

異なるデータ ID の統合・統一化に関する調査研究 最終報告書

2023 年 3 月

株式会社アスコエパートナーズ



目次

1	はじめに.....	1
2	関連資料.....	1
3	用語.....	1
4	システムの実体調査.....	3
4.1	調査概要.....	3
4.2	マイ制度ナビ.....	3
4.2.1	ID・データの統合.....	3
4.2.2	運用・オペレーション.....	4
4.3	ぴったりサービス.....	5
4.3.1	ID・データの統合.....	5
4.3.2	運用・オペレーション.....	7
4.4	jGrants.....	8
4.4.1	ID・データの統合.....	8
4.4.2	運用・オペレーション.....	8
4.5	e-MAFF.....	9
4.5.1	ID・データの統合.....	9
4.5.2	運用・オペレーション.....	10
4.6	e-Laws.....	11
4.6.1	ID・データの統合.....	11
4.6.2	運用・オペレーション.....	11
4.7	ミラサポ plus 制度ナビ.....	12
4.7.1	ID・データの統合.....	12
4.7.2	運用・オペレーション.....	13
4.8	実体調査のまとめ.....	14
4.8.1	ID 体系について.....	14
4.8.2	データセットについて.....	14
4.8.3	運用面について.....	15
4.8.4	システムごとの統合可能性の可否.....	15
5	データ ID 統合及びデータプラットフォームの整備に関する研究・検討.....	16
5.1	総論.....	16
5.2	現在のモデル.....	17
5.3	ID・データ統合のフェーズ 1.....	18

5.3.1	データの統合及びデータプラットフォーム整備のモデル検討	18
5.3.2	フェーズ1におけるID・データ統合に必要な作業と課題	19
5.4	ID・データ統合のフェーズ2	22
5.4.1	データの統合及びデータプラットフォーム整備のモデル検討	22
5.4.2	ID・データ統合にかかる負担と課題	22
6	ID・データ統合に関する検討事項	24
6.1	IDの統合に関する検討	24
6.1.1	概要	24
6.1.2	ID体系に関わる要件	24
6.1.3	運用上の要件	25
6.1.4	IDの種類と特性	25
6.1.5	既存のID体系に対する調査	27
6.1.6	ID・データ統合に向けたID体系の提案	28
6.2	データの独立のメリットと重要性	33
6.2.1	利便性向上のメリット	33
6.2.2	技術的なメリット	35
6.2.3	将来的なメリット	37
7	まとめ	38
8	参考資料	38
8.1	政府相互運用性フレームワーク（GIF）	38

1 はじめに

既存システムの効率的な運用及び将来的なデータの効果的な活用を踏まえた、データ ID の統合・統一化に関する調査を実施するとともに、構築することが望まれるモデルについての研究・検討を実施した。

本報告書では、調査対象とする 6 つのシステムの調査及び関連する技術等の調査結果を報告するとともに、調査結果をもとに研究・検討を行った結果として期待される ID・データ統合のモデル及びその構築に係る技術面及び運用面において必要となる作業（コスト）と課題について検討した結果について記述する。

2 関連資料

調査の詳細及び仕様書類については、本報告書には含めず別紙とした。

- 「各システムの登録データ比較表」
調査対象したシステムのデータセットの各データ項目を、GIF を基準に整理した内容
- 「各システムの運用方法の比較表」
調査対象としたシステムの運用オペレーションについて実施したアンケートの内容
- 「統合 ID 体系仕様書」
データの統合に向けて提案する ID 体系の仕様書

3 用語

用語	説明
システム	データベースとその管理機能(データベース自体の設定や、データの登録・変更・削除など)、データベースに保存されているデータの閲覧機能や公開機能などを持ったもの。 例：マイ制度ナビ、ぴったりサービス
ID	データを識別するための識別子。 本書においては、データベースによって自動的に発行される ID の他、データを一意に識別する目的があれば管理番号等も ID とみなす。 例：マイ制度ナビで使用される ID「2000012100001-0099990110」など

統合 ID	<p>各システムの ID をシステムの枠を越えて唯一性が保たれるように統合した ID。</p> <p>別紙「統合 ID 体系仕様書」で仕様を記載する ID は、統合 ID としての使用を想定している。</p>
統合データ	<p>複数のシステムからアクセス可能な ID 体系、データの構造、書式、形式が統一されていて、複数のシステムの間で重複するデータがある場合には、データ項目が統一され単一の仕組みの中で運用できるようにしたデータのこと。</p> <p>例：</p> <p>ここで統一する「データ項目」とは、「概要」や「金額」などの項目名のリストに留まらず、「概要」はテキスト、「金額」は数値、などのデータの形式も併せて考える点に注意が必要となる。</p> <p>統合データのデータ項目を規定する方法としては、全システムの項目の重複を省いて列挙する、全システムで共通して利用する項目を列挙する、などが考えられる。</p> <p>「全システムで共通して利用する項目」には、すべてのシステムで現に使用されている項目の他、一部のシステムでしか使用されていないが、ルール(項目名やデータの持ち方、運用方法など)としては全システム共通で定めておいたほうが良いと考えられるものも含む。</p> <p>※具体的な規定は本書内にて検討のため、上記は抽象的な一例である。</p>
運用	システムを最新の状態で利用できる状態に保つこと。本書においては特にシステム内のデータにおける登録・編集・削除の実施を指す。
システム固有データ	各システムのデータがもつ要素のうち、統合データに含まれないもの
システム管理者	システムの設定、保守などの管理を行う人。
システム開発者	システムの開発や改修を行う人。
データオーナー	システムへのデータ登録や、既に登録されているデータの編集・削除などを行う人。
統合データ管理者	統合されたデータに対して、データ登録や、既に登録されているデータの編集・削除などを行う人。
システム利用者	公開されたシステムを利用する人。ユーザー。単に「利用者」とも呼ぶ。
ローカル ID	システム固有の ID
グローバル ID	システムの枠を越えて唯一性を担保する ID
データプラットフォーム	統合データを集中的に管理するシステム。RESTful API 等を通してデータの提供を行う。必ずしも物理的に 1 つのシステムとは限らない。

システム主体のモデル	データは各システムに従属するという考え方。今回調査対象とした6システムは現時点ですべてシステム主体のモデルとなっている。
データ主体のモデル	データはシステムから独立したものであり、共有される様々なデータを利用してシステムを構築するという考え方。 Web サービスなどで主流の考え方で、例えば、自動車や公共交通機関でのルート検索可能な地図サイト、乗換案内、カーナビアプリなどの一連のシステムは、時刻表、道路図、地図などのデータの上に構築されている。

4 システムの実体調査

4.1 調査概要

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・データの統合の側面と運用・オペレーションの側面から下記の項目について調査を実施した

- 調査対象システム
 - マイ制度ナビ
 - ぴったりサービス
 - jGrants
 - eMAFF
 - e-LAWS
 - ミラサポ plus 制度ナビ
- ID・データの統合
 - ID 体系
外部からの参照の可能性が高いID 体系の基本的な構造及び基本要件に対する考察
 - 登録・管理データ
登録されているデータのうち統合データとしての利用が想定されるデータの調査
- 運用・オペレーション
 - 情報登録・更新のフロー
情報登録や更新のフロー、担当者（組織、部署）や更新の頻度など
 - 問い合わせ・サポートフロー
データ管理者および利用者からの問い合わせなどのサービス状況

4.2 マイ制度ナビ

4.2.1 ID・データの統合

4.2.1.1 ID 体系

マイ制度ナビで利用されている制度情報の ID は下記の構造をもつ

ID 名称	説明
制度番号	制度法人番号 + 制度一連番号
制度法人番号	法人番号 13 文字
制度一連番号	自動附番 数値

4.2.1.1.1 ID の唯一性

この ID は制度法人番号が文字列である一方、制度一連番号に数値を使用している。制度番号はこれら二つをつなげたものであるため制度一連番号が文字列に変換されていると考えられるが、仕様書では文字列への変換書式については記載がなかったため不明である。固定長であるか可変長であるかは変換書式による。固定文字数への変換となっている場合には、潜在的な唯一性のリスクはあるがシステムで十分に考慮されていると仮定すれば問題がないと判断できる。

4.2.1.1.2 ID の継続性

法人番号を含むことで継続性の懸念がある。現時点では自治体が提供する制度・手続きに限定されていて、自治体間での事業移管がないこと、自治体の統廃合に際しては制度・手続き自体が新規となるか廃止となることからリスクはないと判断できる。

4.2.1.1.3 ID の発行

この ID はシステム外部から提供されるものではなく、制度が登録された際にシステムによって自動的に生成される。

4.2.1.1.4 識別対象更新時の対応

識別対象が更新された場合であっても、ID の変更は行わない

4.2.1.2 登録・管理データ（データ情報、アカウント情報等）

主に制度関連のデータを保持している。特に統合対象と考えられるデータとしては「制度情報」及び「事例情報」がある。

詳細については別紙「各システムの登録データ比較表」に記載する。

4.2.2 運用・オペレーション

4.2.2.1 情報登録・更新のフロー

- 所管組織の担当者がデータを登録・編集する
- 自分が所属する部局の制度情報のみ登録・編集可能
- 「入力者」と「承認者」がいて、「入力者」はデータの登録・編集はできるが公開はできない。「承認者」のみデータの公開が可能。
- 管理画面にログインしてデータを登録・編集する
- データの更新は随時行う
- 保持期間は定めていない
- 定期的に古いデータが残っていないか確認している
- 制度登録時に「公開」「非公開」の設定が可能。「公開」の場合は即時反映される。

