

教育における広域なデジタルコンテンツの利活用環境の整備に向けた
調査研究（再度公告）
～デジタルコンテンツの利活用環境のプロトタイプ作成のための調査～

本調査研究事業の実施状況報告

2023年 3月 28日

富士ソフト 株式会社

目次

1. はじめに

- 1-1. ドキュメントの目的 … 2
- 1-2. ドキュメントの構成 … 3

2. 事業概要

- 2-1. 本事業の目的 … 5
- 2-2. 研究調査の手法（将来像の検討） … 6
- 2-3. プロトタイプシステムの利用者 … 7

3. 実証するユースケースと課題

- 3-1. 本事業開始当初のイメージ … 9
- 3-2. 本事業開始当初の将来イメージ（ユースケース） … 10

4. 実証結果報告

- 4-1. 学習指導要領コードとの対応付け調査 - 紐づけ作業の作業前提 … 12
- 4-1. 学習指導要領コードとの対応付け調査 - 紐づけ作業の作業内容（紐づけの対象） … 13
- 4-1. 学習指導要領コードとの対応付け調査 - 紐づけ作業の作業内容（紐づけ準備作業） … 14~15
- 4-1. 学習指導要領コードとの対応付け調査 - 紐づけ作業の作業内容（書籍、画像、動画紐づけ作業） … 16
- 4-1. 学習指導要領コードとの対応付け調査 - 紐づけ作業の作業結果 … 17
- 4-2. メタデータ自動付与システムプロトタイプの開発と検証 - AIの役割と処理フロー … 18
- 4-2. メタデータ自動付与システムプロトタイプの開発と検証 - 作業前提（サマリー情報生成） … 19
- 4-2. メタデータ自動付与システムプロトタイプの開発と検証 - 検証結果と課題（サマリー情報生成） … 20~21
- 4-2. メタデータ自動付与システムプロトタイプの開発と検証 - 作業前提（メタデータの自動付与） … 22
- 4-2. メタデータ自動付与システムプロトタイプの開発と検証 - 検証結果と課題（メタデータの自動付与） … 23~24
- 4-2. メタデータ自動付与システムプロトタイプの開発と検証 - 作業前提（再学習機能の実装） … 25
- 4-2. メタデータ自動付与システムプロトタイプの開発と検証 - 検証結果と課題（再学習機能の実装） … 26
- 4-3. コンテンツ検索APIプロトタイプの開発と検証 - 紐づけ作業の作業前提 … 27
- 4-3. コンテンツ検索APIプロトタイプの開発と検証 - 方式と基盤構成 … 28
- 4-3. コンテンツ検索APIプロトタイプの開発と検証 - 提供機能・画面イメージ … 29
- 4-3. コンテンツ検索APIプロトタイプの開発と検証 - メタデータ設計と構造 … 30~31
- 4-3. コンテンツ検索APIプロトタイプの開発と検証 - データ連携実証の概要 … 32
- 4-3. コンテンツ検索APIプロトタイプの開発と検証 - データ連携実証の成果 … 33~34
- 4-3. コンテンツ検索APIプロトタイプの開発と検証 - 各種ヒアリング概要 … 35
- 4-3. コンテンツ検索APIプロトタイプの開発と検証 - 課題（サービスの観点） … 36~37
- 4-3. コンテンツ検索APIプロトタイプの開発と検証 - 課題（データ連携観点） … 38
- 4-3. コンテンツ検索APIプロトタイプの開発と検証 - 課題（コンテンツ紐づけ観点） … 39

5. 将来に向けた解決すべき課題

- 5-1. 課題の整理 … 41
- 5-2. ロードマップ案 … 42
- 5-3. 社会実装に向けて … 43~46

1-1 . ドキュメントの目的

<プロジェクトの背景（調達仕様書）>

- 政府は、行政の縦割りを打破し、大胆に規制改革を断行するための突破口としてデジタル庁を創設することを柱としたデジタル改革について検討を加え、令和2年12月25日、IT基本法の見直しの考え方やデジタル庁設置の考え方について政府の基本的な方針を盛り込んだ「デジタル社会の実現に向けた改革の基本方針」（以下「デジタル改革基本方針」という。）を閣議決定し、その後、これに基づき、デジタル庁のみならず、政府全体としてデジタル社会の形成の推進のために必要となる施策を盛り込んだ「デジタル社会の実現に向けた重点計画」が閣議決定された（令和3年6月18日、12月24日および令和4年6月7日）。
- また、デジタル庁および関係省庁において、学校内外のデータの将来的な連携も見据えた教育データの蓄積・流通の仕組みの構築に向けて、目指すべき姿やその実現に向けて必要な措置を盛り込んだ「教育データ利活用ロードマップ」を令和4年1月7日に策定・公表した。
- これらも踏まえ、初等中等教育においては、GIGAスクール構想を通じた児童生徒1人1台端末の整備や教職員端末の整備、学校におけるネットワーク環境の改善などが進められているが、教育に関わるデータの利活用環境が整っていないため、校務負担の軽減や効果的な教育内外の分野間のデータ連携が十分に進められていないことが課題となっている。

<本業務の目的（調達仕様書）>

- 小中高等学校の学習指導要領を踏まえた個々の指導内容について、国立国会図書館などに登録されている個々のデジタル書籍・デジタル画像・デジタル動画との関連性を調査し、同調査結果を踏まえ、GIGAスクール端末などやインターネットを通じて容易に検索できるようAPIのプロトタイプを作成する。

<ドキュメントの目的>

本資料では、上記調達仕様に基づいて、実証検証を行い、成果として作成したプロトタイプシステムを用いて、運営委員会やデータ連携実証事業（別調達）、教育関係者からご意見を頂いた結果をまとめ、更なるデータ利活用に向けた課題などを報告するためのものである。

