X. 教育分野におけるデータ仲介機能を用いたデータ連携基盤仕様案

本章では、令和３年度委託調査に基づく本年度調査「データ仲介機能を用いたデータ連携基盤仕様案」について、教育分野におけるデータ連携基盤のための仕様案の検討を行う。検討方法については、教育分野においてデータ連携基盤によって解決が期待される問題を定義し、データ仲介機能およびデータ連携基盤の仕様を明らかにした上で、データ連携基盤の導入時の課題を要件として示し、この対応方法としての仕様案を整理する。

なお、仕様案は記載時点の情報から検討されたものであり、現在並行する関連検討の進捗を踏まえた再整理を必要とするものである。

X.1 教育分野におけるデータ連携に関する検討状況の整理

デジタル社会の実現に向けた重点計画[1]において、データ連携に関する取り組みは「包括的データ戦略[2]に関する具体的な施策」として整理されている。包括的データ戦略において、データ連携基盤は「ツール」と位置づけられており、「利活用環境」と「データ連携に必要なルール」と合わせて「『データ連携』とそれを『利活用したサービスを提供』する基盤（プラットフォーム）」と位置づけられる。また、「各分野のプラットフォーム構築では、アーキテクチャを設計した上で、データ連携、データの標準／品質に関するルールを個別に整備している必要がある。」とされる。

上記を踏まえ、教育分野のプラットフォームを構成する以下の3項目について現在の検討状況を整理する。

- データ連携基盤（ツール）の検討および開発動向
- 教育分野のデータ利活用環境
- 教育分野のデータ連携に必要なルール

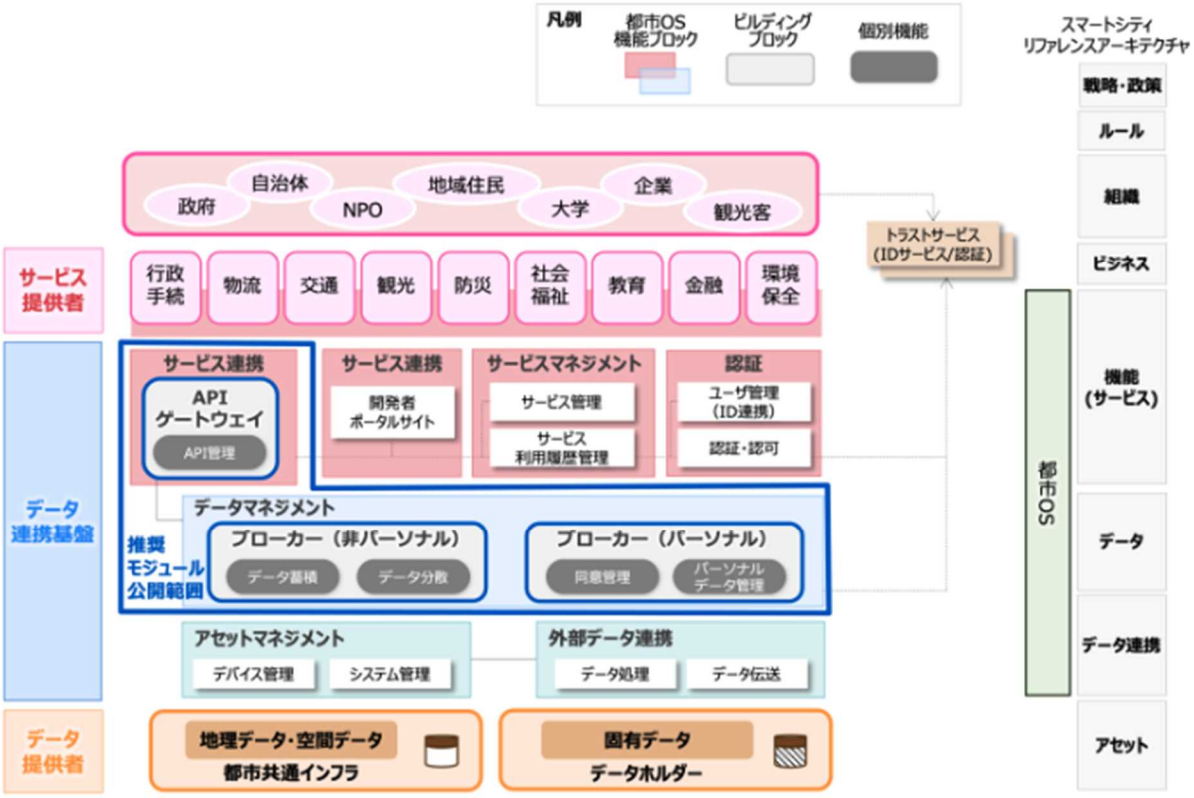
X.1.1 データ連携基盤（ツール）の検討および開発動向

X.1.1.1 データ仲介機能の開発状況

データ連携基盤の機能要件に関する検討は、パーソナルデータを取り扱う生活用データ連携[3]とパーソナルデータを取り扱わない産業用データ連携[4]により行われ、この実装がそれぞ

[1] 2022年6月7日 閣議決定
[2] 2021年6月18日 閣議決定
[3] 日本電気株式会社, 「生活用データ連携に関する機能等に係る調査研究 調査報告書」, 2022/3/3
[4] EYストラテジー・アンド・コンサルティング株式会社, 「産業用データ連携に関する機能及び実装等に係る調査研究 報告書」, 2022/3/30

れ推奨モジュールとして公開されている[5]。この推奨モジュールには、データ連携基盤におけるデータマネジメント機能の中核となる　データ仲介機能（ブローカー）が含まれる。



図x.　データ連携基盤全体像[6]

データ仲介機能は、提供元となるデータへの接続権限を持ち、APIを通じて行われるデータの提供要求に対して要求元の認証やデータ提供の認可などを外部機能と連携して行い、提供先に対してデータ送信を行う。また、データ仲介機能は、パーソナルデータの取扱要否で2種類の要件および推奨モジュールの実装が存在する。

X.1.1.2 パーソナルデータ連携モジュール

教育分野においては個人情報の取り扱いが前提となるため、ここでは本人同意に基づいて第三者提供の制御が可能なデータ仲介機能である「ブローカー（パーソナル）」についての動向を記載する。

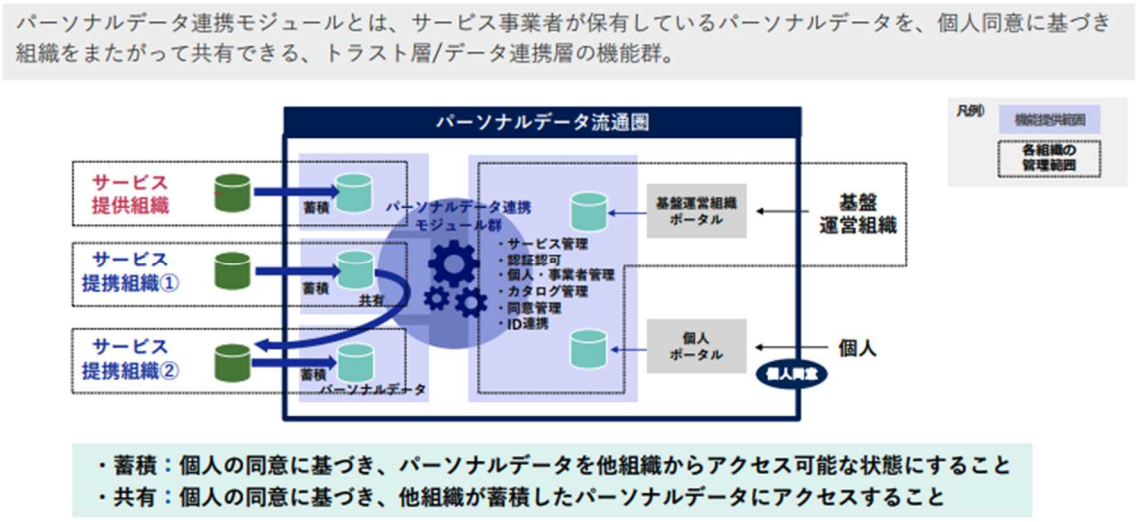
ブローカー（パーソナル）の主な仕様としては、蓄積と共有の2段階の本人同意により、データの構造をデータカテゴリ（例：健康診断データ）、データセット（例：心電図データ）、データ実体（例：2000/3/2の心電図）といった段階に分けて第三者提供を制御することを規定する。この実装として、「パーソナルデータ連携モジュール」[7]が公開されている。

---

[5] https://data-society-alliance.org/area-data/module/manual/

[6] https://data-society-alliance.org/area-data/module/

[7] https://data-society-alliance.org/area-data/module/manual/#personal



図x. パーソナルデータ連携モジュール[8]