詳細については別紙「各システムの運用方法の比較表」に記載する。

4.2.2.2 問い合わせ・サポートフロー

- データ管理者向け
 - 管理画面上に Q&A を用意している
 - システムの所管担当者が電話・メールで対応
 - 委託事業者が電話・メールで対応
- 利用者向け
 - web サイト上に Q&A を用意している
 - お問い合わせフォームがある (委託事業者が回答)
 - コールセンターがある (委託事業者が対応)

4.3 ぴったりサービス

4.3.1 ID・データの統合

4.3.1.1 ID 体系

「ぴったりサービス」システム内では様々な ID・コードが使用されているが、その多くはシステム内（システム固有の xml 仕様を含む）に閉じたもの、又は、一時的に使用されるものである。

ID・データ統合に関係する ID に該当するものとしては、「制度・手続きコード体系」がある。「制度・手続きコード」は各種 API において共通に使用されているコード体系であり、「引越しポータル」などの外部サービスからの API 呼び出しに使用される。

「制度・手続きコード体系」は、以下のような体系の固定長・有意 ID である。

固定部	法人番号部	ID 部	枝番部
2 文字	13 文字	36 文字	3 文字

4.3.1.1.1 ID の唯一性

この ID は固定長ではあるが、ID 部の 36 文字は UUID が使用されることが想定されているため基本的には唯一性を損なうことはない（後述のように厳密には UUID は唯一性を保証しない）。

4.3.1.1.2 ID の継続性

法人番号を含むことで継続性の懸念がある。現時点では自治体が提供する制度・手続きに限定されていて、自治体間での事業移管がないこと、自治体の統廃合に際しては制度・手続き自体が新規となるか廃止となることからリスクはない。

ただし、将来公共インフラに関する民間事業者との連携などが行われるようなケースでは注意が必要である。この ID の特異な点としては、手続き名称などによる検索の結果として呼び出し側が入手するものであるという点である。この性質により ID の変更を比較的容易に行うことができる。呼び出し側が ID を保存できるようになっているため、自由に ID が変更できるという意味ではないが、ID の変更を行う際に呼び出し側が記憶している ID を破棄又は更新してもらうことで対応が可能になっている。

ID 体系を変更する場合、後述の ID の自動発行のロジックの変更などが必要になるが、ID 部が無意 ID であることと合わせ、システムの比較的小規模な変更により ID 体系を変更することが可能と考えられる。ID 体系が変更された場合には、呼び出し側の変更も必要になるが、ID 部が無意 ID であることによりこちらも比較的軽微な変更で対応が可能と考えられる。

4.3.1.1.3 ID の発行

この ID はシステム外部から提供されるものではなく、制度が登録された際にシステムによって自動的に生成される。

4.3.1.1.4 識別対象更新時の対応

識別対象が更新された場合であっても、ID の変更は行わない

4.3.1.2 登録・管理データ（データ情報、アカウント情報等）

主に手続き関連のデータを保持している。特に統合対象と考えられるデータとしては「申請情報」及び「手続き情報」がある。XML を使用した API が整備されている。

詳細については別紙「各システムの登録データ比較表」に記載する。

4.3.2 運用・オペレーション

4.3.2.1 情報登録・更新のフロー

- 基礎データについてはデジタル庁が登録編集を行う。地方公共団体のデータについては各地方公共団体が登録・編集する
- 自分が所属する部局の制度情報のみ登録・編集可能
- 「入力者」と「承認者」がいて、「入力者」はデータの登録・編集はできるが公開はできない。「承認者」のみデータの公開が可能。
- 基礎データについては Excel から DB に反映する。地方公共団体のデータについては管理画面にて登録・編集する
- データの更新頻度は基礎データと地方公共団体が管理する制度で異なる。
 - 制度・手続の基礎データはおよそ 1 か月程度の不定期で更新される。
 - 地方公共団体の制度・手続データについてはデータの登録・編集後にすぐに反映される。
- 保持期間は定めていない
- 基礎データについては依頼を受けて変更作業を行う。地方公共団体のデータについては、古いデータの削除なども含めて各地方公共団体が責任をもつ。
- 公開が不要になったデータは基礎データ・地方公共団体のデータともに「非公開」に設定する。

詳細については別紙「各システムの運用方法の比較表」に記載する。

4.3.2.2 問い合わせ・サポートフロー

- データ管理者向け
 - 委託事業者が電話・メールで対応
- 利用者向け
 - お問い合わせフォームがある
 - 電話での問い合わせを受けている

4.4 jGrants

4.4.1 ID・データの統合

4.4.1.1 ID 体系

この ID はシステム外部から提供されるものではなく、制度が登録された際にシステムによって自動的に生成される。管理番号等は任意に設定可能となるが、現在の運用においてシステムの意味合いでの ID にはあたらない。

4.4.1.1.1 識別対象更新時の対応

識別対象が更新された場合であっても、ID の変更は行わない

4.4.1.2 登録・管理データ（データ情報、アカウント情報等）

主に申請及び制度関連のデータを保持している。

詳細については別紙「各システムの登録データ比較表」に記載する。

4.4.2 運用・オペレーション

4.4.2.1 情報登録・更新のフロー

- 補助金事務局の担当者が登録編集を行う。
- 基本的には管理画面にて登録・編集するが、過去データをコピーすることが可能。
- 申請のタイミングがあるので、データの更新頻度は特に定めていない。
- 補助金は終了日が決まっていて登録時に終了日を指定するため古いデータが残ることはない。
- 募集期間中に変更されることはほぼない。
- 公開が不要になったデータは自動的に「非公開」に設定される。

詳細については別紙「各システムの運用方法の比較表」に記載する。

4.4.2.2 問い合わせ・サポートフロー

- データ管理者向け
 - 管理画面上に Q&A を用意している
 - 委託事業者が電話・メールで対応
- 利用者向け
 - Web サイトに Q&A を用意している。
 - 各データ管理担当者が問い合わせ先となる

4.5 e-MAFF

4.5.1 ID・データの統合

4.5.1.1 ID 体系

大規模なシステムであり多くの ID が使用されている。すでに複数の機能ごとのシステムが連携するように設定されているため全体として混同が起きないように設計されている。

4.5.1.1.1 ID の唯一性

すでに複数のシステムが連携する設計となっているため、全体としては唯一性を損なうリスクは低いと考えられる。ただし、単一の ID ではなく複数の ID によって識別するような部分もあるため、他の省庁のシステム等とさらなる統合を行う場合には注意が必要である。

4.5.1.1.2 ID の継続性

同一 ID の条件が厳格であるため、ID のライフが比較的短い。このことから継続性を損なうリスクは低いと考えられる。

4.5.1.1.3 ID の発行

システムによって自動的に生成される。

4.5.1.1.4 識別対象更新時の対応

識別対象が更新された場合であっても、ID の変更は行わない。ただし、同一 ID の条件は厳格であり手続き内容に変更がある場合には新規の ID となる。

4.5.1.2 登録・管理データ（データ情報、アカウント情報等）

非常に多くの種類のデータを保持している。他のシステムと関連が強いデータでは、制度情報や申請情報などがある。

詳細については別紙「各システムの登録データ比較表」に記載する。

4.5.2 運用・オペレーション

4.5.2.1 情報登録・更新のフロー

- データは必要に応じて随時登録・編集する。ただし、公開は2週間に1回の頻度で実施している。
- 所管組織が登録・編集を行う。
- 管理画面にて登録・編集する。
- 一部に Excel ファイルのインポート機能がある。
- 保持期間は定めていない
- 仕組み上古いデータが残ることがない。
- データ更新の要否は運用モデルにより把握できるようになっている。
- 申請については受付期間の終了とともに非公開となる。
- 古いデータについてはアーカイブ環境に移動される。

詳細については別紙「各システムの運用方法の比較表」に記載する。

4.5.2.2 問い合わせ・サポートフロー

- データ管理者向け
 - システムの所管担当者が電話・メールで対応
 - 委託事業者が電話・メールで対応
- 利用者向け
 - お問い合わせフォームがある
 - 電話での問い合わせを受けている

4.6 e-Laws

4.6.1 ID・データの統合

4.6.1.1 ID 体系

e-Laws では各ドキュメントに 15 文字の固定長 ID を付与している。

以下のような体系の固定長・有意 ID である。

年号	年	法令種別コード	法令号番号
1 文字	2 文字	9 文字	3 文字

4.6.1.1.1 ID の唯一性

この ID は固定長ではあり、省令に対するビットパターンが飽和に近い状態にある。唯一性を保つための方法は考えられるため唯一性が損なわれることはないと思われるが、体系としては複雑になる可能性がある。

4.6.1.1.2 ID の継続性

省令に対するビットパターンが飽和に近い状態にある。継続性を保つための方法は考えられるため継続性が損なわれることはないと思われるが、体系としては複雑になる可能性がある。

4.6.1.1.3 ID の発行

システムは ID を自動的には発行しない。ID は識別対象の特性に基づいて人的に作成される。

4.6.1.1.4 識別対象更新時の対応

識別対象が更新された場合であっても、ID の変更は行わない。修正ドキュメントについては独自の記法に基づいて ID を連結したものが与えられる。

4.6.1.2 登録・管理データ（データ情報、アカウント情報等）

法令のドキュメントを XML 形式で保持している。

4.6.2 運用・オペレーション

4.6.2.1 情報登録・更新のフロー

- デジタル庁が登録・編集する
- 管理画面にて登録・編集する
- 法令の改正のタイミングで随時更新する。
- 保持期間は定めていない
- 未更新のデータの有無を可視化・確認する機能がある。
- 効力を失ったものも含めすべてを公開している。