1. はじめに

1-2 . ドキュメントの構成

ドキュメント名	概要	ページ数
【サマリ版】教育における広域なデジタルコンテンツの利活用環境の整備に向けた調査研究報告書_別紙1_用語集.pptx	本書に記載する用語の定義をまとめたファイル	4
【サマリ版】教育における広域なデジタルコンテンツの利活用環境の整備に向けた調査研究報告書_別紙02_運営委員会.pptx	運営委員会の開催状況をまとめたファイル。	1

02

事業概要

2-1 . 本事業の目的

・GIGAスクール構想を通じた児童生徒1人1台端末の整備や教職員端末の整備、学校におけるネットワーク環境の改善などが進められている中で、教育コンテンツホルダーであるコンテンツ提供者とコンテンツ利用者である児童・教員などの分野間データ連携が十分に進められていないため、メタデータ整備とメタデータ検索を行うための環境整備が必要である。

・今後も継続的な改訂・発展が予想される「学習指導要領」のコードに対し、膨大かつ増加し続ける広範な教育デジタルコンテンツを紐づけるためには、コード付与負担軽減が必要である。

上記課題を解決するために、教育デジタルコンテンツとして利用可能な官民のデジタル書籍・素材などについて学習指導要領コードとの紐づけを行い、メタデータとして利用するための間口の拡張を目的として検索APIのプロトタイプおよび、新しく参画されるコンテンツホルダーに対する学習指導要領コードの自動紐づけを行うためのAIプロトタイプを作成し、公開することで教育デジタルコンテンツの利活用促進を可能とすることを目的としている。

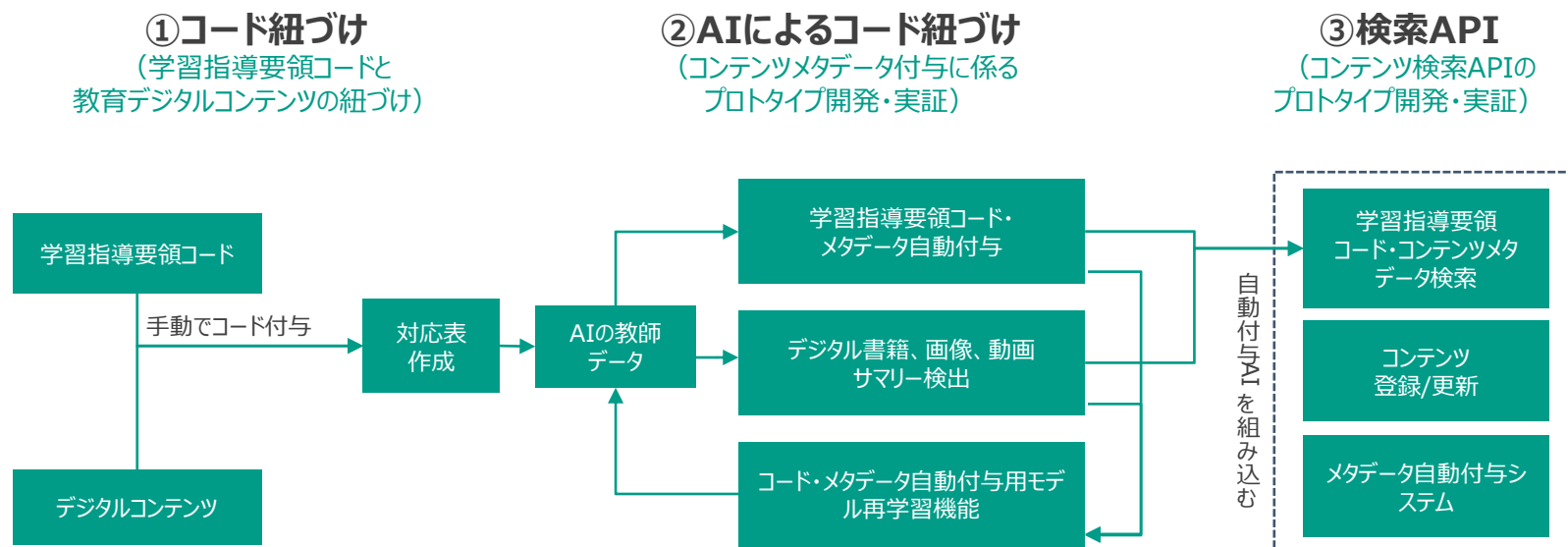


図:教育デジタルコンテンツの利活用環境の整備

2-2 . 研究調査の手法（将来像の検討）

本事業の将来像検討の対応方法として、バックカスティングを採用した。
中長期的な将来の理想像を定義の上、現状とのギャップを課題化し、本実証の目標を設定する手法が、不確実性の高い本事業において、最適であると判断した。
①の将来像を仮定し、「教育データ利活用」に向けた②課題の抽出（バックログ化）を行い、③にて課題解決に向けた作業実施を行い、④で大学教授や教育現場の有識者から実施結果に対する意見などを頂くことで課題の再定義（バックログの再整理）を実施。
この一連の流れを複数回行う事で、有識者からの意見を最大限取り込みながらプロトタイプシステムの作成を行った。

