる問題

基幹業務等で使用されているプリンタ・大量印刷機では、内部に登録されているフォントデータを使う印刷形態をとっているものがあるため、その場合は登録されていない文字を印刷することはできない。

また、印刷を行う業務システムが、その提供ベンダ独自の文字セットを使用している場合、同じ文字セットに登録したプリンタ・印刷機でないと印刷できない文字が発生する可能性があり、結果的に業務システムの提供ベンダ以外のベンダのプリンタ・印刷機を使用することができず、ベンダロックインの原因となっている。

自治体では、住民へ発送する書面や封筒の印刷など、大量の印刷を外部委託する場合があるが、印刷する内容に外字が使われている場合、印刷するデータの他に外字データを委託先の事業者提供しなくてはならないなど、余計なデータのやりとりが発生していることもある。

(5) 外字を含むデータの業務利用に関わる問題

専用端末ではなく通常業務用端末から外字の検索を行う場合、日本語入力ソフトで変換候補となるのは、端末のオペレーティングシステムで扱える標準文字のみであるため、漢字による検索ができない場合がある。従って、漏れのない検索を行うためには、カナ文字による検索をしなくてはならないというケースがある。

また、業務システムで取り扱うデータに外字が含まれている場合、外字に対応したソフトウェアでしか当該外字の表示ができないという問題もある。例えば、ある業務システムで外字を含むデータを出力し、汎用的な表計算ソフトウェア等に読み込むと、外字の部分が文字化けするため、データの二次利用を困難にしている場合がある。

1.2 システム移行時の課題と現状の対処

(1) 文字データの移行が困難

業務システムを移行する際に、新旧システムで提供ベンダが異なる場合は、外字の同定作業やコードの割当て作業が発生することがある。さらに、旧システムで標準文字となっており、新システムで標準文字に含まれていない文字については、新システムで必要な文字であるかどうかを調査（文字の見える化）した上で、必要な文字については紙で提供された情報から外字データを新規作成・登録する必要がある。

また、新旧システムにおいて文字フォントの管理方法も異なる場合がある。その際はデータ移行が不可能であり、外字データを再作成しなくてはならず、移行費用が膨大なものとなる可能性があるが、この場合も、必要な外字かどうかを調査（文字の見える化）した上で、作成・登録する必要がある。

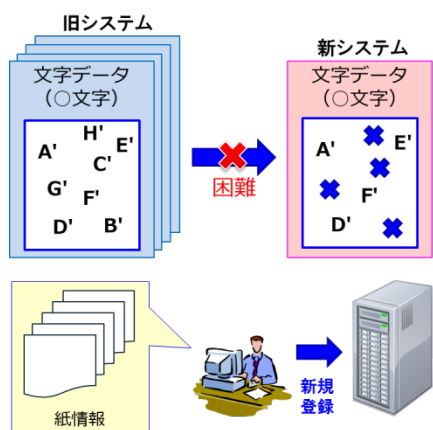


図 2-4 文字データ移行イメージ

(2) 基幹業務用プリンタ・印刷機の外字対応困難

旧システムから新システムへの移行後、新システムと既存のプリンタ・印刷機の提供ベンダが異なる場合、プリンタ・印刷機に登録されているフォントデータと新システムのフォントデータではコード体系が異なる可能性があり、外字が正しく印刷されない場合がある。

また、プリンタ・印刷機が、導入した新システムにおける文字フォントの管理方式と互換性がない時は、既存のプリンタ・印刷機を継続して使用することは困難となる可能性がある。

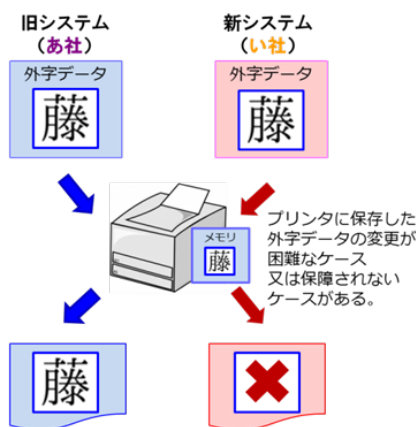


図 2-5 印刷機のシステム互換性の問題

既存ベンダのシステムから新しいベンダのシステムに切り替える際において、これらの問題に対処するためには、以下のような方法が考えられる。

① プリンタ・印刷機をすべて新しくする

新システムの導入に合わせて、プリンタ・印刷機を導入した新システムに対応するものにすべて置き換える方法である。ただし、この方法は置き換えるプリンタの数や価格によっては、旧システムから新システムへの移行の際のコストを増加させる要因となり得る。

② 外字データをすべて再登録する

導入した新システムに合わせて、すべての外字データを再登録する方法である。ただし、この方法は多大な外字登録作業が旧システムから新システムへの移行時にかかることとなり、新システム導入の際のコストを増加させる要因となり得る。

また、プリンタ・印刷機の文字フォント登録可能数が不足する場合や、移行した新システムにおける文字フォントの管理方式と互換性がない時は、既存のプリンタ・印刷機を継続して使用することは困難となる可能性がある。