詳細については別紙「各システムの運用方法の比較表」に記載する。

4.6.2.2 問い合わせ・サポートフロー

- データ管理者向け
 - 管理画面上に Q&A を用意している
 - コミュニケーションツールを使用
- 利用者向け
 - Web サイト上に Q&A を用意している。
 - お問い合わせフォームがある。

4.7 ミラサポ plus 制度ナビ

4.7.1 ID・データの統合

4.7.1.1 ID 体系

柔軟性のある次のような可変長の ID 体系を採用している。

《所管組織》-《パート1》-《パート2》

4.7.1.1.1 ID の唯一性

ID 体系に柔軟性があるため、その点では唯一性のリスクは低い。ただし、制度番号が再利用可能であることから、システムを越えての唯一性は保証されないためデータ統合に際しては対策が必要となる。

4.7.1.1.2 ID の継続性

所管組織が含まれていて、所管組織が変更になった場合には ID が変更になってしまう。しかし、国が所管する場合には省庁ごとに分けることはせず国の単位で固定の番号が付与されるため、省庁間の移管が問題になることはない。

4.7.1.1.3 ID の発行

人が指定することもできる。指定されていない場合は、システムが自動的に ID を発行する。

4.7.1.1.4 識別対象更新時の対応

識別対象が更新された場合であっても、ID の変更は行わない

4.7.1.2 登録・管理データ（データ情報、アカウント情報等）

主に制度関連及び事例のデータを保持している。

詳細については別紙「各システムの登録データ比較表」に記載する。

4.7.2 運用・オペレーション

4.7.2.1 情報登録・更新のフロー

- 管理画面にてデータの登録編集を行う。
- CSV のインポート機能もあるが、新規の場合のみ使用できる。
- 基本的には各所管担当者がデータの登録編集を行う。
- データの更新は必要な際に随時行われる。
- 保持期間は定めていない
- 定期的にデータの不足等を確認している。
- 制度登録時に「公開」「非公開」の設定が可能。
- 「公開日」「公開終了日」を設定することで、「公開」「非公開」を自動的に切り替える機能をもつ。

詳細については別紙「各システムの運用方法の比較表」に記載する。

4.7.2.2 問い合わせ・サポートフロー

- データ管理者向け
 - システムの所管担当者が電話・メールで対応
- 利用者向け
 - 電話での問い合わせを受けている

4.8 実体調査のまとめ

4.8.1 ID 体系について

ID については、当然ながらいずれのシステムにおいてもシステム内における唯一性のみが考慮された ID 体系を用いている。そのためデータ統合を行うためには、別紙「統合 ID 仕様書」に示すような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。

ID については、ID の基本要件（唯一性、継続性）の観点から懸念があるものもいくつかあった。特に、マイ制度ナビ、ぴったりサービス、ミラサポ plus 制度ナビなど多くのシステムがデータ管理者の情報を含んでいるために継続性の懸念があることが分かった。地方公共団体間での業務移管はほぼ無いと考えてよいこと、及び、地方自治体の統合に際しては新しい ID を付与するなどの対応ができるため必ずしもリスクが大きいとは言えない。

しかし、データ統合を行った場合には、1 システム内に閉じた運用を行う場合に比べ、ID の変更による影響範囲が非常に広がるため、統合データでは継続性の面からも、別紙「統合 ID 仕様書」に示すような業務移管、組織統合などに影響されにくい ID 体系を導入する必要がある。

また、e-Laws のように ID 体系の設計上の上限（唯一性、継続性の維持のために ID 体系、もしくは、運用面で何かしらの対応が必要になる可能性）に近づいているシステムでは、データ統合により影響範囲が広がることを考慮し早めに対応する必要がある。

4.8.2 データセットについて

今回調査対象とした 6 システムの間では、部分的に重複するデータは存在するもののデータセットの大半が共通するようなデータは少ないという調査結果となった。重複が少ないこと自体は管理上の観点からは非常にポジティブなことである。一方、各システムが提供できる情報が限られているため、利用者にとっては必ずしも利便性が高いといえない状況になる。例えば「マイ制度ナビ」では、「根拠法令」の値が法令名となっているため、根拠法令の内容を確認するためには他のシステムを利用しなくてはならないといった不便さがある。

データを統合することにより、他のシステム（データ管理者）が管理するデータを直接参照できるようにすることにより利用者の利便性を劇的に向上することが可能となる。

システム間のデータの重複及び関係性については別添「各システムの登録データ比較表」を参照。

また、統合を進めるにあたって、連絡先など同一の概念のデータ項目であっても構造、形式、書式が異なるものがある。「政府相互運用性フレームワーク（GIF）」などを参考に構造、形式、書式の統一を検討する必要がある。

特に、手続き関連のデータについては、現在 GIF には参考になるものがないが、住所などの基本要素の構造やコード体系など GIF と親和性のある形式でデータモデルを作成することが期待される。

4.8.3 運用面について

運用面に関しては、システムごとに異なる部分も多いが、公開期限が切れたデータの扱いなどデータの統合に際して問題となるような点に関しては大きな違いはないため、データ統合における技術的な障壁は見受けられなかった。

運用・オペレーションの詳細については別紙「各システムの運用方法の比較表」に記載する。

4.8.4 システムごとの統合可能性の可否

システムごとにみると、システムの構成やデータの内容などからデータ統合のメリットが大きいと考えられるものと、コストに対して統合のメリットが多くないと考えられるシステムがあった。

調査対象の 6 システムについて統合の容易さについて次の表にまとめる。ここで「統合」欄が空欄のものは「統合対象とならない理由」欄に記載の理由により、近々の統合が難しいもしくは、コストに対してメリットが十分に多くないと判断したものである。統合欄が空欄のものであっても、統合が不可能というわけではなく、今回調査項目とした以外の理由により統合を検討することも考えられる。

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	

5 データ ID 統合及びデータプラットフォームの整備に関する研究・検討

5.1 総論

本章以降では、調査結果から統合によるメリットが大きいと考えられる 4 システム（マイ制度ナビ、ぴったりサービス、jGrants、ミラサポ plus 制度ナビ）について、データ ID 統合及びデータプラットフォームの整備に関して研究・検討を行った内容を記述する。

前述の調査結果を踏まえると、どの 2 つのシステムの組み合わせにおいてもデータの重複部分を統合し、将来的にシステム自体を統合するような方向での統合はあまり効果的ではないことが分かった。一方で、利用者の利便性を高めるためには「ID 体系、データの構造・形式・書式・アクセス方法・認証などの統一・統合」という意味でのデータの統合が非常に有効である。（以下、ID 体系、データの構造・形式・書式・アクセス方法・認証などの統一・統合された一連のデータを提供するものを「データプラットフォーム」と呼ぶ）

データプラットフォームの構築にはコストがかかるが、各システムがデータ管理を行う場合に比較し次のような優位性があり将来的なコストダウン及び利便性の向上が期待できる。

- データ管理に関するドキュメントの整備やユーザーインターフェースの強化などに必要なリソースをデータプラットフォームに集中できる

- 統一されたツールを使ってデータの管理ができるため、新しいシステムに対する習熟が早くなる

データの統合を行うためには、データがシステムに従属したものであるという考え方「システム主体のモデル」から、データは独立して価値をもつという考え方「データ主体のモデル」への意識改革が必要である。データの独立によるメリット及び重要性については、「6.2 データの独立のメリットと重要性」に詳述する。

後述のように、データの統合には、システムの改装や新規開発、上述した意識改革を含む運用オペレーションの変更など時間が掛かる。そのため、まずは統合のための要件を満たす統合 ID の体系を導入することが必要となる。続いて、大きく2つのフェーズに分けてデータプラットフォームの構築を行うように逐次的に進めるのが良いと思われる。

まとめると下記のようなステップとなる。

■ 統合 ID 体系の導入

統合 ID 体系の策定と ID の管理体制を整備する。

■ 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す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。

5.2 現在のモデル

現時点では、各システムは完全に独立した実装となっているため、データの共有はなされておらず、また他のシステムが管理するデータへの参照もできない。

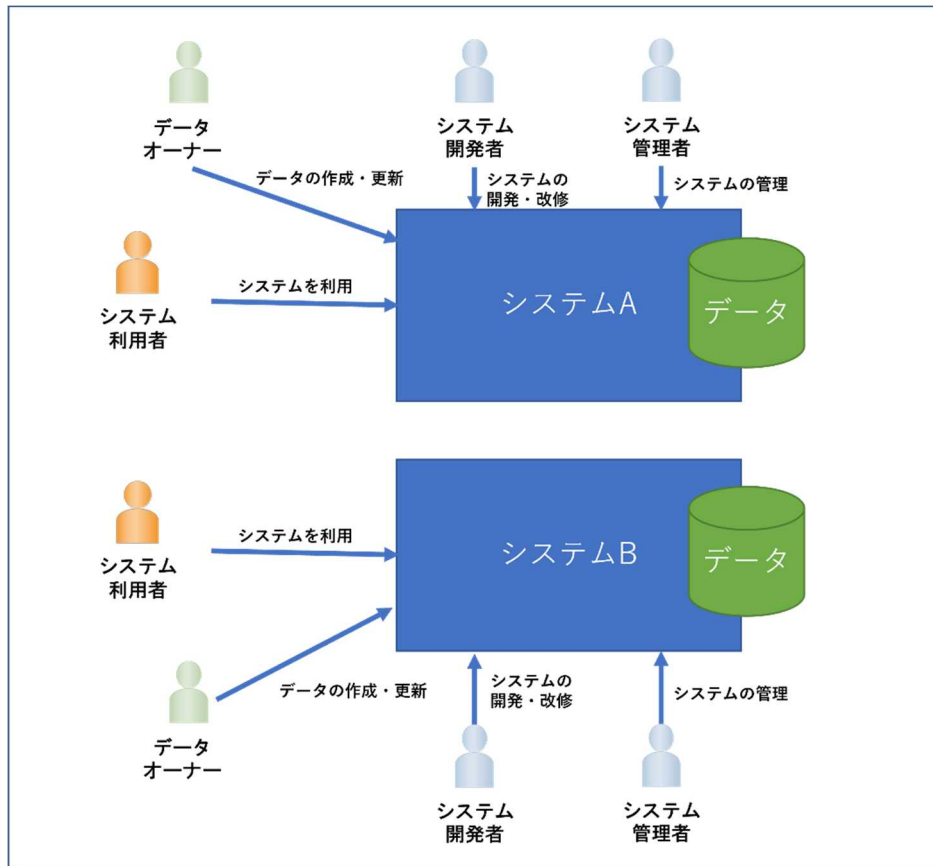


図 1 現在のモデル

5.3 ID・データ統合のフェーズ 1

5.3.1 データの統合及びデータプラットフォーム整備のモデル検討

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。

そのためには、以下の要件を満たすモデル構築が必要である。

- 各システムのコアな部分の変更が不要である。
 - 可能な限りシステムから独立したツールで対応できること。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。
- 各システムのデータオーナーの作業に影響がないようにすること。