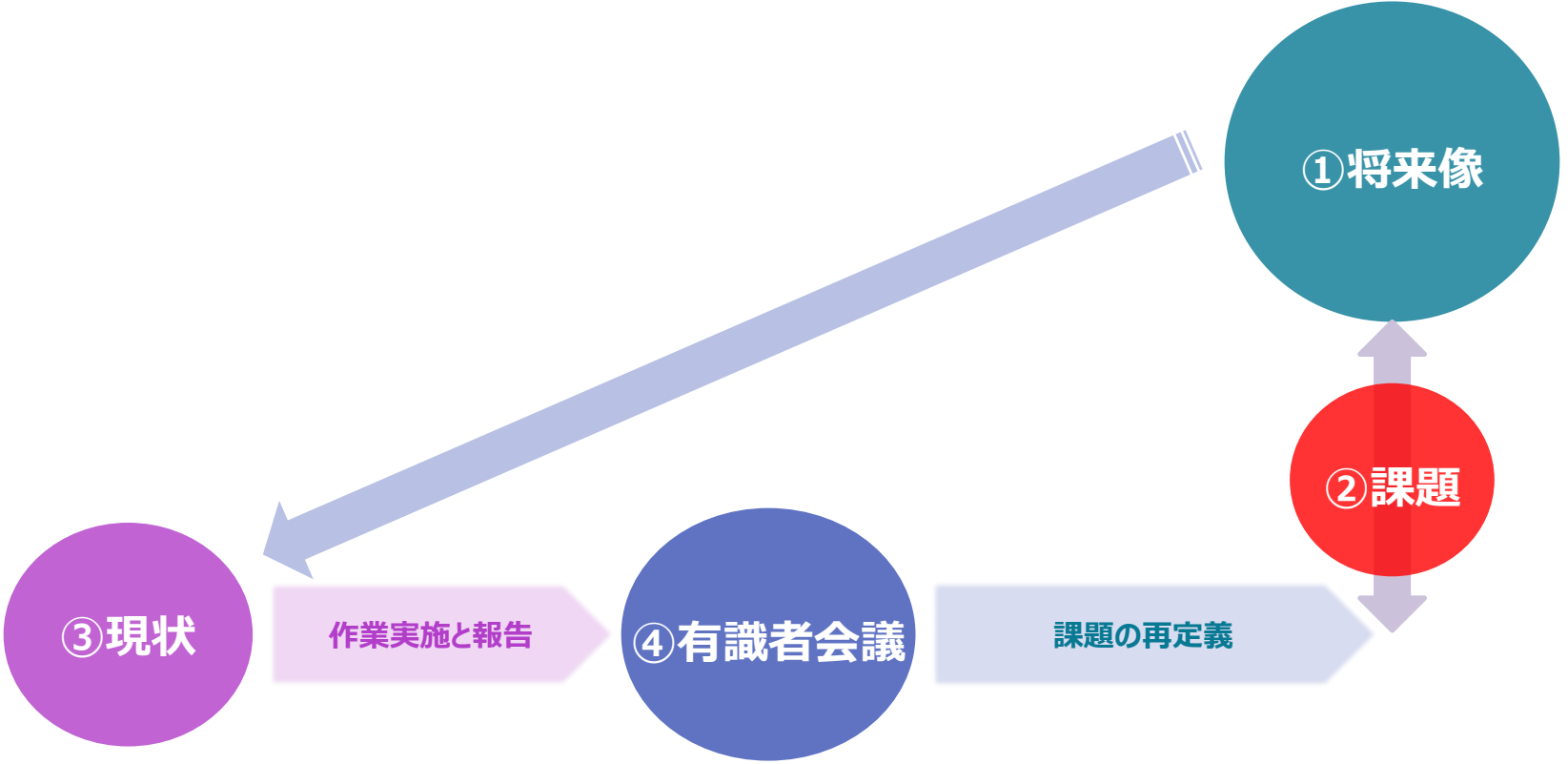


図:研究調査の手法

2-3 . プロトタイプシステムの利用者

プロトタイプシステムから検索APIを通じて連携されるメタデータは、様々な立場の方から参照されるデータとなるため、関係者から意見を伺うことで、プロジェクト開始当初に想定したユースケースとの齟齬やメタデータの在り方などに対して、様々な角度からの意見を基に課題抽出をし、可視化することを目的としている。

そのため、本事業では以下の表に示す利用者から意見聴取を行った。

表:プロトタイプシステム利用者一覧

No	利用者	概要
1	省庁関係者	本調達に関連する省庁 ・デジタル庁 ・文科省 ・経産省
2	データ連携実証事業者	別調達のデータ連携実証事業に係わる事業者 校務支援システム、学習支援システム、学習アプリなどのデータ連携実証事業を実施している事業者
3	教育関係者	大学教授や実際の教育現場で教育をされている教員

03

実証するユースケースと課題

3. 実証するユースケースと課題

3-1. 本事業開始当初のイメージ

本事業開始当初のイメージと課題を以下に示す。

表:課題 (先生)

No	課題	課題詳細
1	指導案 で調べる	<ul style="list-style-type: none"> • どう教えるかの指南書なので分かり易い説明用ではない • 探す時間がかかる
2	ネットで 調べる	<ul style="list-style-type: none"> • ページを探す時間がかかる • 目的のページが見つからない