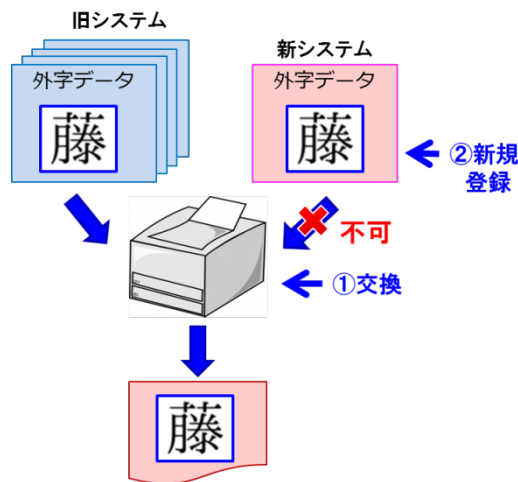


図 2-6 印刷に関する新システム導入の方法

2 外字実態調査の報告

2.1 調査対象の外字情報

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 万字余）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

2.2 外字の同定

市区町村から収集した外字の同定作業は、以下の手順にて作業を進めた。

作業ステップ	概要	外字調査 請負事業者	協力 自治体	研究会
STEP1 外字情報の 収集	自治体で実際に使用されている外字情報（文字データ）を収集する。		外字 情報	
STEP2 包摂基準 (案) の作成	同定作業の具体的なフローとなる包摂基準（案）を作成する。	包摂 基準 (案)		
STEP3 同定作業 の実施	包摂基準（案）をもとに、収集した外字情報と文字情報基盤漢字との同定作業を実施し、「市区町村が使用する外字の実態調査報告書（案）」を作成する。	市区町村 が使用する 外字の 実態調査 報告書 (案)		
STEP4 有識者 からの 意見収集	「市区町村が使用する外字の実態調査報告書（案）」について、本研究会にて説明し、有識者から意見を収集する。			確認
STEP5 報告書 の作成	上記 STEP1 から STEP4 までの成果をとりまとめ、「市区町村が使用する外字の実態調査報告書」を作成する。	市区町村 が使用する 外字の 実態調査 報告書		

図 2-7 同定作業の手順

2.3 文字情報基盤漢字の概要

市区町村が使用する外字の実態調査（以下「本調査」という。）における外字の同定には、文字情報基盤漢字（平成 23 年 10 月 26 日公開正式版フォント情報、<http://ossipedia.ipa.go.jp/ipamjfont/>）を活用した。

この文字情報基盤漢字は、「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決定）における文字に関する課題への取り組みの一環として、平成 22 年度に内閣官房（情報通信技術（IT）担当室）、総務省、法務省、経済産業省、文化庁などの関係府省や関係者が参加する文字情報基盤推進委員会を設置し、文字基盤の在り方について検討を重ね作成されたもので、戸籍統一文字や住基ネット統一文字を網羅した 58,712 文字のフォントと画数等の文字情報が整備されている。

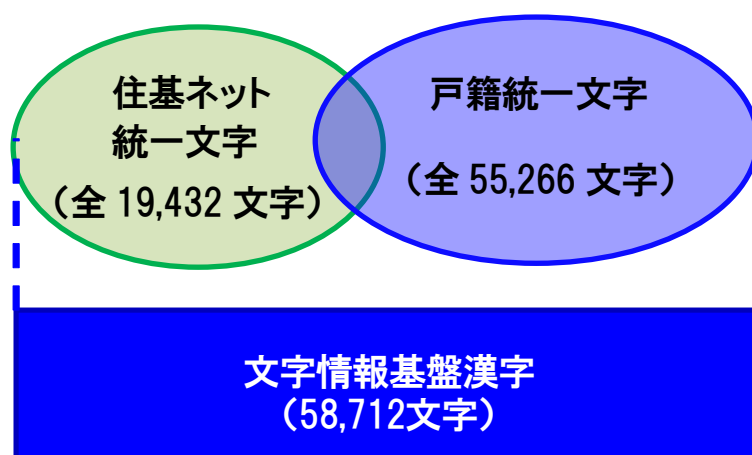


図 2-8 文字情報基盤漢字の概要

2.4 包摂基準の作成

本調査における同定作業による文字の分類

(1) 「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。

(2) 「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。

◆デザイン差の適用例



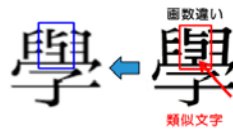
※平成20年度経済産業省委託事業
「汎用電子情報交換環境整備プログラム成果報告書」より一部引用

