

異なるデータIDの統合・統一化に関する調査研究 概要版

2023年3月

株式会社アスコエパートナーズ



既存システムの実体調査

調査対象のシステムとして下記 6 システムを選択し、それぞれについてID・データの統合に向け、次の視点から調査を実施した。

- ID・データ --- ID 体系、登録・管理データ
- 運用・オペレーション --- 情報登録・更新のフロー、担当者（組織、部署）や更新の頻度、問い合わせ・サポートフロー

マイ制度ナビ

ID・データ

- 統合するデータの候補には、「制度情報」、「事例情報」がある。
- 管理者情報を含むためデータの統合に際しては継続性のリスクがある。

運用・オペレーション

- 所管組織の担当者が管理画面にてデータの登録・編集を行う。

ぴったりサービス

ID・データ

- 統合するデータの候補には、「申請情報」、「手続き情報」がある。
- 管理者情報を含むためデータの統合に際しては継続性のリスクがある。

運用・オペレーション

- 所管組織の担当者が管理画面にてデータの登録・編集を行う。

jGrants

ID・データ

- 統合するデータの候補には、「制度情報」、「申請情報」がある。
- IDに関して外部からの参照は想定されていない。

運用・オペレーション

- 所管組織の担当者が管理画面にてデータの登録・編集を行う。

e-MAFF

ID・データ

- 統合するデータの候補には、「制度情報」、「申請情報」がある。
- 複合IDが使用されている部分あり統合に際しては注意が必要である。

運用・オペレーション

- 所管組織の担当者が管理画面にてデータの登録・編集を行う。

e-Laws

ID・データ

- 統合するデータの候補には、法令ドキュメントがある。
- 固定長のID体系であり、仕様ビットパターンが限界に近く継続性のリスクがある。

運用・オペレーション

- デジタル庁の担当者が管理画面にてデータの登録・編集を行う。

ミラサポplus制度ナビ

ID・データ

- 統合するデータの候補には、「制度情報」、「事例情報」がある。マイ制度ナビとの重複が多い。
- 所管組織情報を含むためデータの統合に際しては継続性のリスクがある

運用・オペレーション

- 所管組織の担当者が管理画面にてデータの登録・編集を行う。

※ 調査結果の詳細については、別紙「各システムの登録データ比較表」「各システムの運用方法の比較表」に記載。

実態調査のまとめ

ID体系

- 既存のIDはシステム内における唯一性のみが考慮されている。
- ID内に所管組織、管理組織などの情報を含んでいるものが多い。
(課題) 上記の調査結果から、データ統合を行うためには、システムの枠を越えて唯一性を保証するID体系を導入する必要がある。
(リスク) 所管組織の情報を含んでいる、統合にあたってIDの適用範囲が拡張されることなどから継続性のリスクがあるID体系のものが多く見受けられた。

データモデル

- 制度、申請、手続きなど大きくは同様なデータを扱うシステムであるが、調査の結果、データの所管、項目など実際に重複するデータは多くなかった
- システムの対象が異なるがデータモデルとしては統一が可能なもの、根拠法令が法令データの参照ではなく法令名になっているなどデータ間の参照を行うことでデータの有効活用ができるようになるものなどが多く見受けられた。
- ✓ 調査結果から重複を取り除く方向の統合（データの内容の統合）のメリットは大きくない。一方、1つのデータを複数のシステムが利用できるようにすること（システムからのデータの分離、データの構造、形式、書式の標準化）によるメリットが大きいことが分かった。
(課題) 全体としては多くはないが重複するデータがあるため、データのすり合わせや担当の振り分けなどが必要となる。
(リスク) データの統合によりデータ間で更新タイミングのずれが生じる可能性がある。

運用・オペレーション

- 運用・オペレーションに関しては調査対象のいずれのシステムもデータオーナーが各システムの管理画面によりデータの登録・編集を行う、非公開となったデータも保持しているなど基本的な運用モデルは同一であった。
- ✓ 調査結果から、データ統合により運用・オペレーションが変更となる点については考慮が必要だが、データ統合にあたって、大きな機能的な障壁はない。
(課題) データの統合により登録・編集の作業が変更になるため円滑な移行が可能になるようにマニュアル類の整備などが必要になる。
(課題) システム主体の考え方（データは各システムに従属する）から、データ主体の考え方（データは独立したものとして共有され、各システムは共有されたデータを利用する）への移行が必要となる
(リスク) データ主体の考え方が浸透するためには時間が掛かる可能性がある。