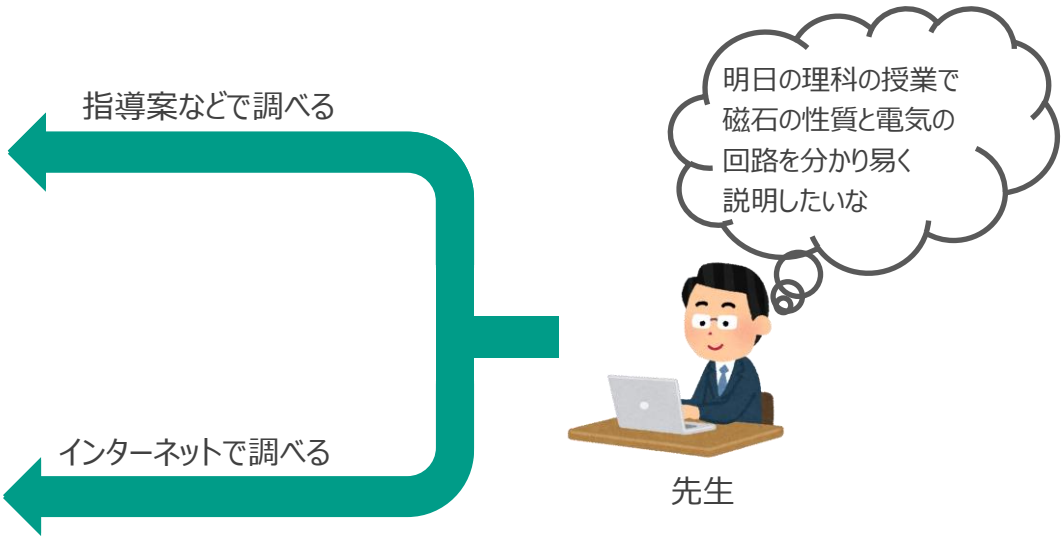


図:課題 (先生)

表:課題 (生徒)

No	課題	課題詳細
1	本で調べる	<ul style="list-style-type: none"> • 本を見る場所までの移動時間がかかる • 本を探す時間がかかる • 目的の本が見つからない
2	ネットで 調べる	<ul style="list-style-type: none"> • 検索の仕方が分からない • キーボードが打てない • ページを探す時間がかかる • 目的のページが見つからない

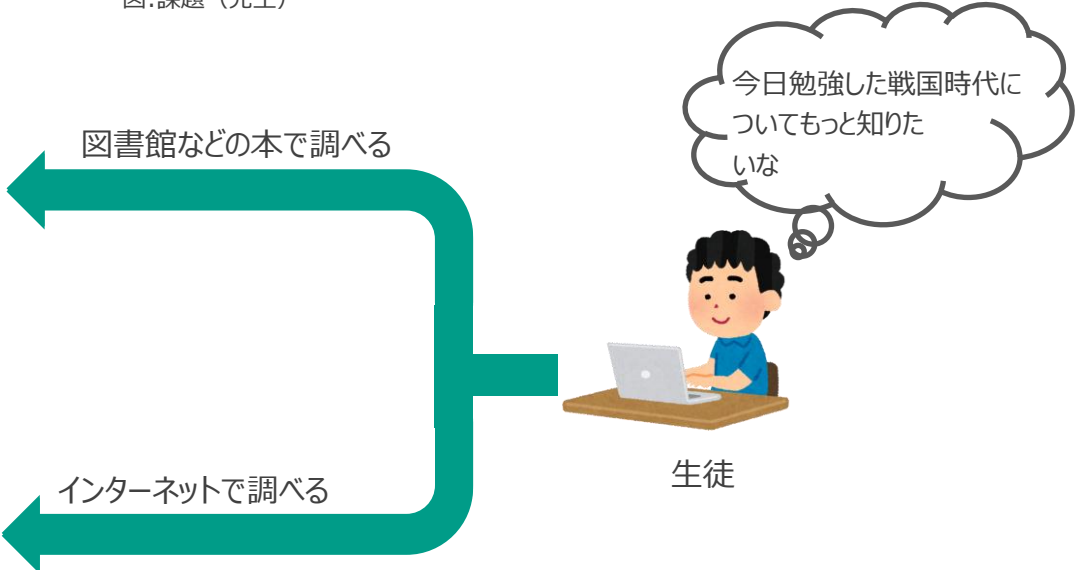


図:課題 (生徒)

3. 実証するユースケースと課題

3-2 . 本事業開始当初の将来イメージ (ユースケース)

本事業開始当初の将来イメージ (ユースケース) とメリットを以下に示す。

表:将来のメリット (先生)

No	メリット
1	<ul style="list-style-type: none"> ・探す時間がかからない ・目的のページが見つかる

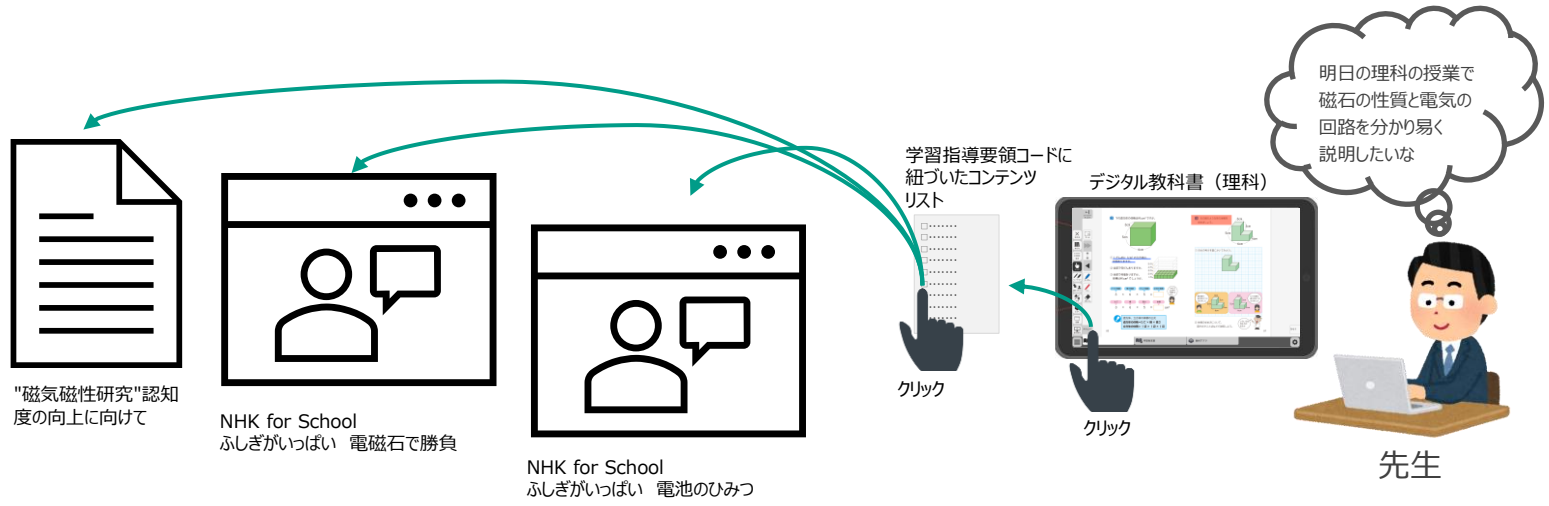


図:将来のメリット (先生)