図 2-9 デザイン差の適用例

(3) 「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。

◆類似文字の適用例



※平成20年度経済産業省委託事業
「汎用電子情報交換環境整備プログラム成果報告書」より一部引用

図 2-10 類似文字の適用例

(1) 文字構成要素の配置の例

文字構成要素の画数・形状が同じであるが、配置が異なる文字は、異なる字形の文字とした。

例) 「松」と「叅」は、文字構成要素の画数・形状は同じであるが、配置が異なるので同一字形ではない。





(2) 文字構成要素の画数・形状の例

文字構成要素の画数・形状による、デザイン差と類似文字の判断例

例) 「松」は、木+八+厶から成り立っている。「木」、「八」、「厶」を文字構成要素と定義し、それぞれの要素の形状と画数を包摂基準に用いた。



例えば「八」が、の場合は、「八屋根」のデザイン差の基準の範囲内と判断した。

「厶」が、の場合は、画数違いにより類似文字と判断した。

※平成20年度経済産業省委託事業「汎用電子情報交換環境整備プログラム成果報告書」より一部引用

図 2-11 文字構成要素の配置・画数・形状

(4) 「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、変体仮名及び記号など。

2.5 同定作業

(1) 同定作業の手順

包摂基準に基づき、文字同定を行った手順を以下に示す。

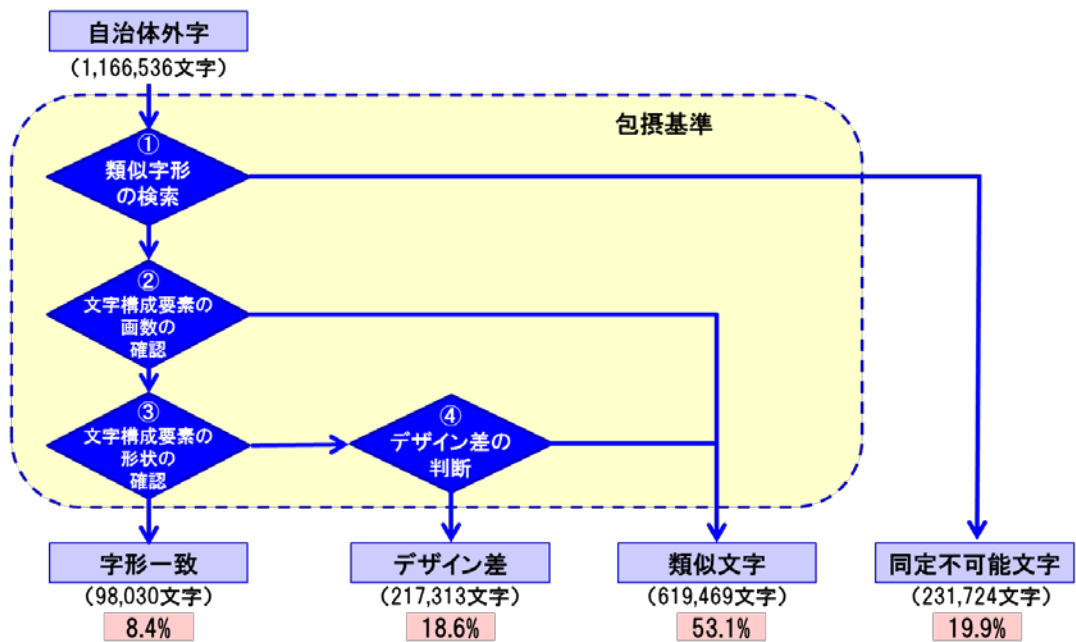


図 2-12 同定作業のフロー

<作業手順>

- ① 「類似字形の検索」
 - すべての文字構成要素の配置が一致する／しない
 - 類似字形の漢字が文字情報基盤漢字に存在する／しない
- ② 「文字構成要素の画数の確認」
 - すべての文字構成要素の画数が一致する／しない
- ③ 「文字構成要素の形状の確認」
 - すべての文字構成要素の形状が一致する／しない
- ④ 「デザイン差の判断」
 - 文字構成要素の形状の違いが、デザイン差の基準の範囲内に収まる／収まらない

(2) 同定作業の結果

自治体外字（1,166,536 文字）の約 80%（934,812 文字）が、文字情報基盤漢字に「字形一致」、「デザイン差」、「類似文字」という分類で関連づけることができた。「同定不可能文字」（231,724 文字）は、不明文字（52,294 文字）、変体仮名（97,791 文字）、記号等（81,639 文字）として整理した。