各システムの統合によるメリットと難易度の考察

- システムによって統合のメリットが少ないものや、統合の難易度が高いものがあった。

システム名	統合	統合対象とならない理由
マイ制度ナビ	可	
ぴったりサービス	可	
e-MAFF		規模が大きく複数のシステムの連携によって成り立っていることからID・データ統合による影響範囲が大きく、統合に際してシステムの比較的大規模な改修が必要となるため短期的な統合対象とするのは難しいと思われる。
jGrants	可	
e-Laws		データの内容が法令の文書であり他のデータとは性質が大きくことなるからデータの書式や形式などの統一の必要はない。また、URIやAPIの利便性向上などの課題はあるものの、すでに文書を参照するためのAPIが用意されていることから統合によるメリットが少ない。
ミラサポplus制度ナビ	可	

ID・データの統合 1

ID・データ統合にあたっては、先ず統合ID体系の策定した後、統合データモデルの構築及び実際にデータを各システムに提供するようなプラットフォーム（以下、「データプラットフォーム」）を構築することが必要になる。データモデルやデータプラットフォームの構築とシステムからの利用を平行して行うことには大きなリスク伴うことやデータのすり合わせを短期で行わないといけないといったことから、下記のように大きく2つのフェーズに分けてデータ統合を推進するのが良い

統合ID体系の策定

- 統合データの各要素をシステムに依存することなく、データセットを越えて一意に特定できるようなID体系の策定

統合データモデルの構築とデータプラットフォームの構築

■ フェーズ1（準備段階、統合データとしてのデータの収集とデータモデルの構築）

- アプリ主体のモデルを維持したまま、ID体系、データの構造・形式・書式などを標準化したデータモデルの構築及び構築したデータモデルによるデータ収集を行い、アクセス方法・認証などの統一・統合などの要件をみたま形でのデータの収集しデータプラットフォームを構築することを主な目的とする。

■ フェーズ2（統合段階、各システムによる統合されたデータの活用）

- データ主体のモデルに移行する。フェーズ1で構築されたデータプラットフォームのデータを直接的又は間接的に利用するようにする。この変更に伴い、データオーナーはシステムを通してではなく、データプラットフォームのデータ管理を行うようにする。

統合ID体系の策定

実体調査の結果から、各システムは独立して動作しているためID体系はシステム内での唯一性のみが考慮されていることに加え、データの共有にあたってはより一層の継続性が求められる点からID・データの統合に際してはシステムの枠を越えて唯一性を担保し、かつ、十分な継続性が保証されるID体系が必要となる。

本事業では、上記を踏まえ下記の要件を満たす次の要件を満たすID体系を提案する。

- システムを越えて基本的な要件（唯一性、継続性、悉皆性）を担保できること
- 事業の移管、組織の統廃合があった場合であっても基本的な要件を満たす
- システムは独自に新しいIDを発行できること
- システムの大規模な改変を必要としないこと
- 識別対象の更新があった場合にもデータ連携に支障をきたさないこと
- インターネット上のリソースとして公開する際に利用できること
- 手動でのID管理を行うような場合にも、ある程度の分かり易さがあること

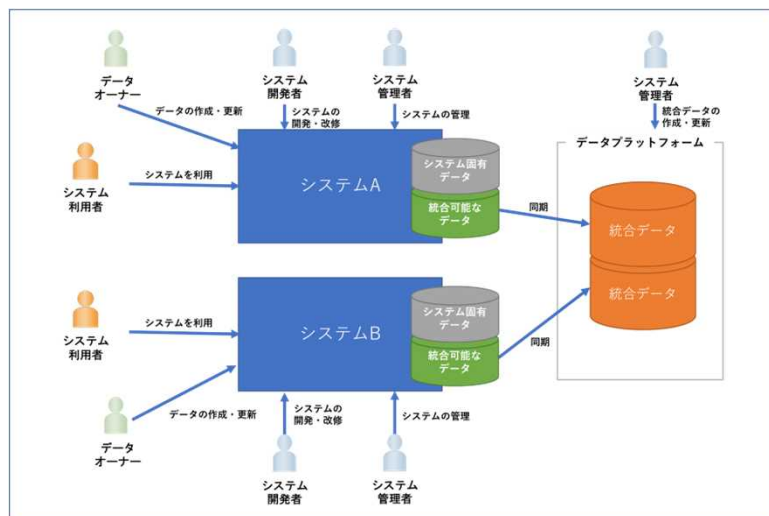
具体的には次のような構造をもつ無意コードの複合IDとなる。詳細については、別紙「統合ID仕様書」に記述する。