本検討では、パーソナル連携モジュールの開発元である日本電気株式会社にブローカー（パーソナル）の要件外の実装仕様に関してインタビューを行った。当該モジュールは、複数の自治体等の事業者をマルチテナントで対応することができるが、複数のモジュール間における相互連携については対応外であることがわかった。また、データセットが蓄積用データベースのレコードに対応しており、データセット単位の制御を行う際は、パーソナルデータ連携モジュール外のデータベースに直接連携することは現時点で未対応であった。

X.1.1.3 とっとり「人づくりDX」構想

本検討では、データ連携基盤の運用事例である鳥取県の「とっとり『人づくりDX』構想」[9]にインタビューを行った。鳥取県では、Googleアカウントを全公立校で利用しており、Google Workspaceをベースにした LMS[10]構築を目指している。本事例では、児童生徒の健康観察情報のリアルタイム共有といったデータ連携で成果が認められている。

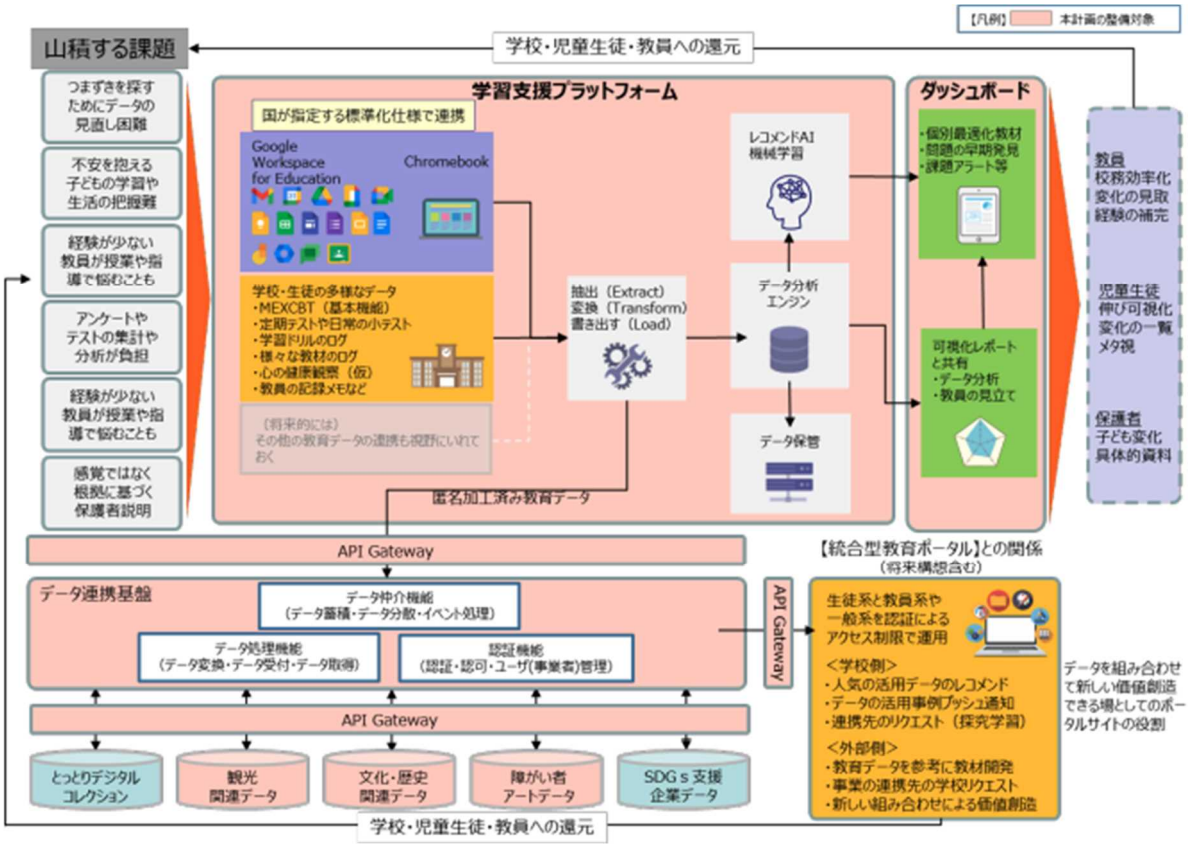
当該事例のデータ連携の課題に関するコメントとして、県外の異なるデータ連携基盤と連携の際に、県内外のデータ形式の不一致や認証連携の実現方法が未知であることが示唆された。

---

[8] 日本電気株式会社, 「パーソナルデータ連携モジュール 説明資料」, 2023/2/2

[9] https://www.chisou.go.jp/sousei/about/mirai/pdf/denenkouhukin_saitaku_type23saitakujirei.pdf

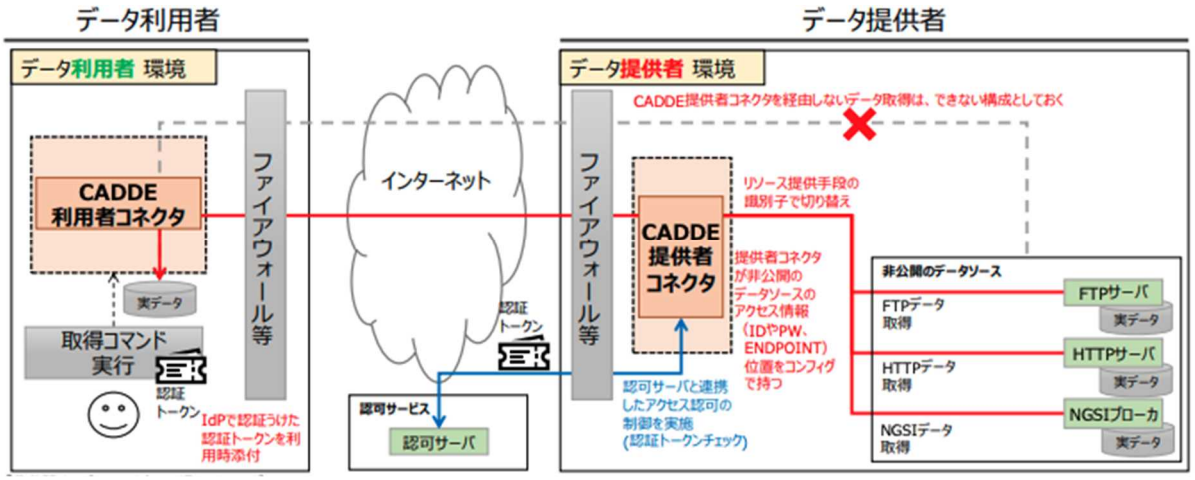
[10] Learning Management System



図x. 「とっとり『人づくりDX』構想」学習支援プラットフォーム概要図

X.1.1.4 SIP分野間データ連携基盤

データ連携基盤の発展的な研究として、SIP[11]第２期において、異なる分野間でデータの発見と利用ができる仕組みをCADDE[12]として提案されている。CADDEは、複数のデータ提供者のデータカタログを収集し、データ利用者に対して「データカタログ分野横断検索サービス」を提供する。アクセス制御方法については、「コネクタ」を通じて認証認可を行うこととしている。