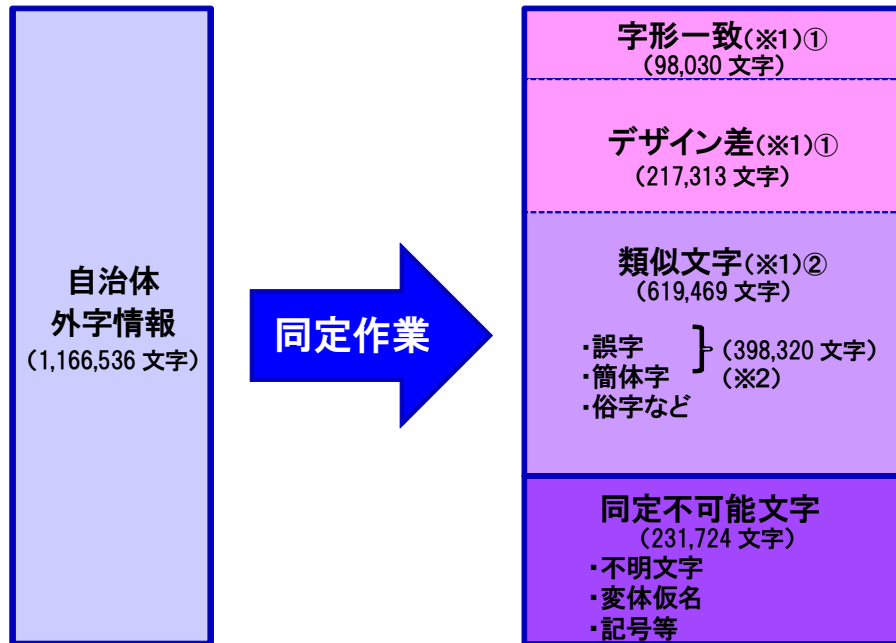


図 2-13 同定結果の概要

(※1) 分類した文字の種別について

① 「字形一致」・「デザイン差」

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した文字

② 「類似文字」

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な文字

(※2) 「類似文字」の分類について

① 誤字

戸籍では実際に利用されていない可能性の高い文字

② 簡体字

住民基本台帳法の一部を改正する法律（平成 21 年法律第 77）附則第 3 条により多くが置き換えられ、実際に住民票では利用されなくなる可能性の高い文字

③ 俗字など

現に戸籍で利用されている可能性が高い文字
(戸籍電算化時にもそのまま利用できる文字)

※ 上記※2①及び②（合計 398,320 文字）については、戸籍電算化の進展と電算化した戸籍の氏名に合わせた文字への住民票氏名の更正及び保存期間経過やシステム更新による住民票の改製により、その多くが解消されるであろうと考えられる。

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 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 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 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 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 89,280 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

図2-14 外字の実態調査結果

3 今回の同定結果の活用

3.1 外字発生抑制

(1) 文字運用ポリシーの策定

自治体では、使用する外字の取り扱いに関する文字運用ポリシーを策定することが必要である。具体的には、本調査の成果物である「包摂基準書」や「外字実態調査の活用方法」を参考とし、文字の確認基準や外字作成の基準などを文字運用ポリシーに盛り込んで作成する。この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、文字運用負荷の軽減につながると思われる。

また、自治体においては、人事給与システムなど住民用の帳票を発行しない内部事務においては、外字を使用しないなどの方針を文字運用ポリシーへ盛り込むことも外字発生抑制につながると思われる。

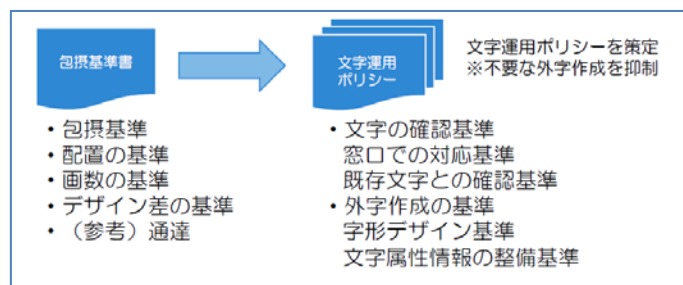


図 2-15 文字運用ポリシーの策定イメージ

(2) 窓口業務での説明への活用

IPA MJ 文字情報検索システムでは、文字情報基盤漢字について、文字属性情報が設定されており、簡単に類似文字が検索できるばかりでなく、戸籍統一文字や住基ネット統一文字であることもわかるので、窓口での住民への説明に活用できる。

また、市区町村に文字字形を伝達する際、文字情報基盤漢字の MJ 文字図形名を使うことで、正確に伝達することができる。

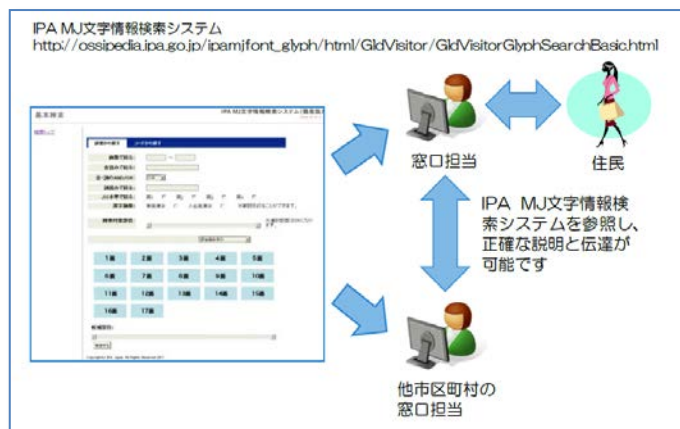


図 2-16 窓口業務での説明への活用イメージ

3.2 外字データの移行における活用

システム移行時における外字データの受け渡し方法は大きく2種類ある。一つは電子データで受け渡しをする方法で、もう一つは紙媒体で受け渡しをする方法である。外字データを異なるベンダのシステム間で移行する場合、電子化された外字データは紙媒体に印刷され、移行先システムのベンダは、紙情報から外字データを新たに作成、登録するという膨大な作業を行っている事例がある。

しかし、本調査の同定結果を活用することにより、文字情報基盤漢字に関連付けられた文字については、独立行政法人情報処理推進機構（IPA）が無償で提供するフォントデータを利用できるため、外字の新規登録作業が大幅に軽減されることが期待できる。（自治体として独自対応が必要とされる外字は、主に同定不可能文字に分類された文字となるため、外字データの作成、登録の業務が大幅に削減されることが予想される。）