《ID》 := 《シグネチャ》《スキーマバージョン》 "+" 《データセット作成者コード》 "+" 《データセット識別コード》 "+" 《個別ID》

リスクと課題

- IDの利用に際しては、どのようなID体系を定めたとしても、正しく運用されない場合には唯一性や継続性の問題が生じるため、運用方法を徹底する必要がある。
- 完全な唯一性を担保するためには、何かしらの一極集中管理が必要となる。上記のID体系は管理コストを抑えるよう考慮されているが、管理体制の整備は必要である。

データ統合のフェーズ 1



フェーズ 1 では、可能な限りデータオーナーの作業に影響を与えないようにしつつ、既存のデータを収集し、ID体系、データの構造、形式、書式が統一されたデータプラットフォームを構築することを主目的とする。

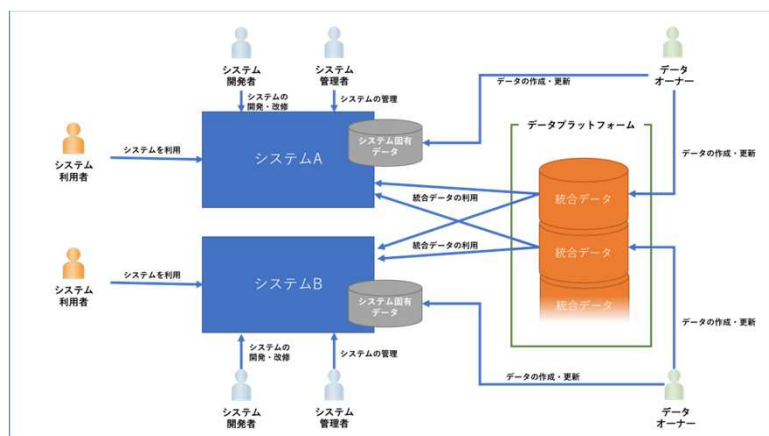
作業

- データプラットフォームの簡易実装（データ収集に特化）
 - 可能な限りシステムから独立したツールで対応できると。
 - システムの改修が必要な場合は稼働時の動作に影響がない部分での改修で対応できること。
- 各システムからのデータ取り出し
 - データ取得API又はエクスポート機能をもたないシステムについてはいずれかを実装する。
 - 取り出したデータを同期するためのツールの実装。
 - ※ 当初データの即時更新は求められないため、手作業での取り出しでも可
- データの収集とデータモデルの構築

リスクと課題

- 運用面ではデジタル庁における作業比重が大きい。運用面を支援するツールがあるとよい
- 重複データのすり合わせには、各データオーナーの協力が必要になる場合がある。

データ統合のフェーズ 2



フェーズ 2 においては、データ統合の最終モデルとして各システムから独立したデータ提供システムとしてのデータプラットフォームを構築する。データの管理はデータプラットフォームに対して行うことを中心とし、各システム上ではシステムに固有の情報のみを管理するモデルへの移行を目的とする。

作業

- データプラットフォームの強化
 - データ提供APIの整備、データの追加・編集を行うための管理機能の強化、各データオーナーが利用することになるため、直感的な利用が可能でかつ利便性の高いユーザーインターフェースの構築、データの変更などを各種通知する機能の実装
 - データセキュリティ（認証機能とアクセスの許可・拒否のコントロールなど）
- 各システムの改修
 - 独自のデータを利用するモデルからデータプラットフォームが提供するデータの利用へ変更
 - 運用モデル移行のためのマニュアルの整備