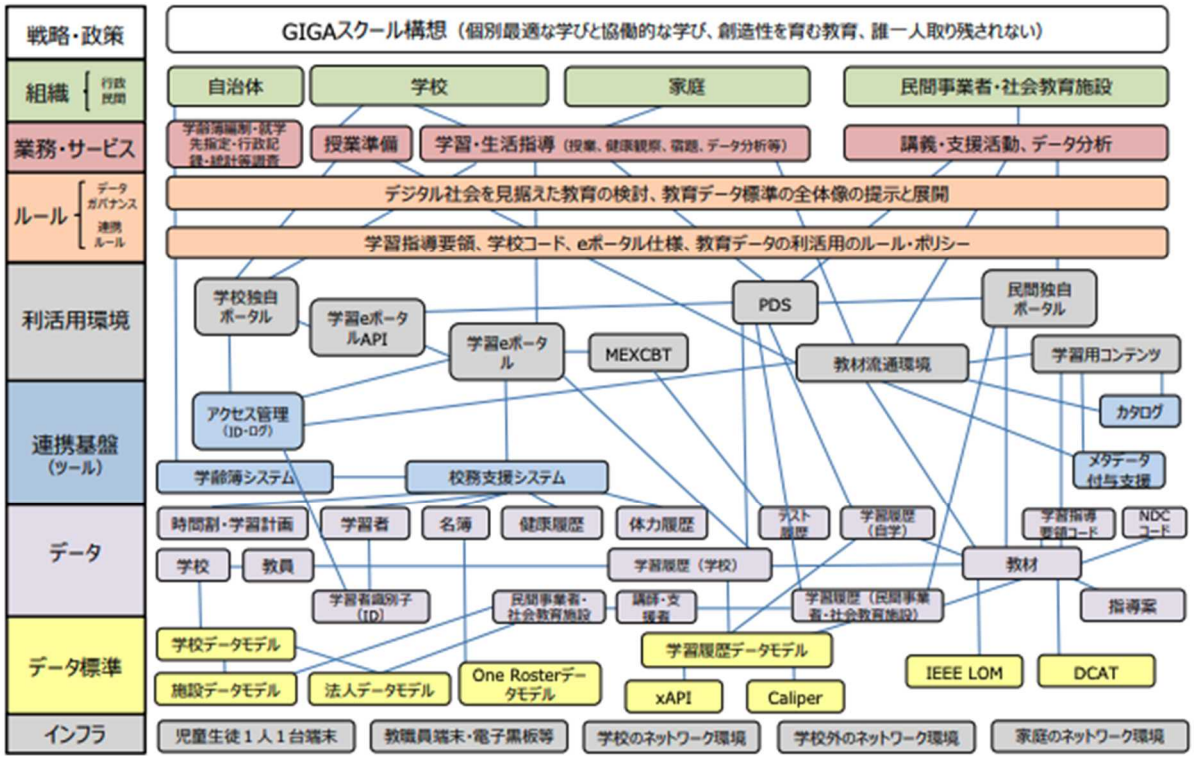
[11] 総合科学技術・イノベーション会議の戦略的イノベーション創出プログラム

[12] Connector Architecture for decentralized Data Exchange

図x. CADDEによる非公開データの連携[13]

X.1.2 教育分野のデータ利活用環境

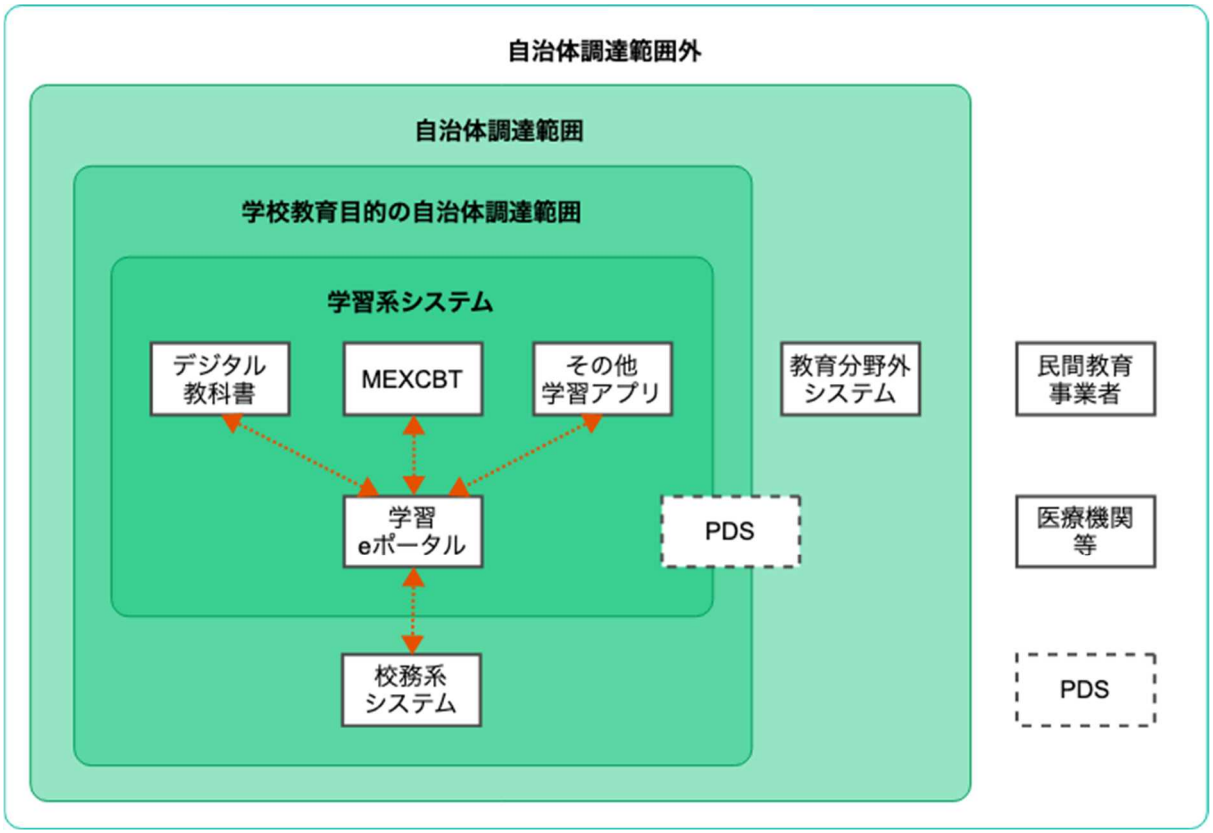
教育データ利活用ロードマップ[14]における教育分野のデータ利活用環境は、主として既存の教育ICTシステムに一致することから、初中等を中心に現在運用または議論される主要な利活用環境をシステム種別を以下に記載し、関係性を整理する。



図x. 初中等の教育データの蓄積と流通の将来イメージ
（教育データ利活用ロードマップより抜粋）

- 校務系システム
- 学習系システム
- MEXCBT
- 学習eポータル
- 民間教育事業者システム
- 教育分野外システム
- PDS（Personal Data Store）

[13] SIPデータカタログ項目仕様V2.0ガイドライン(2022年3月31日版)
[14] デジタル庁，総務省，文科省，経産省，「教育データ利活用ロードマップ」，2022/1/7



図x. 教育分野のデータ管理システム

X.1.2.1 校務系システムと学習系システム

校務系システムとは、主に職員室の事務処理で用いるシステムであり、出欠、成績、授業時数、保健情報等を取り扱う。これらの取り扱いデータ形式についてAPPLIC[15]により標準化がされている。学習系システムとは、デジタル教科書に代表されるような主に授業時間に用いられるシステムであり、主に児童生徒の学習を目的とする。学習システムのデータ構造は、教材開発会社ごとに異なり、現時点で統一化はされていないが、本事業で推進されるxAPIによるスタディ・ログとしての標準化が検討されている。文科省が運営するCBTシステムであるMEXCBTもxAPIによるログ出力を行う学習系システムの一つである。

X.1.2.2 学習eポータルと学習系システム

学習eポータルの現在の実態として、連携システムへのSSO[16]やMEXCBT等の各学習システムから出力されるxAPIのstatementの収集するLRS[17]を具備する。そのため、本検討では学習eポータルを学習システムが提供するスタディ・ログ（教育データ）[18]のデータ集積基盤とみなす。

---

[15] https://www.applic.or.jp/

[16] Single Sign On, 連携する複数のシステムへの認証を統合する機能

[17] Learning Record Store

[18] 児童生徒や教職員等による「主体情報」、学習内容を示す「内容情報」、学習行動のみに限らず関連行動をう組む「活動情報」により構成される行動ログ

X.1.2.3 民間教育事業者システム

民間教育事業者システムは、学校以外の教育機関である塾・予備校・各種通信教育などが該当する。校務系および学習系システムについては主に自治体が管理するため、管理データも自治体保有情報となるのに対し、民間教育事業者システムは管理主体が民間事業者となる。

X.1.2.4 教育分野外システム

教育分野外システムとは、自治体内の教育用途以外のシステムや医療機関等の公共性の高いシステム、または一般的な商用アプリケーションなどのデータ連携基盤を利用する可能性のあるシステム群を意味する。例えば、地域医療連携の一種として学校と病院の連携[19]などが想定される。

X.1.2.5 PDS（Personal Data Store）

PDSとは、「他者保有データの集約を含め、個人が自らの意思で自らのデータを蓄積・管理するための仕組み（システム）であって、 第三者への提供に係る制御機能（移管を含む）を有するもの」[20]と定義されるため、本人同意に基づいてパーソナルデータの第三者提供を行うシステム全般が該当する。このため、本人同意に基づいてパーソナルデータを制御するデータ連携基盤も広義のPDSと位置づけることができる。

X.1.3 教育分野のデータ連携に必要なルール

教育データ利活用ロードマップでは、データ連携に関する以下のルールが明示されている。

1. 学習履歴を含めた個人の教育データは、学校や自治体、民間事業者といった主体（関係者）ごとに「分散管理を基本」とする。
2. 機関間の個人情報等の連携は、法令に基づく場合等を除き、原則として本人の同意により提供する。

上記のルールは、学校や自治体、民間事業者に分散管理されたデータの連携実現のための認証連携と本人同意情報の連携が要件となる。

X.1.3.1 自治体や校種を跨ぐ校務情報やスタディ・ログのデータ連携

自治体や校種を跨ぐ転校を行う際、校務系システムの校務情報や学習eポータルのスタディ・ログを共有・連携する方法は確立していない。実現に向けた具体的な課題として、自治体や私学などの学校法人、またはシステムベンダーなどを跨ることによる以下の2点が上げられる。