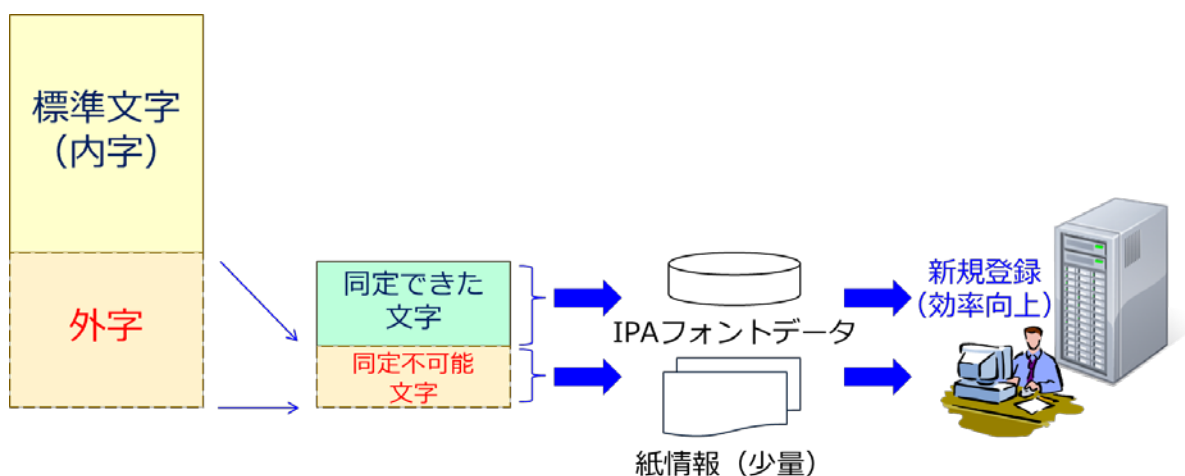


図 2-17 外字データ量の縮小によるシステム移行に関する効果

3.3 外字の整理に活用

外字情報を提出された自治体に対して、本調査で実施した外字同定作業の当該自治体分の同定結果リスト及び同定結果 CSV ファイルを提供した。

各自治体では今回の同定結果を参考にして、逐次外字の整理を行っていくことで、外字の維持対象を大幅に減らすことも可能となり、職員の負担軽減が期待される。

(1) 住基ネット統一文字に一致した外字（約 24 万文字）

住基ネット統一文字に一致した外字（字形一致及びデザイン差）は削除し、本来の住基ネット統一文字コードに変更し整理することができる。整理する際には、データの変更を伴うため十分な確認が必要だが、これらの外字をすべて解消できると、全国で約 24 万文字の外字解消につながる。

3861文字)					
外字	MJ文字図形名	補足	外字	MJ文字図形名	補足
E001	MJ021897	住基:J+8352 戸籍:045436 UCS:U+8352 UCS実装:○	E00B	MJ023148	住基:J+8A9F 戸籍:060476 UCS:U+8606 UCS実装:
荒	荒	類似文字	蘆	蘆	字形一致
E002	MJ029104	デザイン差 住基:J+9B8E 戸籍:321116 UCS:U+9B8E UCS実装:○	E00C	MJ059309	デザイン差 住基:J+AD41 戸籍: UCS: UCS実装:
鮎	鮎	デザイン差	今	今	デザイン差
E003	MJ002878	字形一致 住基: 戸籍:263700 UCS:U+405C UCS実装:○	E00D	MJ006569	デザイン差 住基:J+4F0A 戸籍:006260 UCS:U+4F0A UCS実装:○
睨	睨	字形一致	伊	伊	デザイン差
E004	MJ016776	デザイン差 住基:J+B572 戸籍:223116 UCS:U+7247 UCS実装:	E00E	MJ022690	字形一致 住基:359520 戸籍:U+851A UCS実装:
片	片	デザイン差	蔚	蔚	字形一致
E005	MJ008581	デザイン差 住基:J+5609 戸籍:047020 UCS:U+5609 UCS実装:○	E00F	MJ015127	字形一致 住基:J+6C84 戸籍:104260 UCS:U+6C84 UCS実装:○
嘉	嘉	デザイン差	汙	汙	字形一致

図 2-18 住基ネット統一文字と一致した外字の例

(2) 戸籍統一文字に一致した外字（約 7 万文字）

戸籍統一文字に一致した外字（字形一致及びデザイン差の約 7 万文字）は、現に戸籍で利用されている文字が住基ネット統一文字に存在しないため、外字を作成し利用している文字と推測される。外字の整理には十分に戸籍を確認し注意して行う必要がある。

(3) 類似文字中の誤字及び簡体字（約 40 万文字）

類似文字中の誤字については、既に戸籍電算化などで正しい文字に置き換えられている可能性がある。当該文字が現に使われている文字かどうかを確認した上で、住民票等のデータを修正することで、外字解消につながる。

類似文字中の簡体字については、住民基本台帳法の一部を改正する法律（平成 21 年法律第 77 号）附則第 3 条によりその多くが正字等に置き換えられる可能性がある。仮住民票等を確認して文字の確認をしたのちに、データを修正することで、外字解消につながる。