- 統合データ部分の衝突がないこと
 - システム間にて重複するデータに対し、すり合わせを行うための技術的、運用面的な支援があること。

これらの要件を満たすモデルとして下図のようなモデルを提案する。

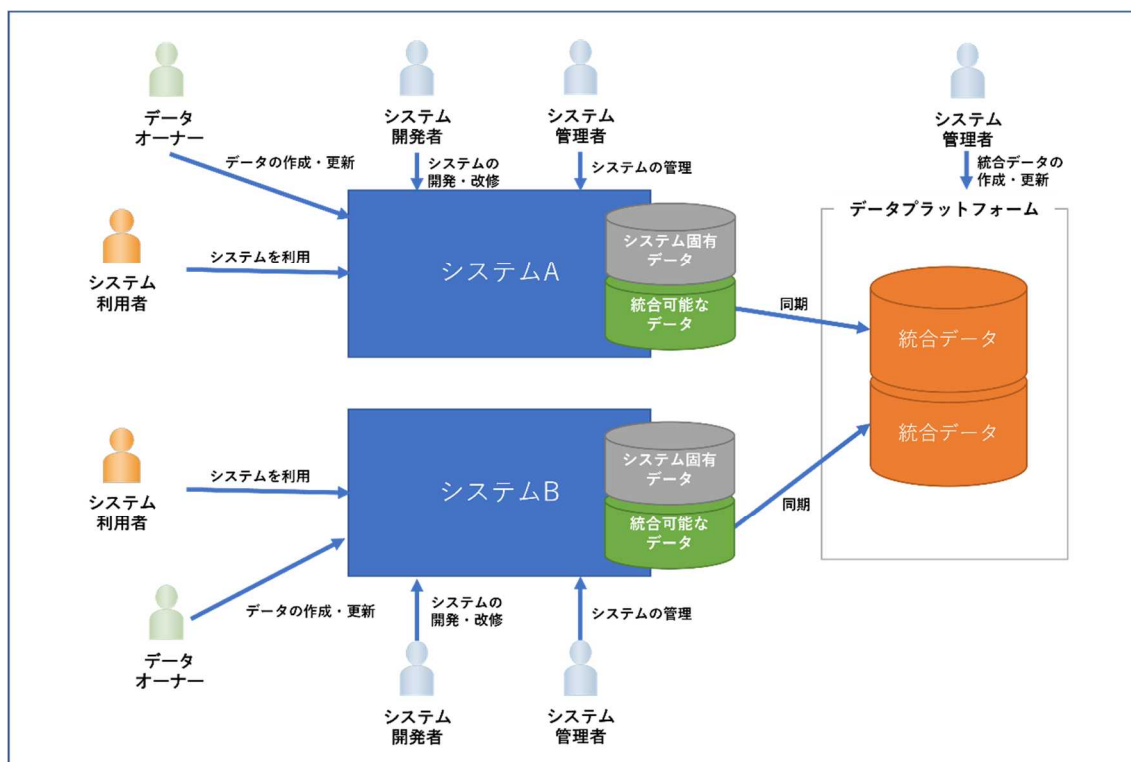


図 2 データの収集とデータプラットフォームの構築

5.3.2 フェーズ 1 における ID・データ統合に必要な作業と課題

このモデルを実現するためには、下記のような作業が必要となる。

5.3.2.1 データプラットフォームの構築

データプラットフォームは各システムから共通に使用されるものであるため、システムごとに構築する必要はない。直接公開されるものではないが、機能的には一定の範囲（国が管理するシステムなど）において自由にデータにアクセスできるようにするオープンデータのような環境を提供するシステムであり、下記のような機能をもつ。

5.3.2.1.1 データ提供 API

各システムがデータプラットフォームに格納されたデータに自由にアクセスできるような API を提供する。ただし、フェーズ 1 においては主にデータ収集が目的でありシステムから直接アクセスされることは想定されていないため、簡易な実装が良い。実装には多くの選択肢があるが利便性と実装コストから RESTful API が望ましいと思われる。

5.3.2.1.2 データの追加・編集を行うための管理機能

基本的にはデータは同期ツールによって収集される想定となるが、データの修正や追加のための簡単な管理機能をもつことが望まれる。

5.3.2.2 各システムに対応した同期ツールの構築

各システムからエクスポートされたデータ又は API により提供されたデータをデータプラットフォームに同期するためのツールが必要となる。

データを統合するためには、まずデータの各要素を識別するための ID を統合し、様々なシステムのデータ要素をシステムの枠を越えて識別できる環境を構築することが必要となる。そのためには、システムの枠を越えて唯一性・永続性・悉皆性など基本的な要件を満たせるような ID 体系にしたがった ID が必要となるが、各システムが採用する ID はシステム内でのみ唯一性が保証された ID であるため、データのエクスポート又は API の実装にあたっては統合 ID を生成する必要がある。

統合 ID の要件を満たす ID 体系については 6.1 にて提案する。

また、データの構造、形式、書式については必要に応じてデータプラットフォーム内のデータ規定にしたがってデータの変換を行う必要がある。

5.3.2.3 各システムの改修作業

5.3.2.3.1 データのエクスポート機能又はデータを取り出すための API の整備

同期ツールがシステムのデータを取り出すことができるようにするため、各システムはデータのエクスポート機能又はデータを取り出すための API を整備する必要がある。ただし、すでにエクスポート機能等を備えているシステムについては改修の必要はない。

5.3.2.4 共通データの内容のすり合わせ

複数のシステムに共通の項目については、更新タイミングのずれや担当者の違いなどにより内容に不整合が生じる可能性がある。このような不整合がある場合には同期の前に不整合の原因を調べどのシステムのデータを採用するかなど内容のすり合わせや必要に応じて各シス

テムが保持するデータの修正など調整を行う必要がある。

同期ツールなどで不整合を発見した際にデータオーナーに通知する機能があると良い。

5.3.2.5 統合データのデータモデルの定義

5.3.2.4 のステップの結果として、統合データのデータモデルを定義する必要がある。データモデルの定義は、項目ごとに、名称、書式、使用するコードなど可能な限り詳細に定義しておくことが重要である。別紙「各システムの登録データ比較表」に政府相互運用性フレームワークとの対応を示してある。（参考：政府相互運用性フレームワーク

[https://www.digital.go.jp/policies/data_strategy_government_interoperability_framework/]

5.3.2.6 データ管理・運用体制の整備

データのすり合わせや更新タイミングの調整など、データプラットフォームの管理・運用体制を整備する必要がある。フェーズ 1 においては、データオーナーの作業に変更はないため、運用体制はデータプラットフォーム及び各システムの所管組織の範囲での体制整備に留まる。

フェーズ 1 においてデータの重複や不整合を完全に排除することが重要である。

データプラットフォームを整備するにあたり、統合データ管理者をおく必要がある。この統合データ管理者の役割は主にデータの重複・不整合の検出及びデータモデルの精査・更新となる。不整合については、一時的なものかどうかを見極め、必要に応じてデータオーナーと連絡をとり解決する必要がある。

5.3.2.7 必要な対応のまとめ

	データオーナー	各システム開発者	統合実施者
データプラットフォームの構築			システム構築
同期ツールの構築			システム構築
各システムの改修		システムの更新	システム改修
データ内容のすり合わせ	不整合データへの対応		データオーナーへの通知とデータの編集
データモデルの定義			データモデルの定義
データ管理・運用体制の整備		管理・運用マニュアルの作成	管理・運用マニュアルの作成

5.4 ID・データ統合のフェーズ2

フェーズ2においては、データが各システムから独立するため、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。