---

[19] https://www.mhlw.go.jp/kokoro/parent/consultation/way/way_02.html

[20] 官民データ活用推進基本計画実行委員会，「データ流通・活用ワーキンググループ第二次とりまとめ」，2019/6

- 児童生徒の本人識別情報の共有問題
- 要求元の認証問題

児童生徒の本人識別情報とは、いわゆる本人識別ID等で、従来の校務系システムでは学籍番号等の形で所属機関ごとに発行し、外部共有を目的としなかったため共有ルールが存在していない。また、学習eポータルにおいても、UUIDによる共通フォーマットが利用されたが、厳密なユニーク性や学習eポータル事業者間のデータ連携ルールが定まっていない。

要求元の認証問題とは、外部のデータ連携基盤に所属する事業者からデータ提供要求を受けた際に自身の所属するデータ連携基盤で認証ができないデータ提供要求者をどのように認証するか、という問題である。

X.1.3.2 学校保有情報と民間保有情報の相互活用

公教育機関と学習塾や予備校等の民間教育機関がデータ連携を行う際、個人情報の取り扱いに関する課題が存在する。異なる個人情報取扱事業者のデータを利用する際は、それぞれの個人情報取扱事業者が第三者提供に対する本人同意を得る必要がある。　本人同意に在り方については、前段の自治体連携などで複数のシステムを跨る際の同意情報の管理方法に関する課題や段階的同意[21]の技術的実現方法などが課題となる。

本件に関して、ブローカー（パーソナル）の開発を行った日本電気株式会社にヒアリングを行ったところ、本人同意の取り扱いは基本的には包括同意であり、データ実体内のデータ項目ごとや連携方法の種類に応じた段階的同意には対応していないことが確認された。また、現行実装はデータ連携基盤内にデータセット[22]を収集する前提となっており、データカタログ[23]のみを取り扱う設計ではないことについても確認された。

X.2 データ連携基盤に対する仕様案

教育分野におけるデータ連携基盤の取り扱いデータは、主として児童生徒の個人情報であり本人同意管理が必要とされるから、データ連携基盤の種別としては「生活用データ連携に関する機能等に係る調査研究」で検討されたブローカー（パーソナル）の実装仕様に近い。

ここでは、上記調査研究およびブローカー（パーソナル）の実装仕様の利用を前提に教育分野におけるデータ連携基盤の要件を検討する。

X.2.1 データ連携基盤間の構成要件

教育分野のデータ連携基盤については、教育データ利活用ロードマップにより「データの分散管理を基本」とし「国が一元的にこどもの情報を管理するデータベースを構築することは考えていない」ことを前提にし、データ連携基盤自体も分散型を前提とする。その上で分散型の構成を階層型か非階層型か、といった構成要件を考える。

---

[21] 森田瑞樹，「患者中心の情報管理とそれを可能にする新しいインフォームドコンセント」，2014/1/24

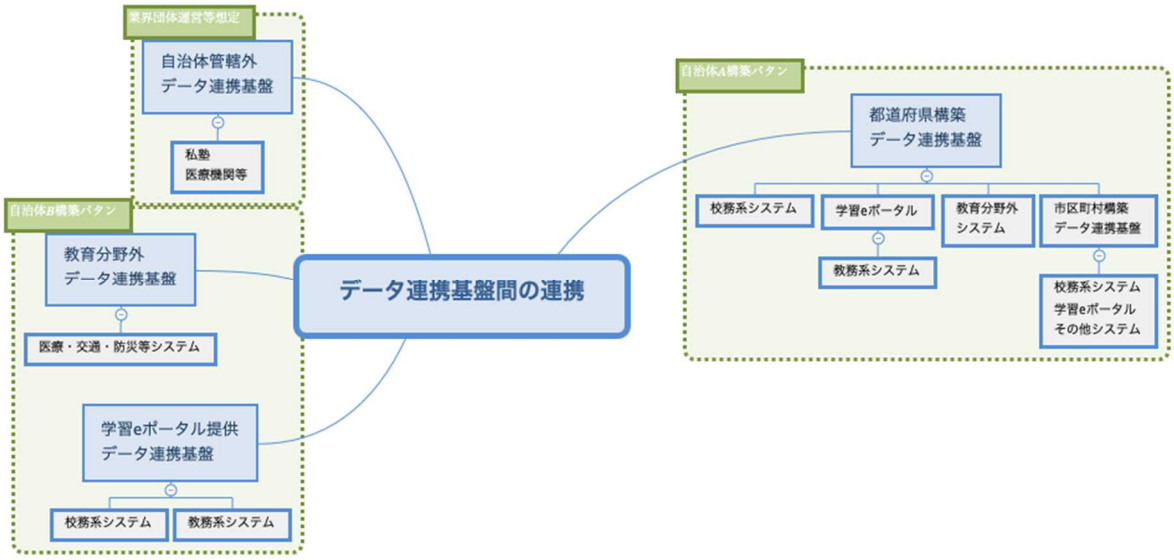
[22] 複数のデータから構成されるデータの集合

[23] データセットを対象に記述されたメタデータの値の塊

自治体によってはデータ連携基盤のデータセットに健康・医療・介護分野などの隣接分野を含めたい場合を考慮する必要があることを踏まえ、自治体単位で統制された複数のデータ連携基盤を有することとする。階層型か非階層型かに関しては、システム規模としてある程度の集約利用するものであることから、都道府県や一部の広域自治体などの単位を基本形とし、データ連携基盤の運営主体の裁量で更に階層化するかどうかを決定できるような構成とする。

構築例としては、以下のような自由度があることを保証することを要件とする。

- 自治体において教育分野に限らず横断的に利用可能なデータ連携基盤を構築する
- 自治体において教育分野に限定して学習eポータルなどの外部ベンダーが提供しうるソリューションを導入する
- 自治体において県と広域自治体を階層化または同列化して提供する
- 自治体において分野単位で複数のデータ連携基盤を提供する
- 自治体外の民間等が運営するデータ連携基盤の存在を許容する。



図x. 複数のデータ連携基盤間の関係

X.2.2 データ連携基盤間の識別子連携

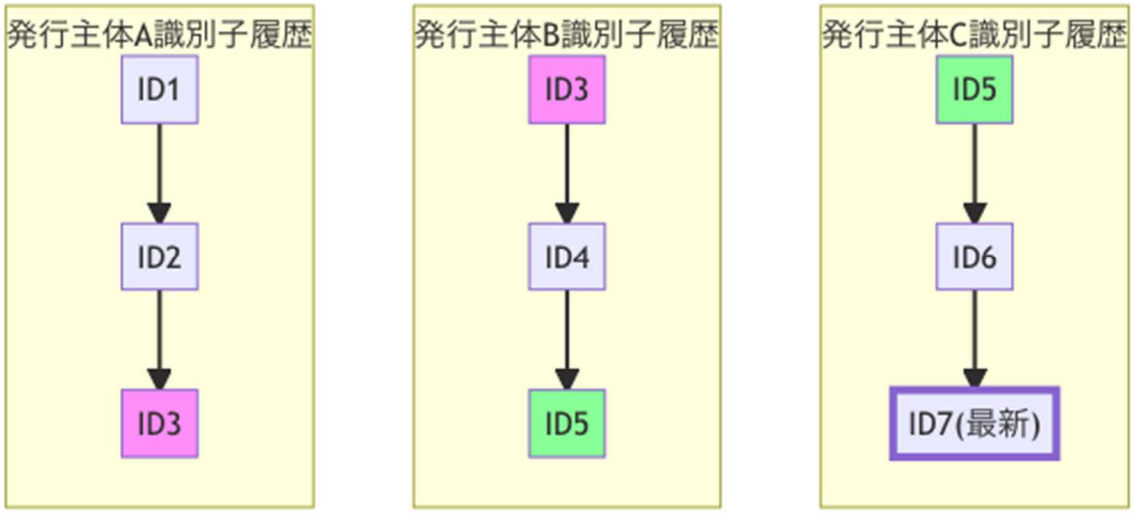
データ連携基盤が分散型である場合、データ連携基盤ごとに児童生徒に異なる識別子が発行されることを考慮する必要がある。例えば、進級・進学・転校などによって識別子が変更される場合、過去の自身の識別子を取得可能な仕様が必要となる。

また、場合によっては過去のデータ連携基盤のデータを連携させたくない場合については、本人同意に基づいてこれを拒否することも可能であることが望ましい。この解決のためには、データ連携基盤を横断して個人を識別する識別子の管理運用を要する。