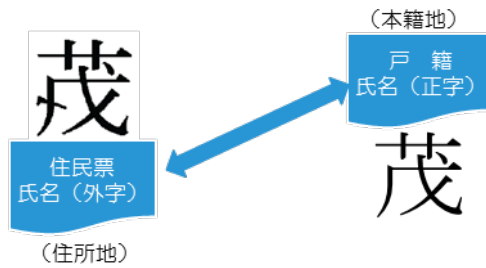


図 2-19 類似文字中の誤字確認



図 2-20 類似文字中の簡体字確認

3.4 文字環境の見える化に活用

同定結果リスト及び CSV ファイルを用いて、市区町村内の各種業務システムや市区町村外とのデータ連携に係る外字の見える化に活用できる。

(1) 提供外字と文字情報基盤漢字の見える化

同定結果リストでは、提供外字と文字情報基盤漢字の文字字形を印字している。同定結果リストで実現している、整理の方法や、見える化の在り方について、市区町村内の文字環境の見える化を行う際の参考として活用することができる。

3861文字)					
外字	MJ文字図形名	補足	外字	MJ文字図形名	補足
E001	MJ021897	類似文字	E00B	MJ023148	字形一致
荒	荒	住基: J+8352 戸籍: 345430 UCS: U+8352 UCS実装: ○	蘆	蘆	住基: J+BA9F 戸籍: 368470 UCS: U+8606 UCS実装:
E002	MJ029104	デザイン差	E00C	MJ059309	デザイン差
鮎	鮎	住基: J+9B8E 戸籍: 521110 UCS: U+9B8E UCS実装: ○	今	今	住基: J+AD41 戸籍: UCS: UCS実装:
E003	MJ002878	字形一致	E00D	MJ006569	デザイン差
睨	睨	住基: 戸籍: 263700 UCS: U+405C UCS実装: ○	伊	伊	住基: J+4F0A 戸籍: 005250 UCS: U+4F0A UCS実装: ○
E004	MJ016776	デザイン差	E00E	MJ022690	字形一致
片	片	住基: J+B572 戸籍: 223110 UCS: U+7247 UCS実装:	蔚	蔚	住基: 戸籍: 359520 UCS: U+851A UCS実装:
E005	MJ008581	デザイン差	E00F	MJ015127	字形一致
嘉	嘉	住基: J+5609 戸籍: 047920 UCS: U+5609 UCS実装: ○	汙	汙	住基: J+6C84 戸籍: 194260 UCS: U+6C84 UCS実装: ○

図 2-21 同定結果リスト

(2) 各種業務システムにおける文字環境の見える化

同定結果 CSV ファイルと「既存住基システム文字」と「住基ネット文字」の文字対照表を連結し、さらに「既存住基システム文字」と「各種業務システム文字」との文字対照表などを連結し組み合わせることで、各種業務システムの文字と文字情報基盤漢字の対照表を作成することができる。

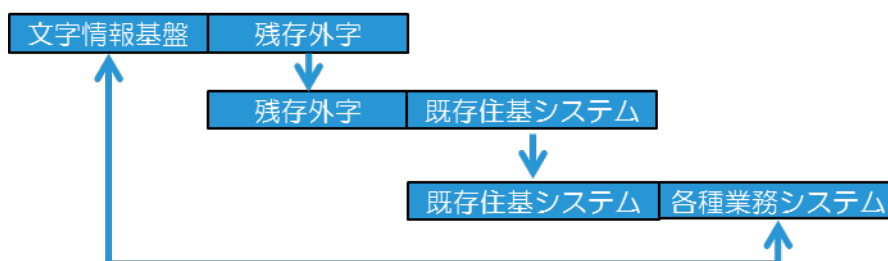


図 2-22 各業務システム文字対照表の作成イメージ

(3) 低解像度外字の再作成時の字形参照

文字情報基盤漢字と関連付けられた外字には、MJ 文字図形名が付されているので、低解像度外字の再作成時に文字情報基盤漢字のデザインを参照することができる。証明書などに印字する文字のデザインが簡単にかつ綺麗にできるため、住民との文字トラブルの解消も期待できる。

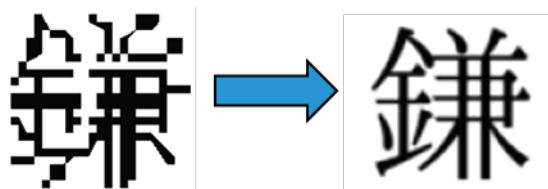


図 2-23 低解像度外字の再作成への活用

3.5 システム間連携やシステム移行時の標準文字に活用

同定結果リスト及び CSV ファイルには、文字情報基盤漢字の MJ 文字図形名及び UCS を掲載しているため、これらの情報を用いて各種業務システムにおける文字環境の見える化を行うことができる。各種業務システムの文字を文字情報基盤漢字に関連付けすることにより、文字情報基盤漢字をシステム間連携やシステム移行の標準文字に活用できる。

(1) システム間の連携用データの標準文字として活用

事前に文字情報基盤漢字と各種業務システムの文字字形を可視化し、各種業務システムの文字コードと共に見える化を実施し、文字情報基盤漢字と各種業務システム文字コードを関連付けておく。