表:将来のメリット (生徒)

No	メリット
1	<ul style="list-style-type: none"> ・移動時間がかからない ・探す時間がかからない ・タッチやクリックなのでキーボードが不要 ・目的のページが見つかる

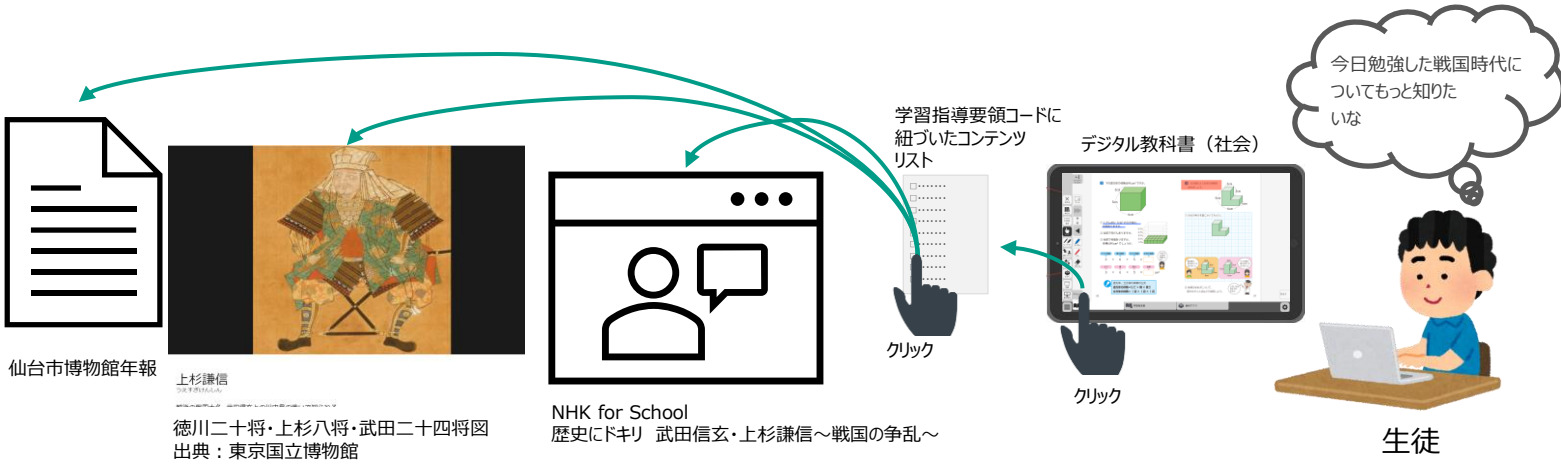


図:将来のメリット (生徒)

04

実証結果報告

凡例) 実証における問題の識別

- [実証内検知] : 実証の中で認識した部分
- [想定とブレ] : 計画に対してうまくいかなかった部分
- [ヒアリング指摘] : ヒアリングで指摘された部分
- [本来やるべき] : 計画を立てる上で諦めたが本来やるべき部分

4-1 . 学習指導要領コードとの対応付け調査 - 紐づけ作業の作業前提

(1) 作業前提

調達仕様 (抜粋)

- ・デジタル書籍に関する学習指導要領コードとの対応付け調査
指導用又は学習用に参考となるデジタル書籍（電子書籍）を、国立国会図書館サーチ、国立国会図書館デジタルコレクション又は STEAM Libraryより抽出し、学習指導要領コードと対応する書籍との対応表を作成する。
- ・デジタル画像に関する学習指導要領コードとの対応付け調査
指導用又は学習用に参考となるデジタル画像を、JAPAN SEARCH又は、STEAM Libraryにより検索可能な画像の中から抽出し、学習指導要領コードと対応するデジタル画像との対応表を作成する。
- ・デジタル動画に関する学習指導要領コードとの対応付け調査
指導用又は学習用に参考となるデジタル動画を、STEAM Library又は、JAPAN SEARCHにより検索可能な動画の中から抽出し、学習指導要領コードと対応するデジタル動画との対応表を作成する。

本実証におけるアプローチ

上記の調達仕様に記載のコンテンツホルダーに対して紐づけの準備作業として以下の手順で紐づけ作業を実施した。

- ①学習指導要領コードと紐づけ先のコンテンツホルダーの各コンテンツ数の調査。
- ②調達仕様に対して充足した件数の紐づけコンテンツ数となるかを確認し、不足している場合はデジタル庁と協議の上、対象とするコンテンツホルダーを追加。
- ③書籍・画像・動画それぞれに対する紐づけ作業ルールを制定し、紐づけ作業を実施。
紐づけ作業に関しては専門知識を必要とするため、教科書会社である株式会社文理に委任。
- ④定期的に関催される運営委員会（有識者として学芸大学の教授）より、適宜紐づけ結果に対する意見を頂くことで紐づけ作業の改善を実施。