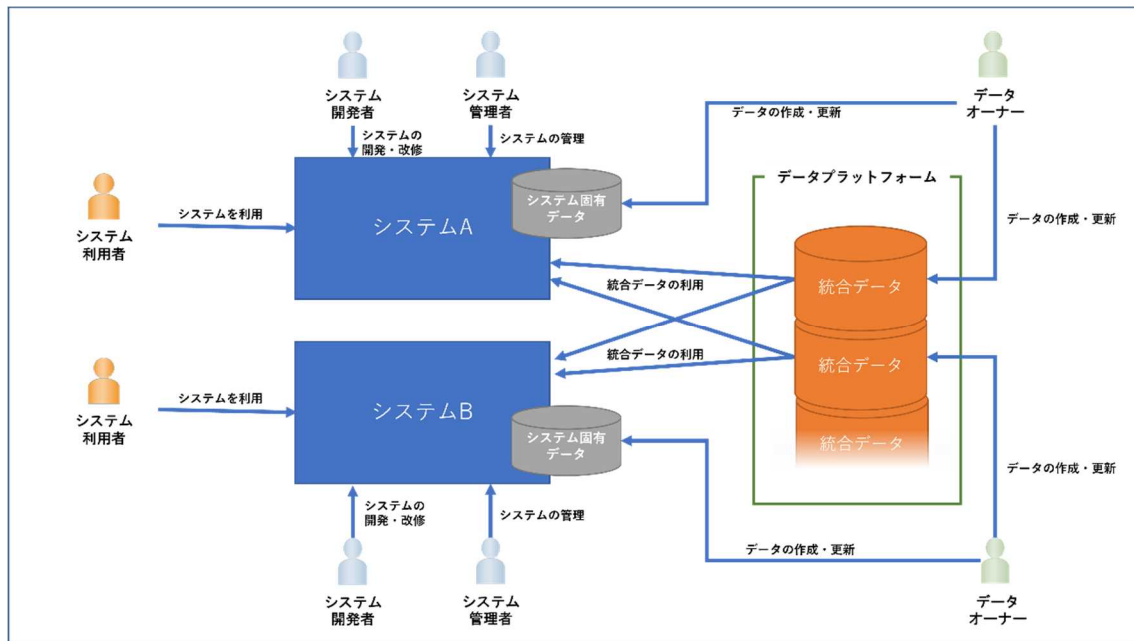


図3 フェーズ2におけるデータ統合のモデル

5.4.1 データの統合及びデータプラットフォーム整備のモデル検討

5.4.2 ID・データ統合にかかる負担と課題

5.4.2.1 データプラットフォームの構築

データプラットフォームはフェーズ1の実装に加えて各システムからの利用に対応するため下記のような機能の追加・拡充を実施する必要がある。

5.4.2.1.1 データ提供 API

各システムからのアクセスに対応するため、後述の認証機能に応じたアクセスの許可・拒否を実装する。また、非機能要件として適切な処理能力を有する必要がある。

5.4.2.1.2 データの追加・編集を行うための管理機能

実運用を行うにあたって、データの反映管理及びプレビュー（ステージング環境）の提供が必須となる。

また、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が求められる。

5.4.2.1.3 データの変更等を通知する機能

円滑な運用を行うために、関連するデータが変更された際に関係するデータオーナーに通知を送信する機能を実装することが必要となる。

5.4.2.1.4 認証機能

各システムからの API アクセス及び各データオーナーからのアクセスに対応するための適切な認証機能を実装する必要がある。

5.4.2.2 各システムの改修作業

5.4.2.2.1 データプラットフォームからのデータ取得機能

フェーズ 1 ではシステムごとに管理するデータを利用していたが、フェーズ 2 ではデータプラットフォームのデータに直接アクセスするモデルへ移行するため、それに対応したシステムに改修が必要となる。

5.4.2.2.2 データ管理機能の変更

各システムによるデータ管理はシステム固有のデータに限定され、それ以外のデータの管理はデータプラットフォームの管理機能を利用することになるため、データ管理機能の改修が必要となる。

5.4.2.3 データ管理・運用体制の整備

フェーズ 2 では、データオーナーは各システムにデータを入力するのではなく、データプラットフォームに対してデータ入力を行うことが求められる。これには、システム主体の考え方からデータ主体の考え方への移行を含め、データの管理・運用体制が大きく変化することになる。そのため、データ管理の全体マニュアル、データプラットフォームの管理画面のマニュアルなど各種の管理・運用マニュアルを作成するなど、データオーナーの業務の円滑な移行を支援する必要がある。

5.4.2.4 必要な対応のまとめ

	データオーナー	各システム開発者	統合実施者
データプラットフォームの構築	管理機能の学習	システムの更新	システム改修
各システムの改修	管理機能の変更対応	システムの更新	システム改修

データ管理・運用体制の整備	運用モデルの学習 運用体制の整備	管理・運用マニュアルの 作成	管理・運用マニュアルの 作成
----------------------	---------------------	-------------------	-------------------

利用者には直接的な影響はないが、既存システムの改修や新規システムの構築を通して、システムを移動することなく必要な情報が 1 か所で得られる、関連手続きについて 1 つのシステム上で行うことができる、さらには同じデータを繰り返し入力する必要がなくなるなどの利便性の向上が期待できる。

6 ID・データ統合に関する検討事項

6.1 ID の統合に関する検討

6.1.1 概要

データプラットフォームの構築にあたっては、グローバルな一意性及び十分な継続性などをもつ共通の ID 体系を導入することが必須。

6.1.2 ID 体系に関わる要件

ID・データ統合を行うための ID 体系に必要な技術要件には次のようなものがある。

- (1) 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

効率的な ID 体系の構築に際しては、基本的な要件の範囲及び期間を明確にすることが重要であるが、ベースレジストリにおける活用などを考慮し、効率性よりも十分に広い範囲（世界、少なくとも国内）かつ半永久的な期間を想定する。

- (2) システムの大規模な改変を必要としないこと

システムがもつ ID（以下「ローカル ID」）を変更せずに他のシステムやベースレジストリの連携時にのみ連携用の ID（以下「グローバル ID」）を使用できるようにし、かつローカル ID とグローバル ID は双方向かつ機械的に変換できること

- (3) システムは独自に新しい ID を発行できること

システムは ID 発行システムなど他のシステムに依存することなく、新しいローカル ID を発行することができ、かつ対応する唯一性が担保されたグローバル ID を機械的に生成できること。

- (4) 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

事業（システム）の移管や組織の統廃合などがあつた場合であっても、ID の変更が必要ないこと。

(5) 識別対象の更新があつた場合にもデータ連携に支障をきたさないこと

識別対象更新の際に ID を更新するかしないかはシステムの要件や目的によって異なる。識別対象更新時の ID の扱いに関する運用方法の違いを吸収し、連携に支障をきたさないようにすること。

(6) 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

インターネット上でデータを公開する際にグローバル ID から機械的に一意の URL を生成できること。これは ID 体系そのものの必須要件ではないが、機械的に URL を生成可能とすることにより、利用者にとって直感的なデータの公開が可能となるとともに、データ公開に関わるコストを削減することができる。

また、種別コードなど ID が手動で管理されるケースでは、上記技術要件に加え技術的な基本要件を満たす範囲において次のような人にとっての分かりやすさを考慮する必要がある。ただし、これらの要件の多くは技術的な要件と相反するものであり、制度情報などのような構造をもつデータに対しては、人向けの情報は ID とは別にデータの一部としてもつことが望ましい。

- (1) 特定の桁に特定の意味を持たせる（ID による抽出、概要の把握、ソードなど）
- (2) 桁数を固定する（1. による抽出を容易にする、印刷時の見栄え）
- (3) 使用できる文字種を制限する（入力が容易）
- (4) 意味のある文字列にする（記憶が容易）
- (5) 短くする（記憶が容易）

6.1.3 運用上の要件

ID・データ統合を行うための運用上の要件には次のようなものがある。

- (1) 運用・管理コスト及び負荷の集中を低く抑えること ID を運用・管理するにあたって個々の ID の一極集中管理や大規模なシステム開発などを必要としないこと。
- (2) 運用・管理体制の構築が容易かつ段階的構築が可能であること ID の基本的な要件を維持するためには、運用・管理体制の整備が必須であり、長期的には運用・管理体制を整える必要がある。しかし、運用・管理体制の整備には時間とコストが掛かるため、既存の ID の活用などにより完全な運用・管理体制を整備することなく ID 体系を利用できるようにすること。

6.1.4 ID の種類と特性

6.1.4.1 有意 ID と無意 ID

有意 ID とは ID の中に種別や分類などを内包した ID で、無意 ID とは連番のように ID そのものはなんら意味をもたない ID である。1. 有意 ID の利点 - ID のみで種別や分類を知ることができるため人にとって利便性が高い - 唯一無二の ID を生成するのが容易

有意 ID の利点

- ID のみで種別や分類を知ることができるため人間にとって利便性が高い
- 唯一無二の ID を生成するのが容易

有意 ID の弱点

- 分類軸の変化や事業移管等の影響を受けるため継続性を維持するのが難しい
- 情報量が多くなると ID が長くなり最大の利点である利便性が落ちる

無意 ID の利点

- 様々な変化に影響を受けないため、基本的な要件である、継続性を維持するのが容易である。
- ID そのものは短く効率が良い

無意 ID の弱点

- ID だけではデータの内容が分からない
- 唯一性を担保するためには、ID を管理する体制が必要となることから運用・管理コストが高くなる傾向がある。

6.1.4.2 固定長 ID と可変長 ID

固定長の ID とは ID 文字数が一定な ID のことで、可変長の ID とは文字数が ID によって変化する ID である。固定長の ID とは表示上の長さが一定という意味ではないことに注意する必要がある。

固定長 ID の利点

- 人が入力する際に間違いに気が付きやすい
- ハードウェアでの処理に向いている
- 極めて高速な処理を求められる際に有利となる場合がある
- 組み込みなど処理速度が遅い環境において有利となる場合がある
- 桁数が揃っていることで美しいと感じる人がいる

固定長 ID の弱点

- 情報の最大量が決まっているため、ID の適用範囲の変化等に弱く唯一性を維持できなくなる可能性がある。このような問題に対処するため、十分に多くのマージンをとると冗長性が高くなる。
- 0 パディングなどパディングが意味をもつことから、処理の際に注意が必要になる。