ここでは、以下のような仕様を前提に考える。

- 児童生徒の識別子は、各発行主体が任意に発行可能とする。

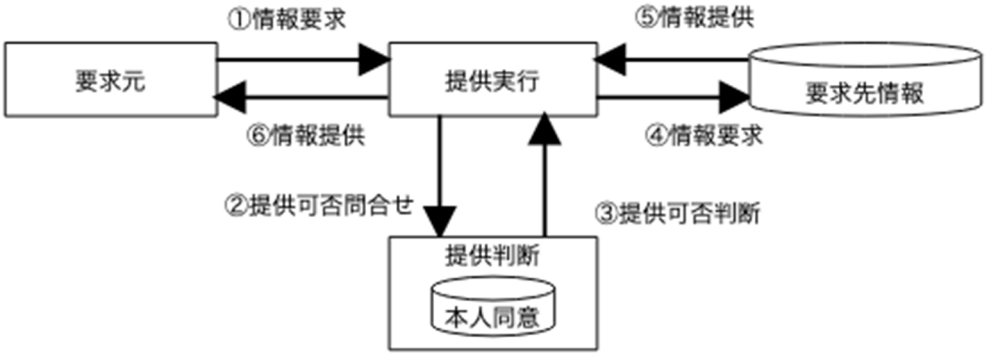
- 児童生徒の各識別子は、複数の連携対象となるデータ連携基盤全体で常に一意であることが保証される。
- 発行主体となったシステムは、同一個人に対する過去の識別子の履歴を発行主体の識別子と共に管理する。
- 新しい発行主体は、過去の発行主体の最後の識別子を運用によって引き継ぐ



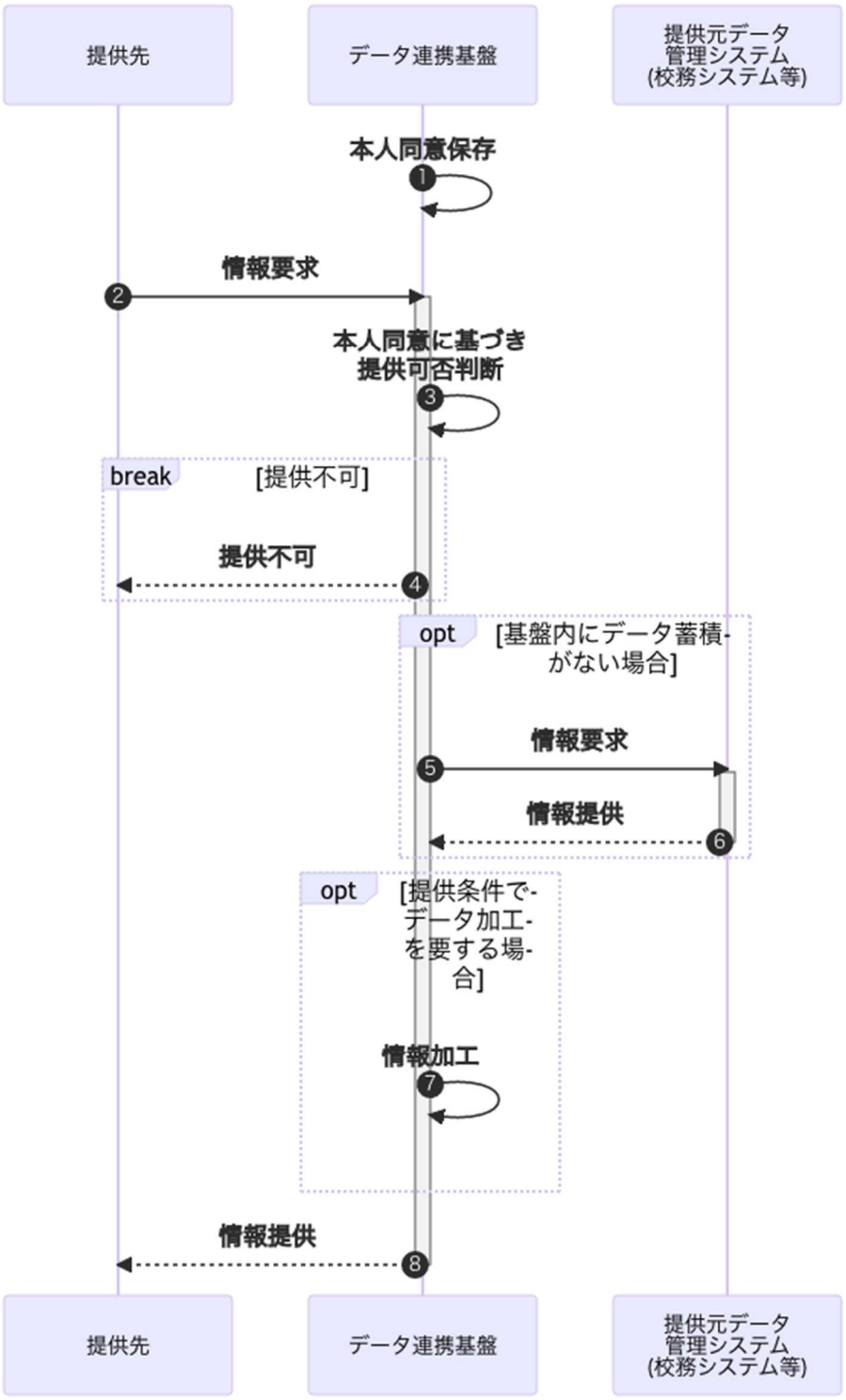
図x. 発行主体の識別子管理

X.2.3 本人同意に基づくアクセス制御の実現要件

教育分野におけるデータ連携基盤は、提供データの制御のためにブローカー（パーソナル）の機能の一部として本人同意情報に基づくアクセス制御の機能を必要とする。下図のアクセス制御機能を一般化したモデルにおいて、要求先情報を管理する各システムが、本人同意に基づいた提供実行機能を担当することが理想できてあるが、各システムの実装負担が現実的ではないため、ここではデータ連携基盤が提供判断と提供実行機能を担当する仕様を検討する。



図x. 一般的なアクセス制御モデル



図x. アクセス制御におけるデータ連携基盤の役割

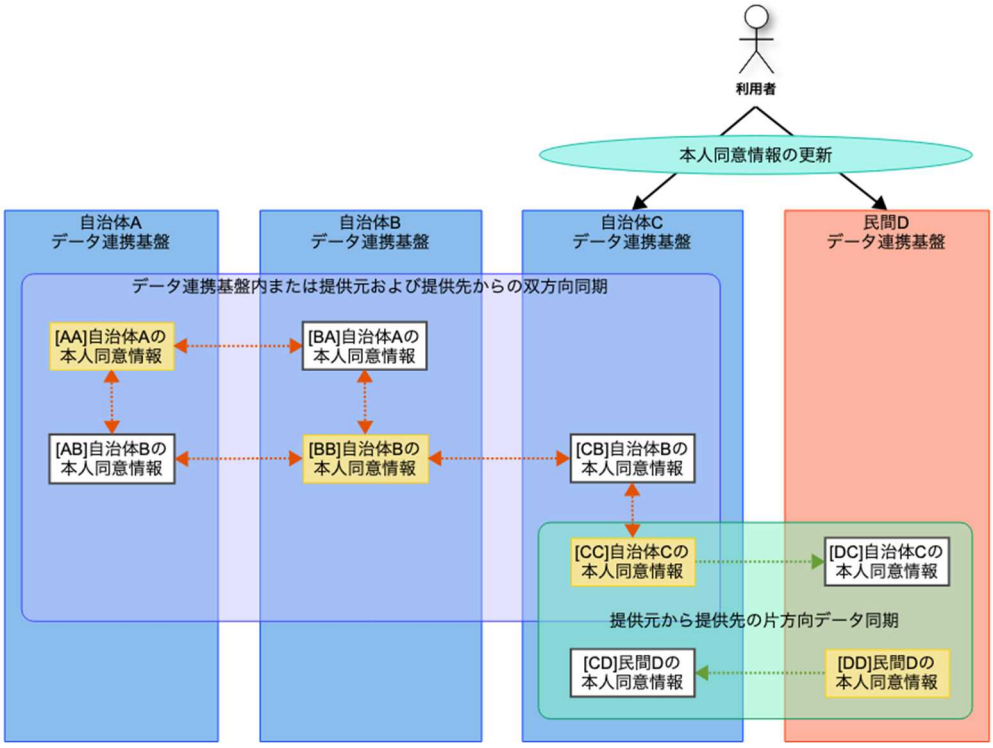
データ連携基盤間の本人同意情報の取扱については、次の前提とする。

- 各データ連携基盤の運営主体は、個人情報取扱事業者であり、本人同意情報の取得と管理の責任を有する。

- データ連携基盤間でデータ授受が発生する際、本人同意情報の取得責任は提供先のデータ連携基盤が有するが、提供元のデータ連携基盤が本人同意の取得を代行できる[24]。
- 他データ連携基盤に対して第三者提供可能な本人同意が取得済みでかつ目的の範囲内であれば、提供先が増えるごとの再取得は不要。