リスクと課題

- データプラットフォームの負荷対応
- データ管理の運用が変更となるため、データオーナーへの情報提供が必要

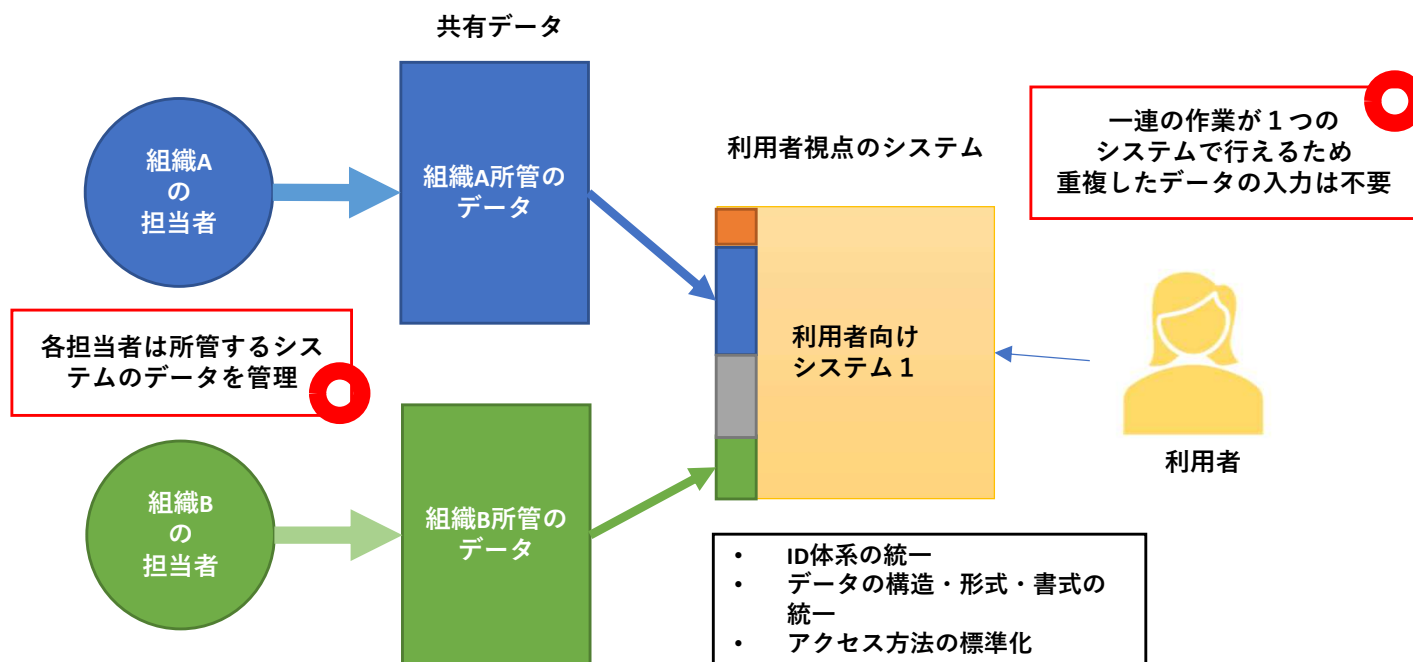
データ主体モデルのメリット

まずは、オフィスのオペレーションを例に考えてみる。あるスタッフ（例えば〇〇さん）が外出する際に他の一人のスタッフ（例えば△△さん）に外出先と戻り時間を伝えるとし、△△さんは自分のメモ帳にこの情報を書いておく。この方法では、〇〇さんの戻り時間を知りたい人は△△さんに聞く必要がある。これは人間が主体の考え方となる。それに対し、〇〇さんが外出時に決められた形式でホワイトボードに書いておくようにすれば誰でもその情報を利用できるようになる。これがデータ主体の考え方である。

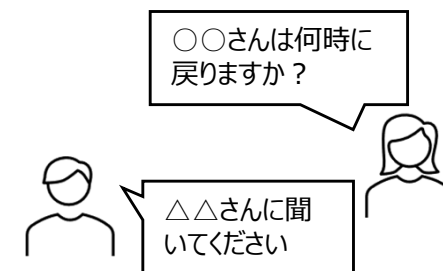
同様なことをコンピューターシステムに当てはめて考えることができる。データを登録・編集する側に都合のよいデータの範囲とデータを利用する利用者やシステム側にとって都合の良い範囲が一对一に対応しない。具体的には、データ管理の観点からは管理者は一か所で担当するデータを登録・編集したい。一方、利用者はデータの所管とは別に関連する作業は一か所で行えるのが良い。システム主体のモデル（データが各システムに従属するモデル）では、実装コストや運用面からこの2つのニーズを両立することができない。

このように、データ（情報）が特定のシステム内にあると、データを有効に活用できないだけでなく利便性の高いシステムを構築することが難しくなる。

データ主体のモデルを有効に運用・活用するにあたっては、データの管理者、利用者（データを利用するシステム）ともに複数となり、かつ異なるデータ管理者によって管理されるデータを同様に利用できるようにするために、ID体系、データの構造、形式、書式を統一することが極めて重要となる。



人間主体



データ主体