可変長 ID の利点

- 短く効率が良い
- データ量に制限がないため、将来に渡り唯一性を担保することができる。

可変長 ID の弱点

- 処理速度の問題が起こり得る。ただし、メールアドレスが可変長であるように極めて高速な処理を要求されない限り問題にならない。
- ハードウェアで処理するのには向かない。

ID 体系については、別紙「ID 体系仕様書」を参照

6.1.5 既存の ID 体系に対する調査

6.1.5.1 Uniform Resource Identifier (URI)

URI は URL(Uniform Resource Locator)、URN(Uniform Resource Name)の総称で、スキーム、ドメイン、ポート、パスなどによって構成される可変長の有意 ID である。

パス部分は入れ子にでき、かつ、自由な文字列で良いため既存の ID をパスに含めることでローカル ID から機械的に生成可能な ID を定義することが容易である。

ただし、スキーム (http: ftp: など) が ID に含まれることから実際のを指し示すのではなくスキーム (到達手段) を通してものを指し示すことになる。独自のスキームを定義することは不可能ではないが現実的ではない。また、ドメインが民間事業者により一極的な管理がなされていて間接的ではなるがドメインが特定の組織・団体に結び付けられていることから永続性の懸念がある。

データ公開時には URI と URL が同一となるのが望ましいが、URL はサーバー管理者に依存することからそのような運用は難しい。また、URI を URL の一部にする (URI を入れ子にする) ことは不可能ではないが、非常に多くのエスケープが必要となってしまうため現実的とはいえない。

このような理由から、ID・データ統合に用いるのは難しい。

6.1.5.2 Universally Unique Identifier (UUID)

UUID は特別な管理システムが不要で手元のシステムで唯一性のある ID を生成できることを特徴とする固定長の無意 ID である。

機械的に唯一性のある ID を発行することで、無意 ID の弱点である運用・管理コストを低く抑えるとともに固定長でありながらデータ量の制限による唯一性のリスクがほぼ無いため多くのシステムで使用されている。

ただし、完全な唯一性を保証するものではなく、特に乱数実装が理想的でない場合には重複の可能性がないわけではないため、トランザクションごとに随時 ID を発行するようなものには不向きである。

ID・データ統合の観点からは、重複のリスクはないと考えて良い。ただし、システムが用いる既存の ID との機械的な変換が不可能であるため ID・データ統合に用いるのは難しい。

6.1.6 ID・データ統合に向けた ID 体系の提案

6.1.6.1 概要

本研究の一環として、グローバル ID として用いるための 6.1.2 の要件に加え下記の要件を満たす ID 体系として設計された ID 体系を提案する。

- 組織、システムの枠を越えて唯一性を担保できる。
- 既存の ID との間で機械的に変換可能である。
- 既存 ID の体系によらず ID を包含することが可能である。
- システムが自動的に発行した ID に対応可能である。
- 事業移管・組織の統廃合においても継続性・唯一性を維持できる。
- 運用・管理のコストが低い

例：

G1.0+8000012010038+1-2731

これは、管理 ID 発行者 ID が 8000012010038、管理番号が 1、個別 ID が 2731 の場合の ID となる。

《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》

ID 体系の詳細な書式と説明は別紙「統合 ID 仕様書」を参照。

6.1.6.2 ID 統合のモデル

複数のシステム間で共通のデータ項目については個別に統合 ID が生成されることになるため、ID の対応付けが必要となる。ID の対応付けの方法としては、次のように一旦すべての ID をグローバル化した上で、マスター ID をハブとしてハブ & スポーク型の対応付けを行うのが良い。

また、対応付けを行うにあたっては、予めマスター ID に「名称」などキーとなる項目をもつデータを関連付けておくことで対応付けを効率良く行うことができる。

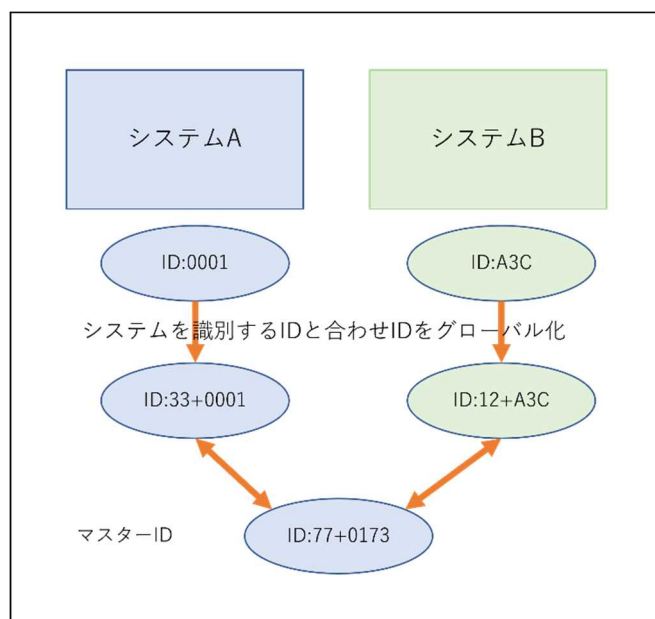


図 4 グローバル ID を用いた対応付け

この方法であれば機械的にシステムを識別できることに加え、マスター ID を統合後のマスターデータの ID としてシステムに依存しないデータを整備することができるようになる。

ID の対応付けは、同一のもの同士の対応付けだけでなく様々な関連性によって関連づけることも場合によって有効である。関連性についてはケースバイケースであるためすべてを列挙することはしないが以下のようなものが想定される。

- 同一
- 類似
- 包含

- 関連
- (旧版)
- (根拠法が同一)

括弧 (かっこ) で囲まれたものは ID の統合に役立つものではないが、グローバル ID の導入による運用上の利点となり得る。

その他の方法には以下のようなものがある。2つのシステムの統合のみを目的とする場合には有効となる方法もあるが、いずれもマスターデータの整備には不十分である。

6.1.6.2.1 ID を直接対応付ける場合

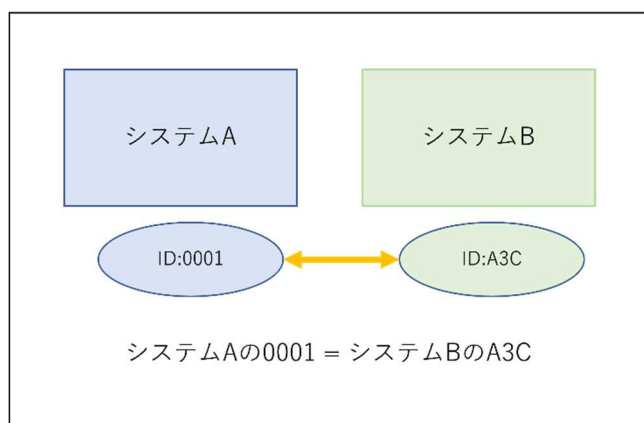


図 5 ID を直接対応付ける

比較的簡易な方法であり2つのシステムを統合する場合には、ID の管理等不要であり有効である。しかし、図中「システム A の」と記載しているようにその ID が使用されているシステムは機械的に識別されておらず運用にて補う必要がある。単に2つのシステムを統合するような場合には問題とならないが、さらに他のシステムとの連携を行うような場合には使用できない。

マスターデータの整備という観点からは、この手法はあくまでもシステム間の統合に留まり別のシステムとのデータの共有や連携が行えないため

6.1.6.2.2 グローバル ID を用いた対応付け

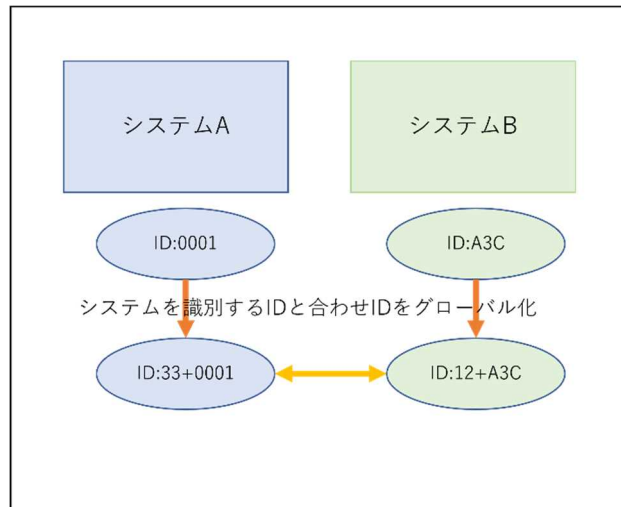


図 6 グローバル ID を用いた対応付け

2つのシステムを統合する場合、基本的には直接対応付ける方法と大きな違いはない。ただし、グローバル ID の生成にはシステムを識別する ID を管理する必要がある。

一方、システムを機械的に識別できることに加えシステムの識別に継続性があるため、複数のシステム間での連携やデータの統合として有効である。しかし、マスターデータの整備という観点からは、システム間がごとの対応関係となるためデータがシステムに依存した状態である。

6.1.6.3 ID の管理体制と運用

6.1.6.3.1 ID の管理体制

この ID 体系を運用するにあたっては「データセット作成者コード ID」の一極管理が必要となる。しかし、管理体制を整備するには時間とコストが掛かってしまう。この問題を解決するにはすでに一極管理されている ID を活用することができる。具体的には法人番号を利用することが第一の候補となる。つまり、法人番号をもつ法人がデータセットを作成する場合は、法人番号を「データセット作成者コード」として使用できることを意味する。

ただし、「データセット作成者」は「管理者」ではない点に注意する必要がある。事業移管等があった場合、ある時点で実際に管理している組織は「データセット作成者」とは異なる組織である場合がある。（参照 6.1.6.3.2）