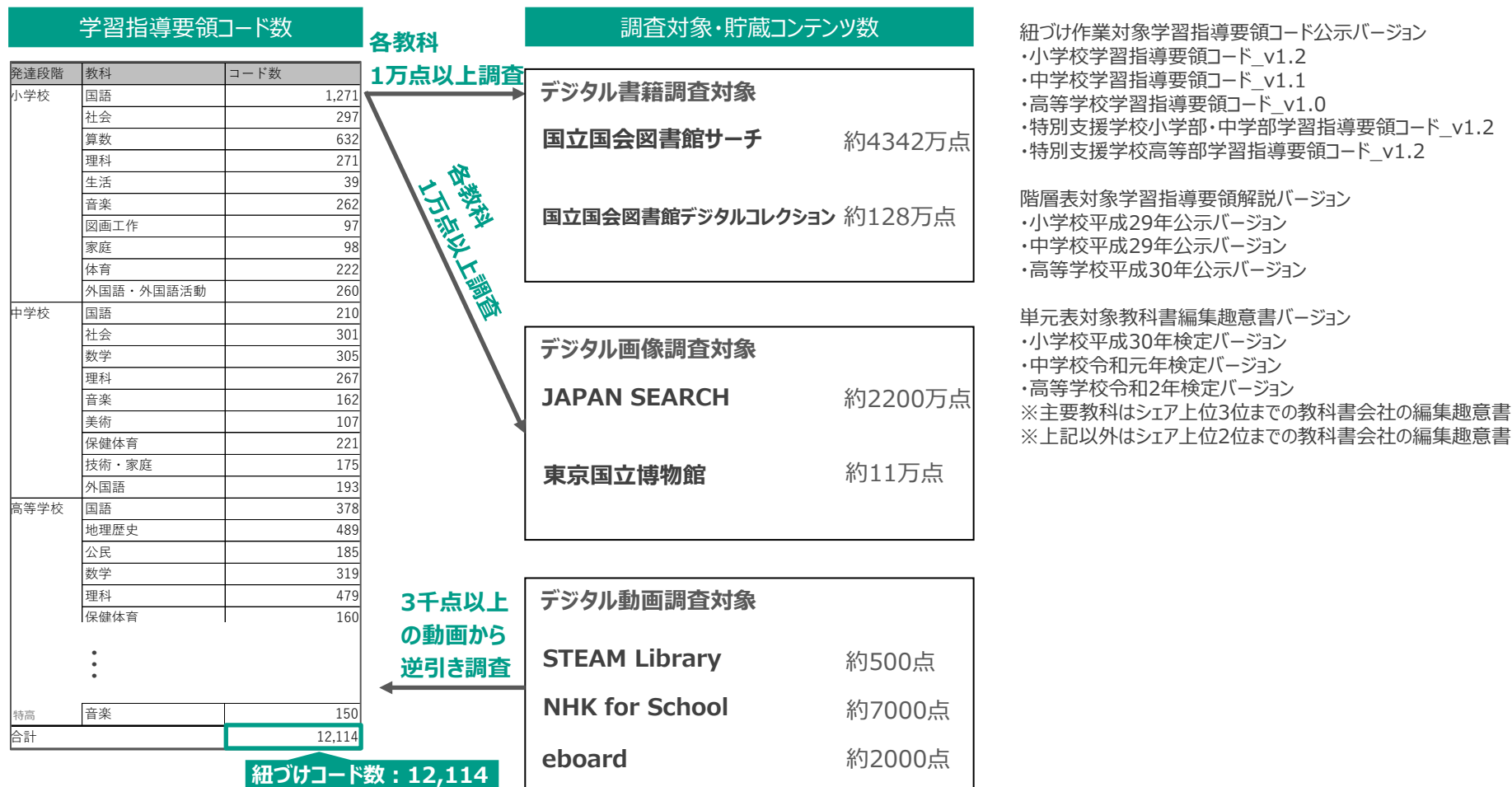
4-1 . 学習指導要領コードとの対応付け調査 - 紐づけ作業の作業内容（紐づけの対象）

(2) 作業内容

<紐づけの対象>

発達段階、教科に分けた学習指導要領コードとデジタル書籍・デジタル画像・デジタル動画との紐づけ調査を実施。

紐づけ調査の対象として、発達段階、教科に分けた学習指導要領コード数（紐づけコード数）、調査対象を下記に示す。



紐づけ作業対象学習指導要領コード公示バージョン

- ・小学校学習指導要領コード_v1.2
- ・中学校学習指導要領コード_v1.1
- ・高等学校学習指導要領コード_v1.0
- ・特別支援学校小学部・中学部学習指導要領コード_v1.2
- ・特別支援学校高等部学習指導要領コード_v1.2

階層表対象学習指導要領解説バージョン

- ・小学校平成29年公示バージョン
- ・中学校平成29年公示バージョン
- ・高等学校平成30年公示バージョン

単元表対象教科書編集趣意書バージョン

- ・小学校平成30年検定バージョン
 - ・中学校令和元年検定バージョン
 - ・高等学校令和2年検定バージョン
- ※主要教科はシェア上位3位までの教科書会社の編集趣意書
 ※上記以外はシェア上位2位までの教科書会社の編集趣意書

図:調査対象

4-1 . 学習指導要領コードとの対応付け調査 - 紐づけ作業の作業内容 (紐づけ準備作業) (1/2)