このような準備を行うことにより、システム間連携の標準文字コードに文字情報基盤漢字コードを利用することができるようになり、システム間連携構築の際のベンダロックインを解消することにつながる。

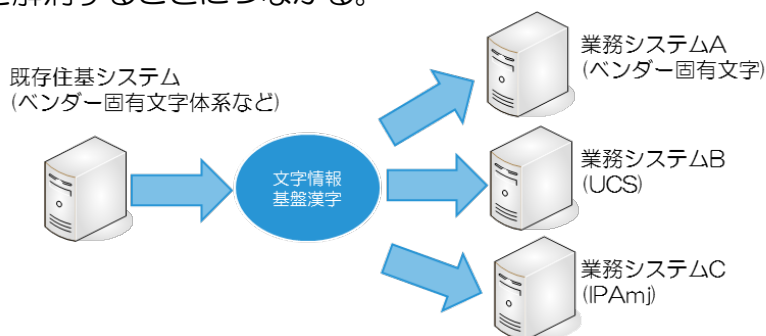


図2-24 システム間連携の標準文字への活用イメージ

(2) 新システムへの移行時の標準文字として活用

事前に文字情報基盤漢字と各種業務システムの文字字形を可視化し、各種業務システムの文字コードと共に見える化を実施し、文字情報基盤漢字と各種業務システム文字コードを関連付けておく。

このような準備と合わせて、別途総務省がまとめたデータ移行用の中間標準レイアウトを活用することで、システム移行時に固有文字コードや固有移行レイアウト等によるベンダロックインを解消することにつながる。

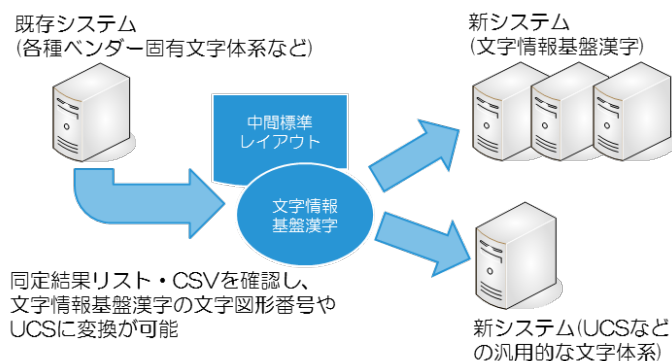


図2-25 システム移行時の標準文字としての活用イメージ

システムの新たな調達時には、中間標準レイアウトや文字情報基盤漢字に対応することを前提条件として調達仕様に入れることができれば、さらに効果が期待できる。

これらが継続されることによって、新システムへの円滑なデータ移行が可能となる。

3.6 プリンタ・大量印刷機における活用

今回の同定結果を活用し、外字を大幅に削減できれば、既存のプリンタ・大量印刷機が内部に登録されているフォントデータを使う印刷形態をとっている場合、残存外字データだけを新規登録することにより、プリンタ・大量印刷機を交換しなくても正しく印字できるようになるケースの増加が見込まれる。