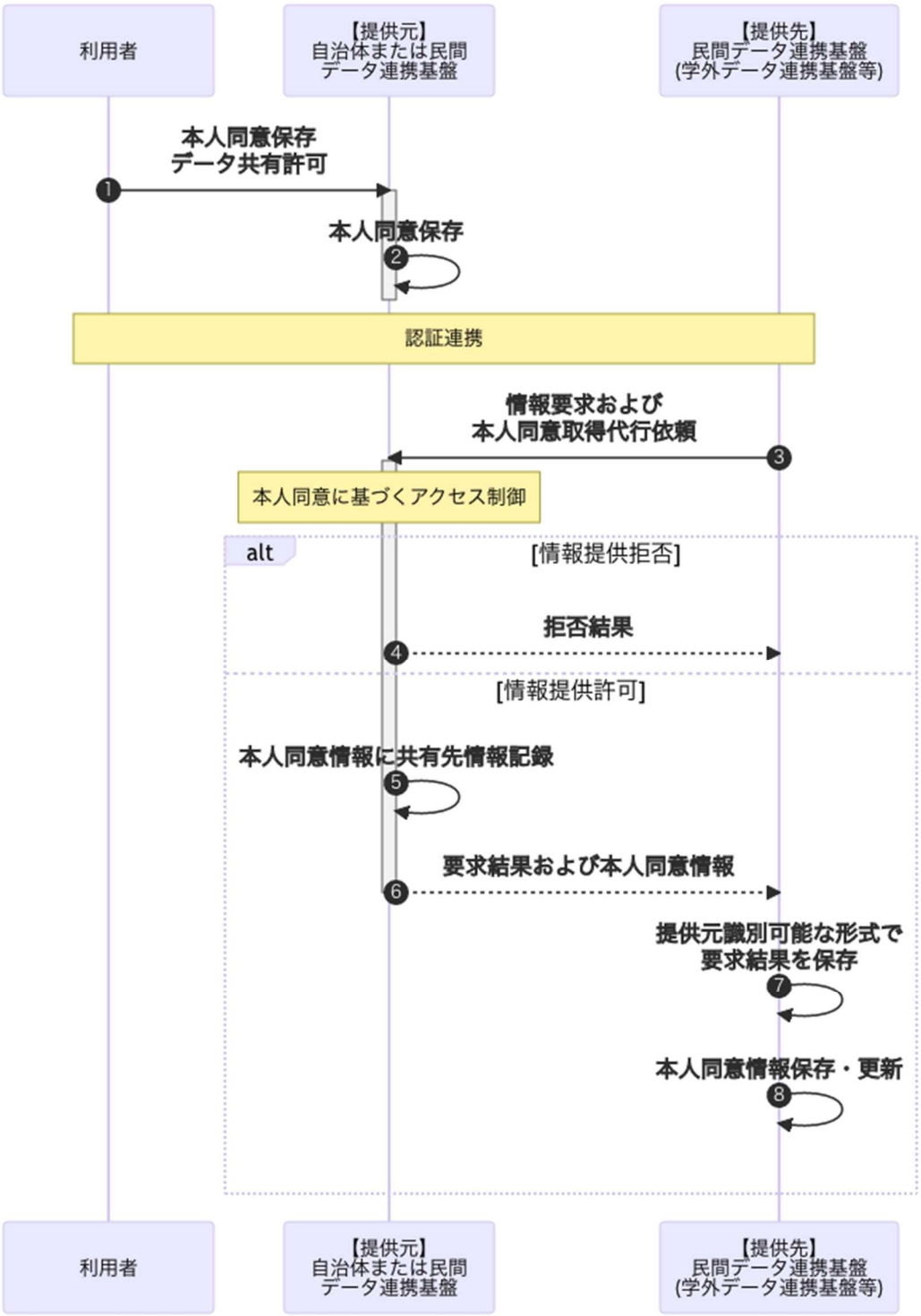
本人同意情報をデータ連携基盤で管理するとした場合、転校等のケースにおいては、異動前後で本人同意情報の内容の変更がない場合、再同意取得のコストや実施漏れを防ぐため本人同意情報はデータ連携基盤間で同期可能とし、一方のデータ連携基盤からの本人同意情報の更新時は、連携対象の全てのデータ連携基盤に対して一括変更できることが望ましい。民間教育サービスの場合などは、サービス単位で独立運用できることが望ましい。

独立管理と一括管理の利用方法を両立させるためには、本人同意情報はデータ連携基盤単位の分散管理を行いつつ、個人単位でデータ連携基盤間の認証連携を契機にデータ同期を選択可能とする必要がある。データ同期を行う場合、データ連携基盤内の本人同意情報に対して、同じ同意条件を設定し、異動先のデータ連携基盤に対しても異動元の本人同意情報が引き継がれる。これにより任意のデータ連携基盤上で更新した本人同意情報が各データ連携基盤で共有される。



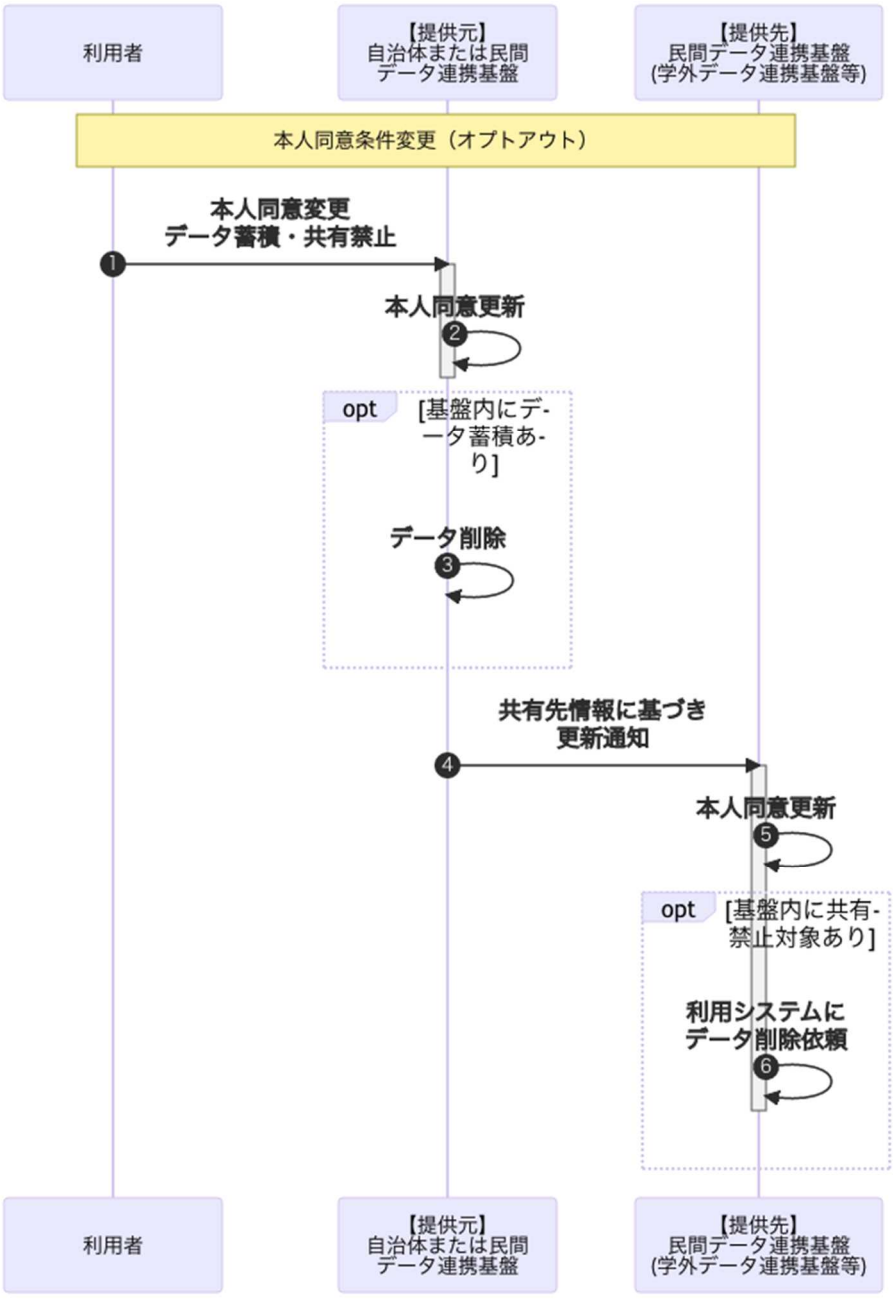
図x. 本人同意情報の同期例

[24] 個人情報保護委員会，「個人情報の保護に関する法律についてのガイドライン（通則編）3-7-2-2 同意を取得する主体」，2022/9一部改定



図x. 本人同意情報のデータ連携基盤間共有

また、本人同意情報は、複数のデータ連携基盤間で同期されていた場合、特定のデータ基盤から個人情報の削除を促す際にも利用できる。例えば、第三者提供の本人同意が失効されたことを検知したタイミングで必要に応じてデータ削除を行わなくてはならない。そのためデータ提供先のシステムは、第三者提供によって取得されたデータを本人同意の変更によって削除可能でなければならない。