(2) 作業内容

① 紐づけ準備作業

効率的に紐づけ作業を行うための紐づけ準備作業を実施

表:紐づけ準備作業

No	紐づけ準備作業項目	作業詳細
1	紐づけコード選別	学習指導要領コードの中には、「内容」などのデジタル書籍などと紐づける必要がないコードが存在する。それらを コード付与作業前に要否選別 を行った。
2	特別、普通校コード照会	特別支援学校小学部・中学部、特別支援学校高等部の学習指導要領コードは、普通校用と重複が多く存在する。 重複部分に普通校の学習指導要領コードを紐づけ、流用することで効率化 を行った。
3	検索対象抽出	デジタル動画紐づけ用に STEAM Library、NHK for School (対象動画を絞り込む)、eboardのコンテンツをリスト化 した。
4	キーワード付与	学習指導要領コード表の学習指導要領テキストは曖昧な表現が多く、コンテンツの検索には適しません。 教材の専門知識を活かして紐づけるためのキーワードを付与 した。
5	メタデータ付与	キーワードを付けた対応表フォーマットに、 メタデータ (発達段階、教科、学年) を学習指導要領コードから自動算出し付与 した。

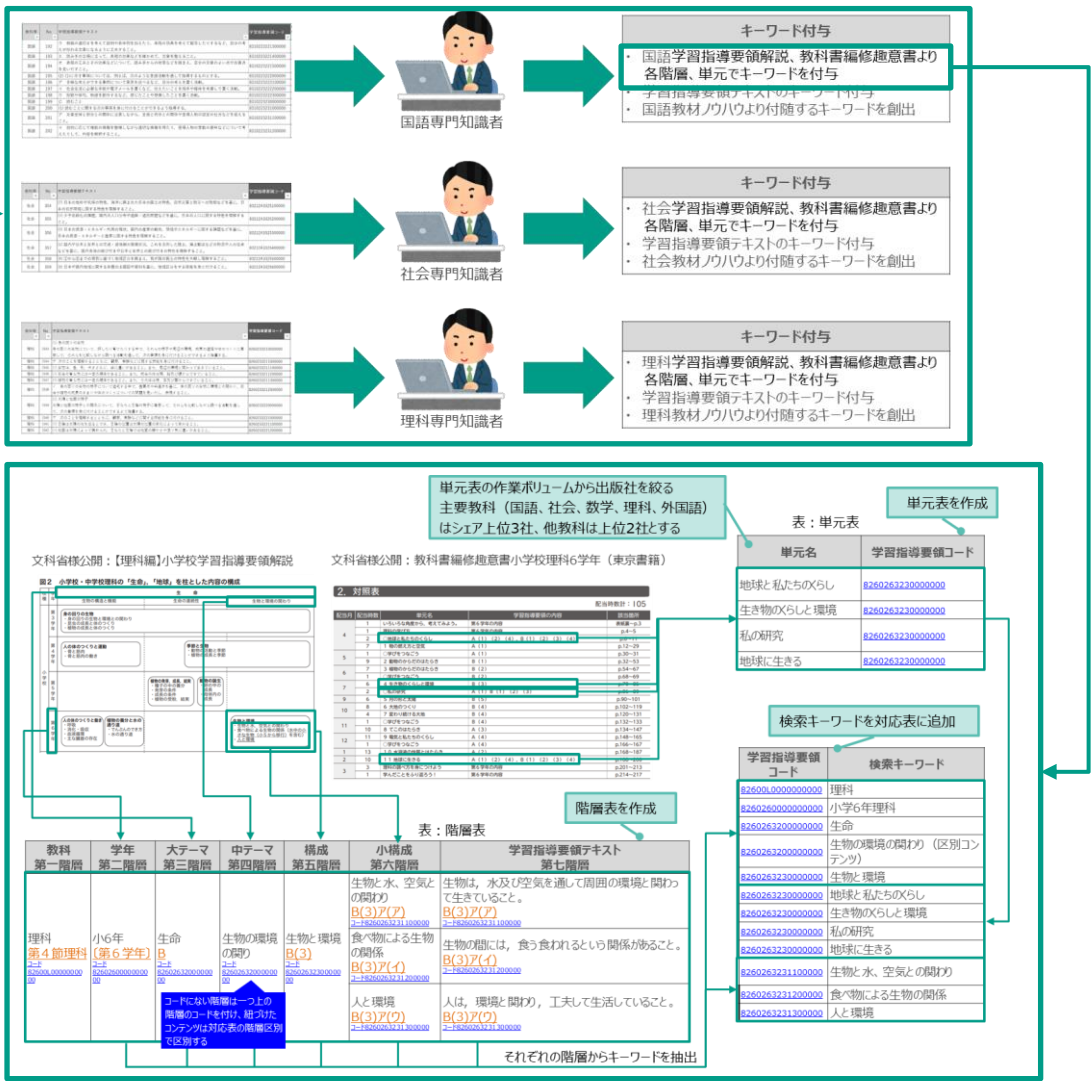


図:紐づけ準備作業

4-1 . 学習指導要領コードとの対応付け調査 - 紐づけ作業の作業内容（紐づけ準備作業）（2/2）

（2）作業内容

①紐づけ準備作業

教育デジタルコンテンツとの紐づけ準備作業中、以下の問題が発生した。

[実証内検知]

学習指導要領テキストだけではコンテンツ側と紐づけが難しい

学習指導要領は抽象度が高い概念や理念の体系であり、ある特定の箇所に相当する学習指導要領コードが単独で指し示す指導内容・教育内容が曖昧になる。

問題への対応

文科省が定める学習指導要領解説から「**階層**」を、同じく文科省が定める教科書編修正趣意書から「**単元**」を、それぞれ引用し、学習指導要領コードに対しコンテンツ検索作業を行う際に使用できるキーワードとして整備する事とした。
本実証では限定的な範囲として**東京書籍の趣意書**を用いている。

今後の課題

今後の課題としては東京書籍だけではなく、**各社の教科書編集趣意書を用いた「単元キーワード」の付与**が必要であると考える。

[実証内検知]