法人登録されていない組織や法人内の組織などがデータセット作成者となることに備えて、

将来的には「データセット作成者コード」を一極管理する体制を整備していくことが望まれる。

6.1.6.3.2 ID の運用

本 ID 体系を利用する者は次の要件を満たす必要がある。

- 重複するデータセット作成者コードを発行してはいけない。
- データセット作成者は重複するデータセット識別コードを発行してはいけない（1つのデータセット識別コードは一度のみ発行できる）
- 同じデータセット識別コードを共有する中において個別 ID は重複してはいけない

これらのルールが守られる限り ID は唯一性・継続性が保証される。

この ID 体系のもっとも特徴的な部分は管理者を意味する ID ではなくデータセット作成者コードを含んでいる点である。管理者を意味する ID は ID が有効であるかぎり有効な意味（管理している者）をもつことが要求されるのに対し、データセット作成者コードはあくまでも発行した時点で有効であればその後の有効性は問われない。

以下、事業の移管を例にとって継続性の維持について考察する。

先ず図 7 のように組織 A がシステム A に管理番号を発行した場合について考える。この時点ではシステム A は組織 A によって管理されているとする。

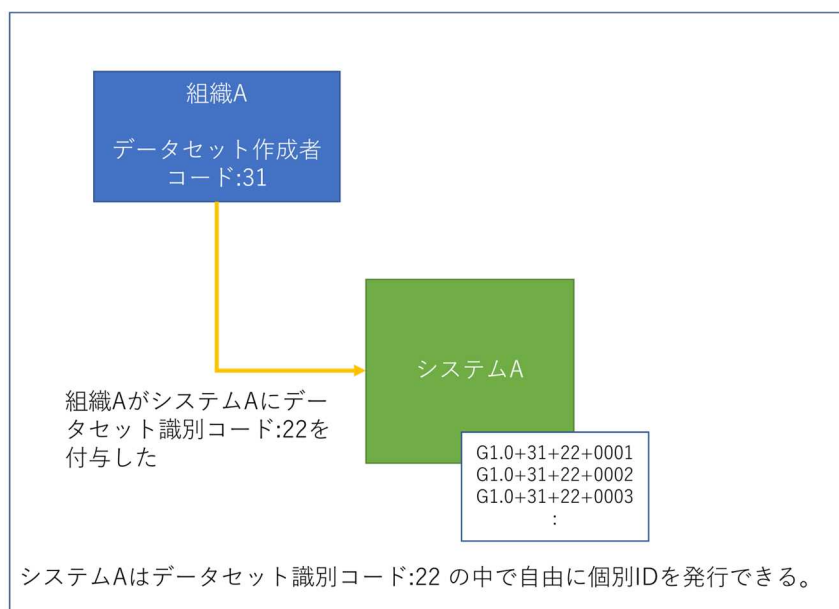


図 7 データセット識別コード及び個別 ID の発行

次にシステム A が組織 B に移管された場合の運用について考える。

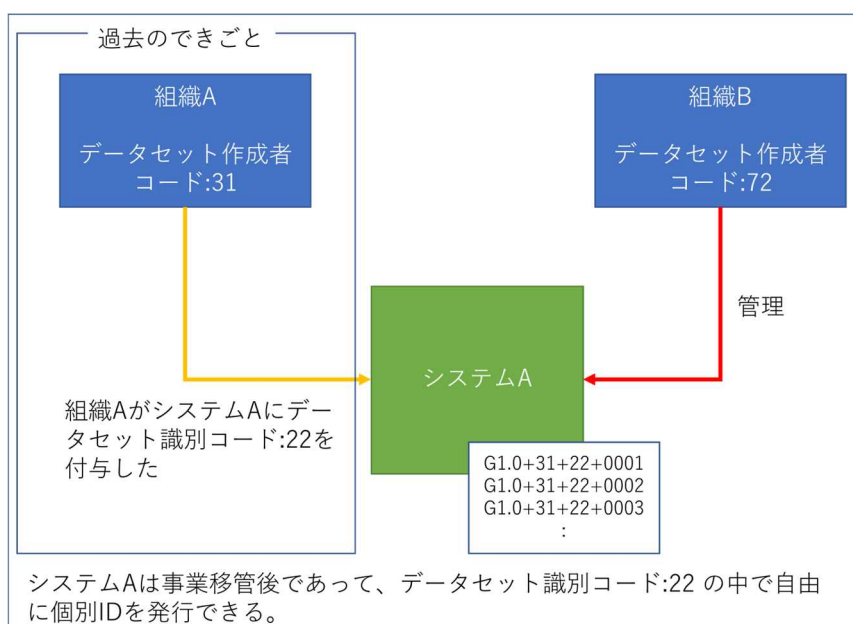


図 8 システムの移管

データセット識別コード:22 は組織 A によって発行されたものであるが、組織 A の存続の有無に関わらずデータセット識別コード:22 が組織 A によって発行されることが無いため、移管された後であってもシステム A は移管前と同様に ID を使い続けることができ、かつデータセット識別コード:22 の中で新しい ID を発行することができる。

注意点としては、データセット作成者コードとして法人番号を使用した場合である。技術的には、データセット作成者コードとして法人番号を利用した場合であっても作成者であり管理者でないため事業移管によって変更する必要はない。しかし、次のような理由によって ID が変更されてしまうことがないよう明確なルールを作成し周知する必要がある。

- 移管したのだから変更する必要があるという勘違い
- 他の組織の法人番号が含まれているのが気持ち悪い

6.2 データの独立のメリットと重要性

6.2.1 利便性向上のメリット

データの管理上の境界と利用者視点でのデータの利用の境界は必ずしも一致しない。データの管理は所管組織ごとによって行われることが一般的であるが、利用者視点からは所管組織に関係なく利用者の状況に応じて必要な情報の取得や必要な手続きを一括して行えるのが望ましい。

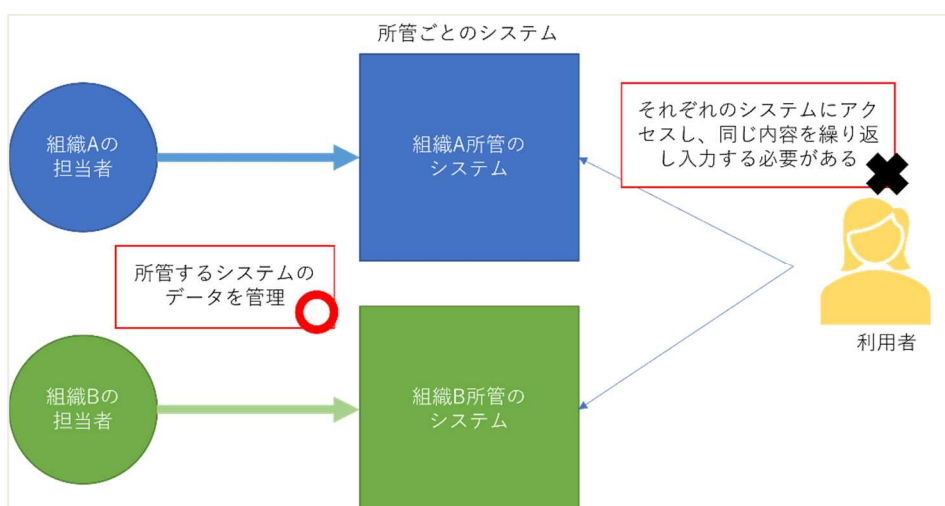


図9 データ管理上中心のモデル

しかし、データとシステムが一体になっている状況では、データの管理上の境界とシステムの境界を切り離すことができない。そのため、データ管理上の境界に合わせてシステムを構築した場合は、利用者の利便性が低くなってしまふ。

また、逆に利用者視点でシステムを構築した場合は、同じデータを複数のシステムに登録が必要が生じるなどデータ管理が煩雑なものになってしまう。

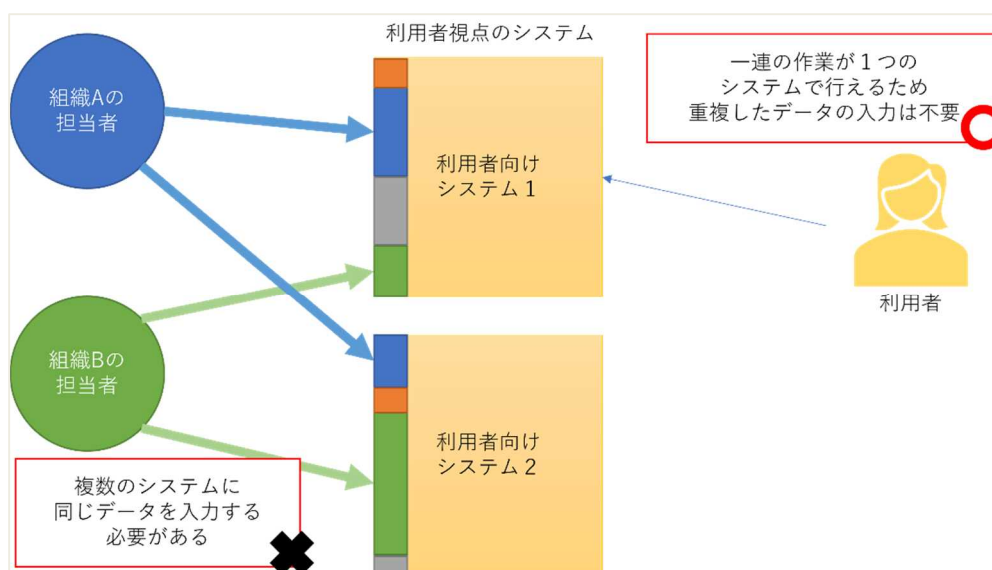


図 10 利用者視点モデル

このような問題を解決しデータ管理の容易さと利用者の利便性を両立するためにはデータをデータプラットフォームのような形でシステムから分離すること必要不可欠となる。

6.2.2 技術的なメリット

また、システム主体のままデータの相互利用を行うようすること、つまり、あるシステムが異なる複数のシステムによって管理されているデータにアクセスするようなシステムを構築することは、各システム間においてデータ形式の正規化などの変換を行う必要があるなど統合されるシステムの数に対し 2 次関数的にコストが増大してしまう。また、それぞれの認証システムへの対応が必要になることで実装が煩雑になってしまうなどの問題がある。