図x. 本人同意変更を契機としたデータ削除

X.2.4 データカタログとデータセットの管理要件

データ連携基盤でデータセットを管理するべきかどうかを考える。データ連携基盤上でデータセットを管理運用することは集計や分析に有利であるが、データ連携基盤が複数の異なる運用システムのデータのアグリゲーションのみを行う場合は、データセットの実体コピーを行う必要があり、コスト的に不利となる。

例えば校務系システムは、外部からの情報参照による性能への影響が許容されない場合やデータ量が限定的である場合、データ連携基盤上のデータ複製が望ましい場合がある。スタディ・ログの場合、データが巨大であり管理システム自体がデータ連携基盤への連携が容易な場合が多く、データ複製は合理的ではない。教材コンテンツ等のコンテンツデータである場合、一般的な検索エンジンシステムと同様に実体の複製は持たずに検索用のメタデータと

してデータカタログのみを持つことが合理的となる。その他のデータカタログのみを用いる優位性としては、対象データセットの管理権限をもつシステムのデータカタログのみを取り扱うことで、データセットの編集時もデータカタログは一貫性を保つことができること、不必要な実体のコピーによる情報漏えいリスクに対策できることなどが挙げられる。

このことより、データ連携基盤は管理権限を持つシステム下のデータカタログのみでも運営できる要件を必要とする。

X.3 ユースケースの検討

ここでは、前段のデータ連携基盤に対する仕様案を用いて、具体的なデータ連携のユースケースについて検討する。ユースケースは記載順に次のデータ連携類型を意図する。

- データ連携基盤が異なる自治体から別自治体へのデータ連携
- データ連携基盤が異なる民間から自治体へのデータ連携
- データ連携基盤が異なる自治体から民間へのデータ連携

X.3.1 転校や進学による異動

データ連携基盤がそれぞれ異なるある自治体から別自治体へのデータ連携の例として、転校と進学に伴う校務系情報およびスタディ・ログの引き継ぎや問い合わせの解決方法を検討する。

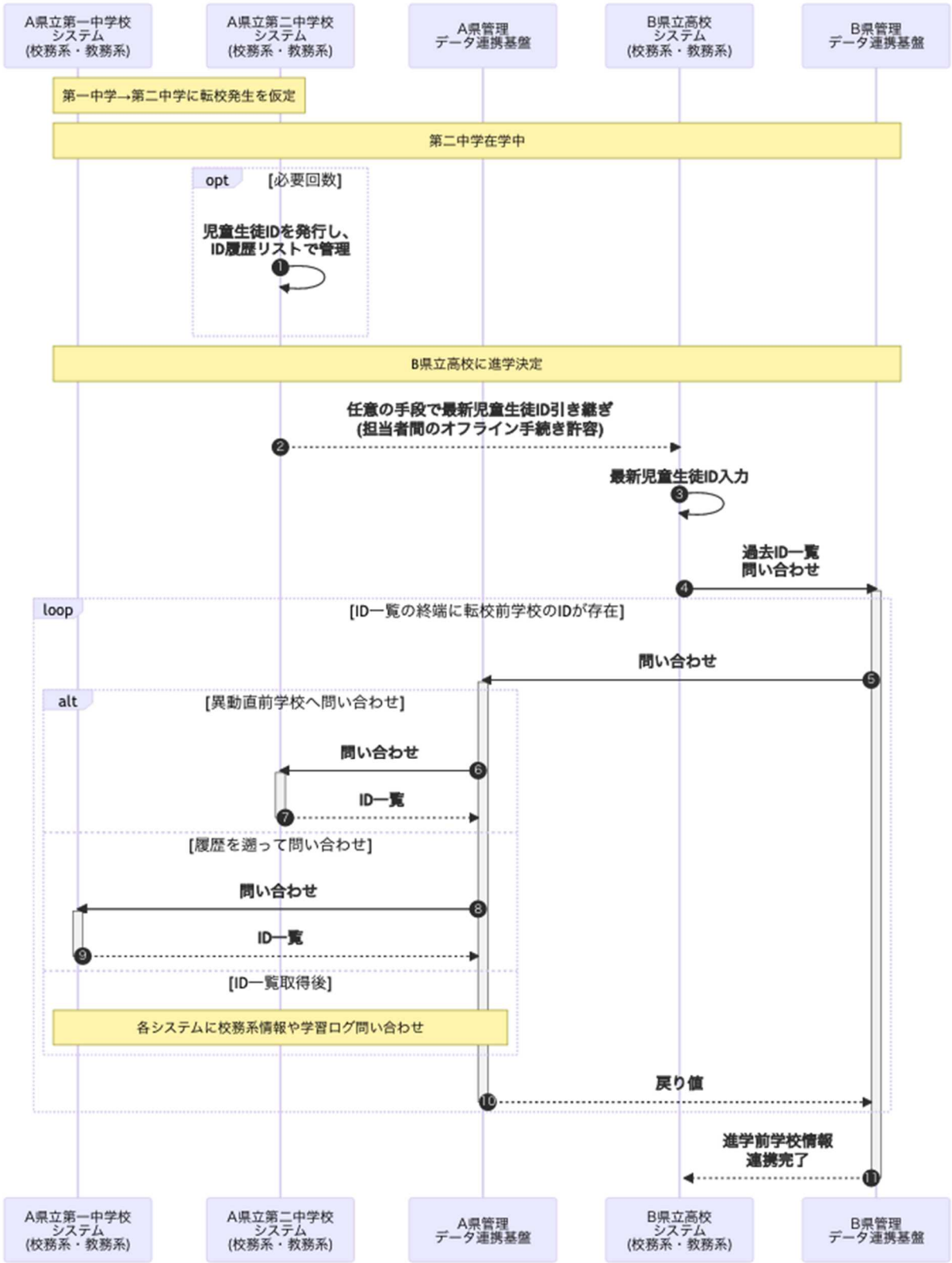
X.3.1.1 データ連携基盤間のトラスフレームワークによる要求元認証

データ連携基盤が異なる要求元の認証は、データ連携基盤同士のトラストフレームワークを形成し、互いを信頼し合うことができれば、相手側データ連携基盤の要求元の認証結果を受け入れることで解決が可能である。トラストフレームワークのための認証基準については、各データ連携基盤の代表団体によって別途定義される。トラストフレームワーク参加のための認証基準については統括ノードの管理団体によって別途定義される。

X.3.2.2 児童生徒用複数識別子の管理

児童生徒の識別子は、自治体・校種・システムごとに新規発行される可能性があるため、次のような手順で運用されることとする。

- 転校や進学直前の児童生徒の識別子は、電子メールなどの非定形による連絡や書類郵送等の物理的な手段による連絡を許容する
- 異動先システムに入力された児童生徒の識別子は、データ連携基盤間の連携により、管理元システムに問い合わせが行われる。
- 問い合わせ先のシステムでは、過去の当該児童生徒の識別子の履歴が管理されており、遡って移動前の学校の情報も物理的な連絡などを通じて識別子を管理する。
- 異動先システムはこれらのリストを順にたぐり、識別子の履歴リストを取得する。
- 識別子の履歴を使い、必要な校務系情報とスタディ・ログを取得する。