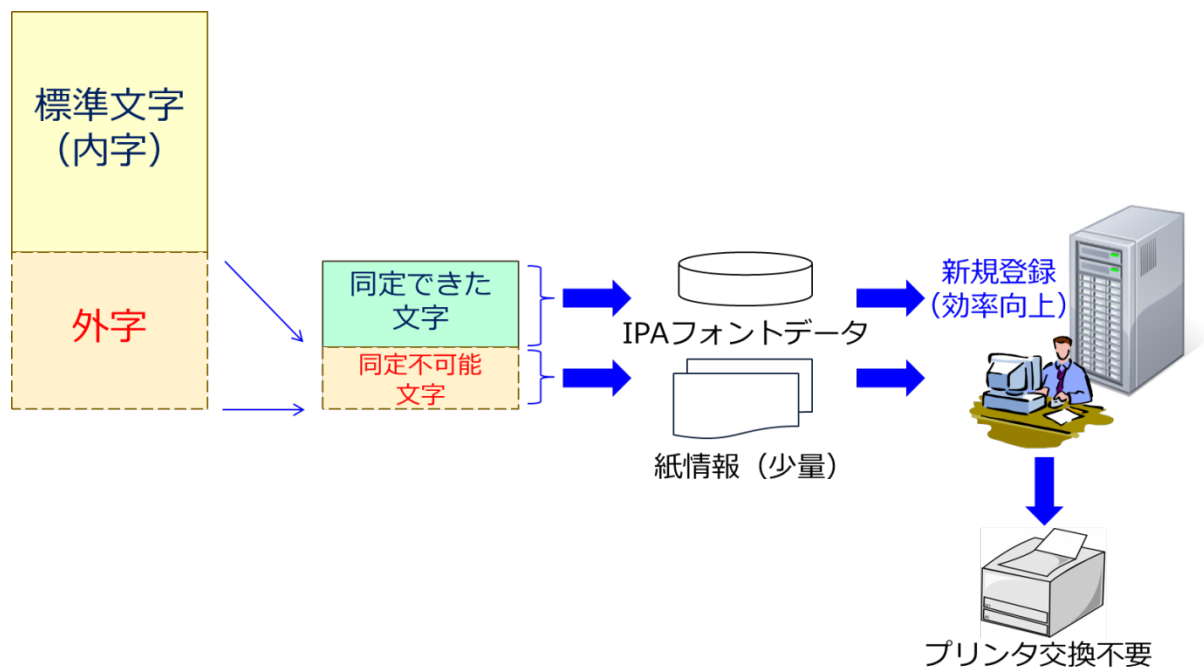


図 2-26 外字データ量の縮小による印刷に関する効果

4 今後のフォローアップ

4.1 常用漢字が変更になった場合の対応

文部科学省による常用漢字表の目的は、漢字使用の目安であり、教育や出版においては大きな意味を持つが、情報流通の観点とは分けて考える必要がある。今後、常用漢字が変更になった場合については、適切な時期に文字情報基盤漢字への反映が検討されることが望ましい。

4.2 その他変更対応

法務省による人名用漢字及び戸籍統一文字の変更等についても、常用漢字と同様に、適切な時期に文字情報基盤漢字への反映について検討されることが望ましい。

また、現に戸籍や住民票で用いられている新字体（※）の漢字についても文字情報基盤漢字への追加対応の検討が望まれる。

（※）新字体及び旧字体とは

「国」（新字体）、「國」（旧字体）のように同じ字種であるが字体が異なる文字をさししており、昭和 24 年に「当用漢字字体表」が制定されたことにより、字体の新旧が交代し、このように表現されるようになった。

「しめすへん」、「くさかんむり」などを構成要素にもつ新字体の中には、現に戸籍や住民票の氏名として利用されている文字が存在するが、文字情報基盤漢字には搭載されていない文字もある。

用語集

No.	用語	語義
1	XML	<p>文書やデータの意味や構造を記述するためのマークアップ言語の一つ。マークアップ言語とは、「タグ」と呼ばれる特定の文字列で地の文に情報の意味や構造、装飾などを埋め込んでいく言語のこと。</p> <p>ソフトウェア間の通信・情報交換に用いるデータ形式や、様々な種類のデータを保存するためのファイルフォーマットなどの定義に使われている。</p>
2	地域情報プラットフォーム	<p>自治体を持つ情報システムをはじめとした、地域内外の様々な情報システムを全国規模で連携させるための共通基盤。総務省主催の「地域における情報化の推進に関する検討会」の中で提言され、「一般財団法人 全国地域情報化推進協会」で標準仕様の作成・管理が行われている。Web サービスやXMLなどの技術を活用して情報システムの基盤を共通化することで、異なる情報システム間でのシームレスなデータのやり取りを実現し、行政・民間を問わず地域の様々なサービスを連携・統合して提供することを目的としている。</p>
3	ベンダロックイン	<p>ある業者（ベンダ）が自社の独自仕様などを搭載した製品及びサービスを提供することなどによって、他の業者の製品及びサービスに移行しにくくすること。本調査では他ベンダのシステムへのデータ移行経費が高上がりになってしまうことなどをベンダロックインの典型的事例として取り上げている。</p>
4	JIS2004	<p>JIS2004（JIS X 0213:2004）は、JIS X 0208:1997（いわゆるJIS97）を拡張し、JIS 第三水準文字・JIS 第四水準文字を含む4344文字を追加した文字コードのこと。JIS97と合わせて11,233文字が規定されている。</p>
5	Shift_JIS	<p>Microsoft社によって策定された日本語文字コードの一つ。文字の1バイト目を見るだけで漢字か1バイト文字（いわゆる半角英数字）が分かる、等幅フォントで表示した場合に画面上の桁数とバイト数が一致するなどの特長から、同社のMS-DOSやWindows、Apple社のMac OSなど広く普及している。</p>
6	Unicode	<p>Apple社、IBM社、Microsoft社など米国の情報関連企業が中心となって提唱し、1993年に国際標準化機構（ISO）でISO/IEC 10646の一部(UCS-2)として標準化された文字コード体系。すべての文字を16ビット(2バイト)で表現し、1つの文字コード体系で多国語処理を可能にしようとするもの。世界の主要な言語のほとんどの文字を収録している。</p>

自治体クラウドの円滑なデータ移行等に関する研究会構成員名簿

(平成24年6月現在 敬称略・五十音順)

【委員】

植村 恭則	広島県福山市企画総務局企画部情報政策課長
小出 弘美	千葉県総務部情報システム課電子自治体推進室副主幹
須藤 修	東京大学大学院情報学環長（座長）
橋本 幸治	神奈川県町村情報システム共同事業組合情報システム担当課長
原田 親一	東京都八王子市総合政策部次長兼政策審議室主幹（前 IT 推進室主幹）
村田 吉則	滋賀県東近江市総務部財政課長（前 情報推進課長）
山澤 浩幸	新潟県三条市総務部情報政策課長

【オブザーバ】

伊駒 政弘	財団法人地方自治情報センター研究開発部主席研究員
田代 秀一	独立行政法人情報処理推進機構技術本部国際標準推進センター センター長
武藤 俊一	財団法人全国地域情報化推進協会企画部担当部長