さらに、システムの終了やリニューアルがあった場合、システムが取得していたデータごと無くなったり、データ構造の変更によってデータ変換システムが動作しなくなったりするなどの影響が発生するなどの問題が生じる。

運用オペレーションにおいても、システム間の相互依存性が高くなってしまい慎重な運用を要求される。

このように、データを独立させることなくデータの相互利用を行うことは、コスト、運用オペレーションの両面から現実的ではない。

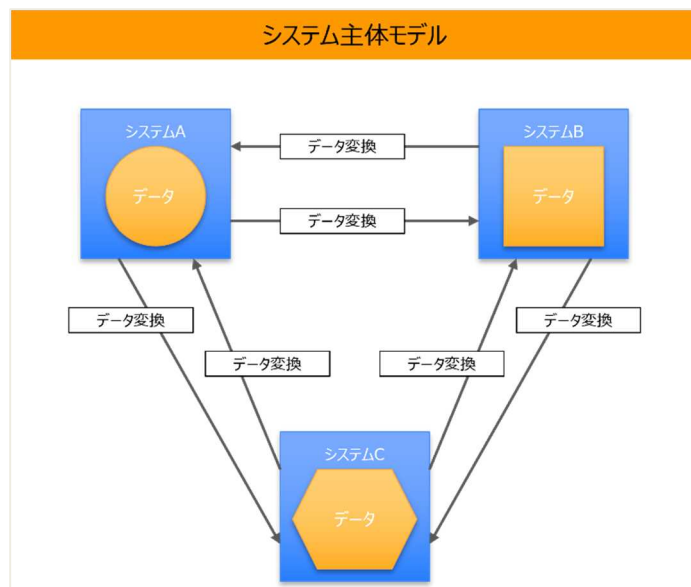


図 11 システム主体のモデル

このような技術的な問題についても、ID 体系やデータの構造・形式・書式などが一定のルールにしたがっていることや一回の認証で全てのデータにアクセス可能になるようなデータプラットフォームをシステムから独立した形で構築することで解決することができる。

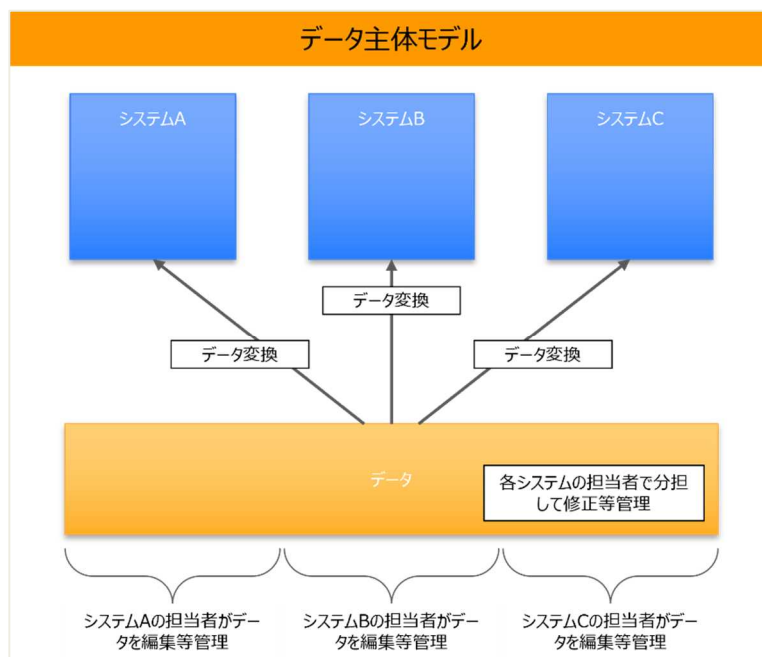


図 12 データ主体モデル

これらのことから、システムからデータを分離し、データ主体のモデルを構築することは実装、管理、利便性のいずれの観点からも大きなメリットがある。

6.2.3 将来的なメリット

6.2.3.1 オープンデータとしての民間利用

データをシステムから独立させることで、データの一部もしくは全部をオープンデータとして民間に開放することが容易にできるようになる。

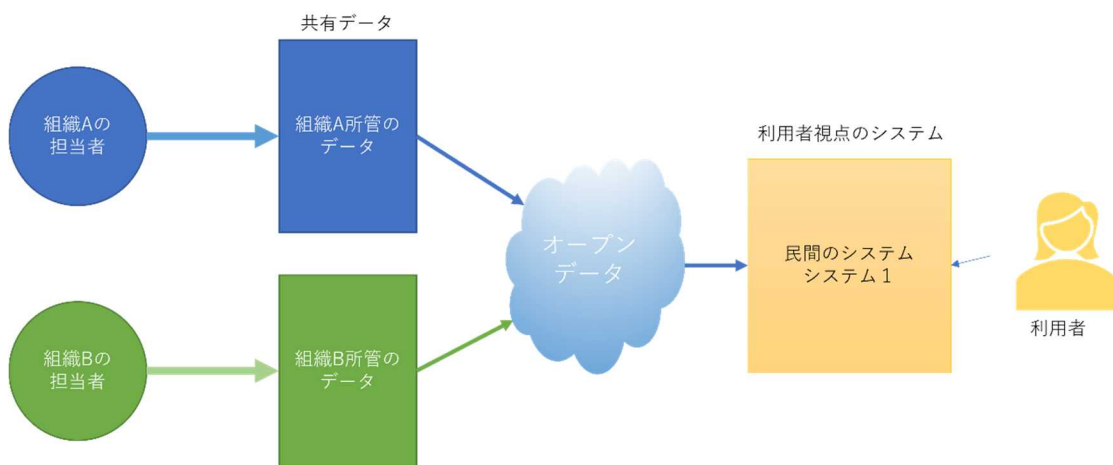


図 13 オープンデータとしての活用

6.2.3.2 ユーザーデータのシステム間共有

公共側からのデータだけでなく、利用者からのデータについてもシステムから独立して管理するようにすることで、手続きの際に同じ内容を繰り返し入力することや、同じ書類を複数回提出することが不要となり利便性の向上が望める（ワンスオンリーの実現）。

公共側からのデータとユーザーデータについてはデータの流れは逆向きであるが、システムとデータの関係は同じであり多くの基盤を共通に利用することが可能である。

ただし、ユーザーの情報を共有するためには、利用に対するユーザーの承諾の管理等が必要となるなど追加の機能実装が必要になる。

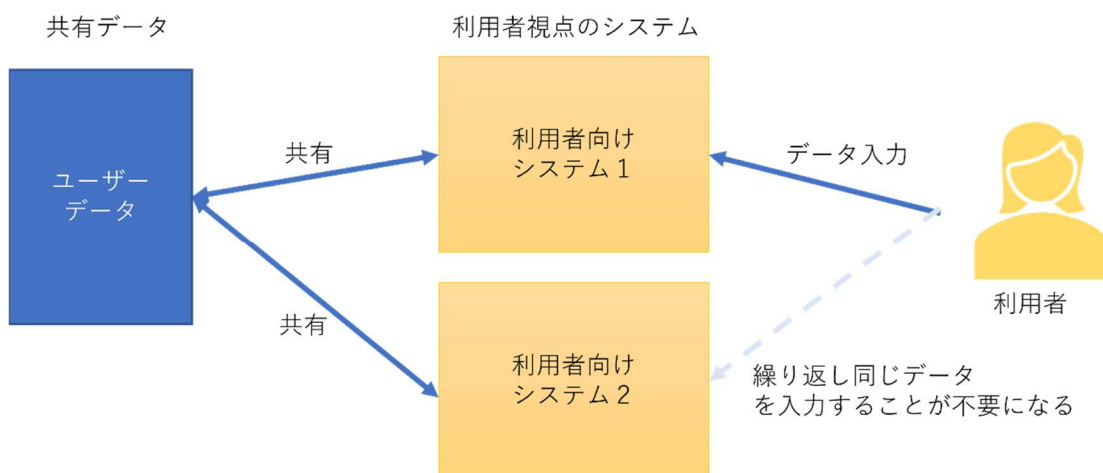


図 14 ユーザーデータのシステム間共有

7 まとめ

6 システムの調査及び調査を踏まえた技術面、運用面からの検討から、各システムからデータを分離し、データをシステムから独立したものとして、公共機関やデータの内容によっては民間も含めて皆が活用できるようにすることが最も重要である。このデータの独立によって次のようなメリットが得られる。

- 将来的な運用コストの削減
- 利用者視点でのシステムの新規構築
- 利用者の利便性を向上するシステムの改修
- オープンデータとして民間へのデータ提供

具体的には、ID 体系、データの構造、形式、書式が統一されたデータを（論理的に）一括管理できるようなデータプラットフォームの構築が期待される。

8 参考資料

8.1 政府相互運用性フレームワーク（GIF）

https://www.digital.go.jp/policies/data_strategy_government_interoperability_framework/