図x.　異動時のデータ引き継ぎ

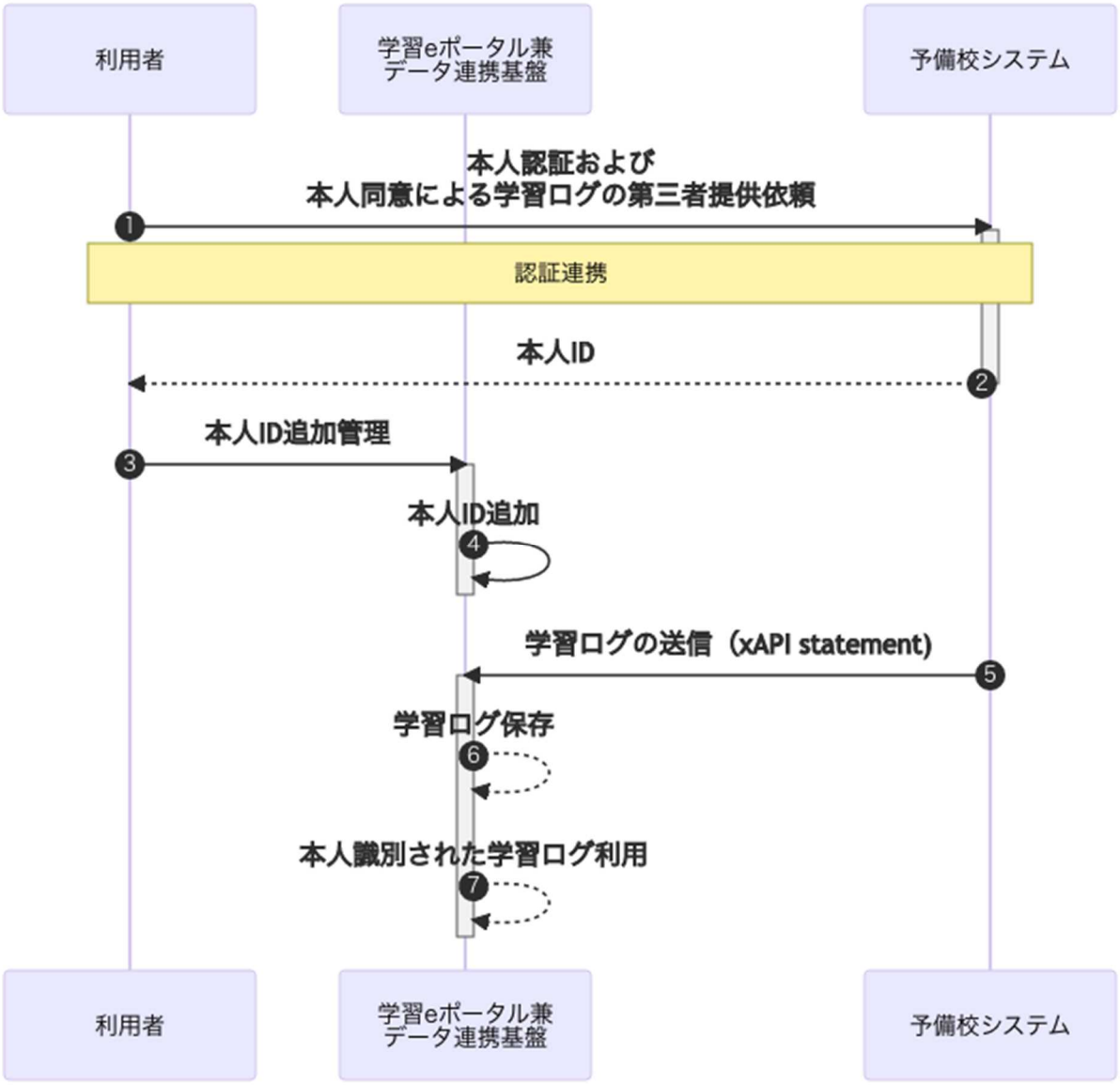
X.3.2 予備校の模試情報共有

データ連携基盤が異なる民間の教育事業者から自治体へのデータ連携の例として、予備校の模試情報共有を検討する。この方法は、学習系システムにおける任意の学習システム連携に親和性があるため、予備校システムからxAPI statementを学習eポータルに送信するまでの手続きについて考える。

予備校システムが学習eポータルのコンテンツとして振る舞う際は、学習eポータル自身の機能によりSSOおよびID連携が行われるが、データ連携基盤を介した外部システムとして振る舞う場合は、SSOを用いずにID連携を行わなければならない。

この場合、利用者のみが各システムへの接続ができることから、利用者自身が予備校システムから取得したIDをデータ連携基盤に登録することでデータ連携を実現することが現実的である。

予備校システムから発行されたIDを学習eポータルに設定する際は、ワンタイムパスワード、生年月日や電話番号等の複数要素による認証がされることが望ましい。

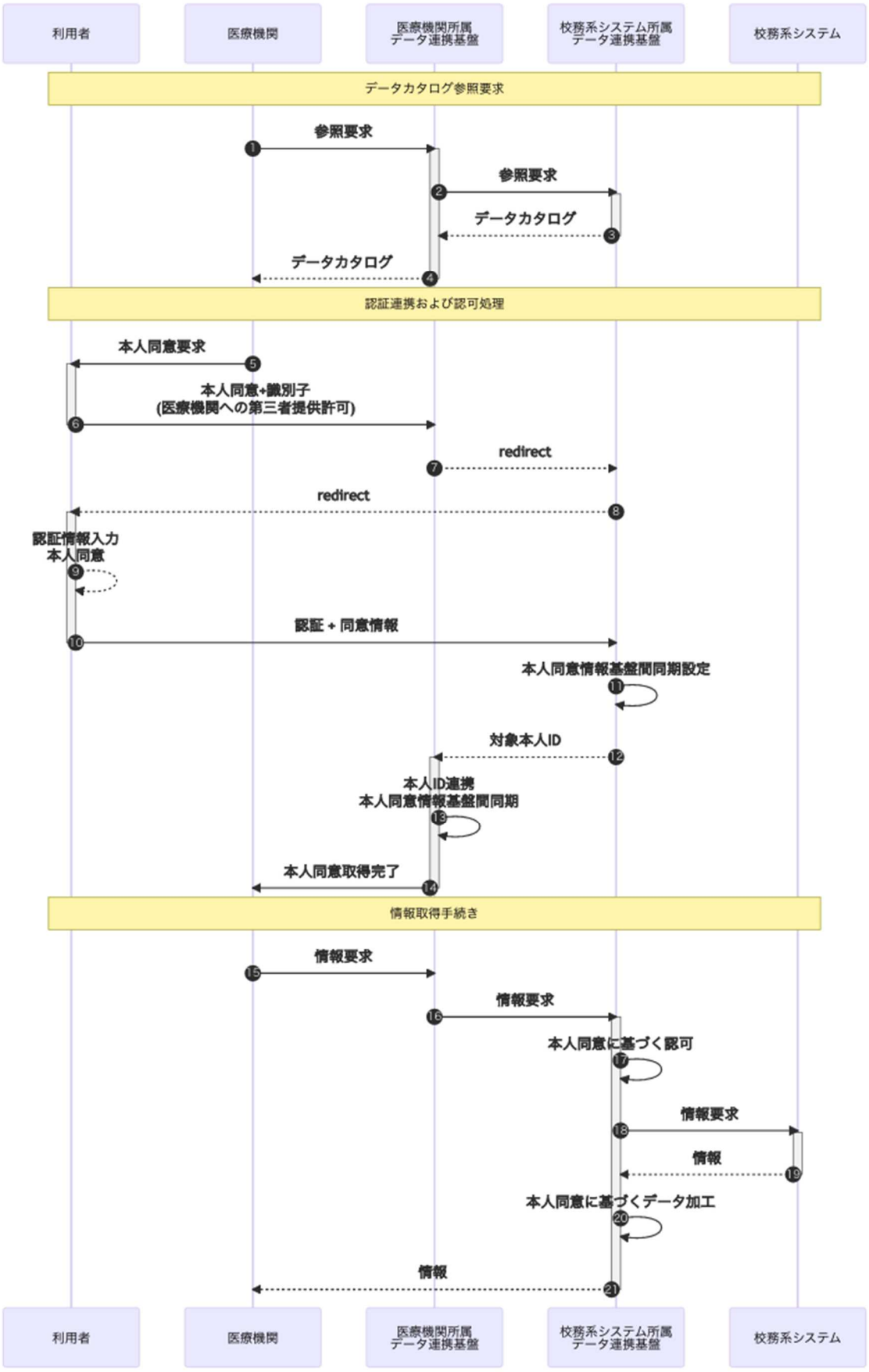


図x. 私教育機関のデータ利用

X.3.3 学校保健情報の医療機関提供

データ連携基盤が異なる自治体から民間へのデータ連携の例として、医療機関が校務系システム内の学校保健情報等を取得する例を考える。医療機関が校務系システムの情報を取得する際、データカタログの参照後に認証連携と本人同意が行われ、データ連携基盤間のID連携と認可処理が行われる。

データ取得手続きにおいては、医療機関が本人の代理人として機能する。本人の代理人からの要求に対して、校務系システムが所属するデータ連携基盤は、本人同意に基づく認可処理を実施する。



図x. 医療機関への情報提供

X.4 まとめと提言

本検討では、データ仲介機能を用いたデータ連携基盤を教育分野に活用するための仕様の検討を行った。その結果、各自治体が一つ以上のデータ連携基盤を運用することでデータ連携基盤間の識別子、認証情報、認可（本人同意）情報の連携が課題になることが明らかになり、この解決のための連携仕様案を検討した。

次年度以降の本検討の方向性として、自治体、学習eポータル事業者、校務系システムベンダー等の運用実態を調査し、各ステークホルダ間で許容可能なデータ連携基盤の教育分野用標準仕様案を策定することで、自治体がデータ連携基盤を導入する際の調達要件として利用できるものを示すことを目的とすべきである。

また、調達仕様化可能な仕様の策定のためには、データ連携基盤間の連携仕様検討と平行して、連携の前提となるデータ連携基盤自体の導入や運用を実証することで、連携対象となる学習eポータルや校務系システムなどの管理データを連携基盤が扱うための実現例を示す必要がある。