教師用、学習者用と言った区分けが必要

紐づけ対象となる書籍には、論文などが含まれており、児童が見て学ぶための教材としては不向きな物が存在した。

問題への対応

紐づけルールの中に**教師用、学習者用などのフラグ**を保持する。
そのフラグをAPIで返却し、UI側でのソートや絞り込みを行うことが可能となるように対応した。

4. 実証結果報告 - 「メタデータを付与する対象コンテンツの調査・整理」

4-1. 学習指導要領コードとの対応付け調査 - 紐づけ作業の作業内容 (書籍、画像、動画紐づけ作業)

(2) 作業内容

② 書籍、画像、動画紐づけ作業

発達段階、教科に分けた学習指導要領コードとデジタル書籍・デジタル画像・デジタル動画との紐づけ方法を下記に示す。

書籍紐づけ作業

キーワード入力

キーワード検索結果

検索結果 (タイトルなど) から紐づけ調査

一致する可能性があるデジタル書籍を選択

紐づけ判定

紐づけOK

No	学習指導要領テキスト	発達段階	教科	学年	キーワード	コンテンツタイトル	コンテンツ種類	紐づけポイント
103	エ アの内の「神話・伝承」については、古事記、日本書紀、風土記などの中から適切なものを取りよげる。	R2020E2 4000000	小学校	社会	6	古事記；日本書紀；風土記；		

出典：国立国会図書館ウェブサイト
 (https://xn--iss-xy9dc47aa5725d.ndl.go.jp/books?any=%E9%83%BD%E9%81%93%E5%BA%9C%E7%9C%BC%E3%81%AE%E7%89%B9%E8%89%B2&op_id=1)

画像紐づけ作業

キーワード入力

検索結果 (画像、タイトル) から紐づけ調査

画像とタイトルで紐づけOK判断

紐づけ

関連コンテンツ紐づけ

No	学習指導要領テキスト	発達段階	教科	学年	キーワード	コンテンツタイトル	コンテンツ種類	紐づけポイント
103	エ アの内の「神話・伝承」については、古事記、日本書紀、風土記などの中から適切なものを取りよげる。	R2020E2 4000000	小学校	社会	6	古事記；日本書紀；風土記；		

出典：JAPAN SEARCHウェブサイト

動画紐づけ作業

キーワード入力

検索結果 (画像、タイトル) から紐づけ調査

タイトル、概要から学習指導要領コードを逆引き検索

紐づけ

教科	No	学習指導要領テキスト	学習指導要領コード
理科	2075	生命の尊さについて、その連続性や有限性なども含めて理解し、かけがえない生命を尊重すること。	83K0400411000000
理科	2076	「自然愛護」	83K0400420000000
理科	2077	自然の崇高さを知り、自然環境を大切にすることの意義を理解し、進んで自然の愛護に努めること。	83K0400421000000
理科	2078	「感動、畏敬の念」	83K0400430000000
理科	2079	美しいものや崇高なものに感動する心をもち、人間の力を越えたものに対する畏敬の念を深めること。	83K0400431000000
理科	2080	「よりよくなる喜び」	83K0400440000000
理科	2081	人間には自らの弱さや弱さを克服する強さや気高く生きようとする心があることを理解し、人間として生きることの喜びを見いだすこと。	83K0400441000000

出典：経済産業省「未来の教室」STEAM Library
 (https://www.steam-library.go.jp/content/175/)

図：紐づけ方法

4-1 . 学習指導要領コードとの対応付け調査 - 紐づけ作業の作業結果

(3) 作業結果

①書籍、画像、動画紐づけ結果

発達段階、教科に分けた学習指導要領コードとデジタル書籍・デジタル画像・デジタル動画との紐づけ結果を下記に示す。

発達段階(教科)	コード数	細部キーワード数	画像キーワード数	単元キーワード数	書籍細部紐つき数	画像細部紐つき数	書籍単元紐つき数	発達段階(教科)	コード数	キーワード数	書籍細部紐つき数	
小学校(理科)	272	381	177	239	210	91	101	特小中(小学体育)	116	113	350	
小学校(体育)	223	341	316	70	222	243	2	特小中(小学図画工作)	75	33	152	
小学校(生活)	40	65	13	36	97	31	6	特小中(小学算数)	274	317	801	
小学校(図画工作)	98	113	101	303	72	26	116	特小中(小学国語)	127	121	745	
小学校(社会)	293	598	332	416	626	416	0	特小中(小学音楽)	181	141	1200	
小学校(算数)	634	421	431	727	121	127	1	特小中(中学理科)	129	42	396	
小学校(国語)	239	292	213	816	169	68	67	特小中(中学保健)	92	63	389	
小学校(外国語)	158	127	91	99	109	56	18	特小中(中学美術)	59	24	138	
小学校(家庭)	98	201	53	44	133	56	6	特小中(中学数学)	222	127	771	
小学校(音楽)	263	315	166	432	194	130	396	特小中(中学職業・家庭)	192	119	2609	
中学校(理科)	267	704	573	221	695	572	153	特小中(中学外国語)	107	113	1618	
中学校(保健)	221	405	859	169	212	442	1	特小中(中学国語)	74	42	500	
中学校(美術)	107	158	113	166	85	32	197	特小中(中学音楽)	173	141	1206	
中学校(数学)	305	226	265	154	140	145	98	特小中(中学社会)	107	101	562	
中学校(社会)	301	683	609	202	896	568	119	特小中(中学図画)	131	99	464	
中学校(国語)	210	206	185	462	125	24	350	特小中(中学英語)	305	1	58	
中学校(音楽)	175	420	284	218	341	199	33	特小中(中学英語)	212	112	636	
中学校(外国語)	193	158	210	397	141	196	0	特小中(中学英語)	212	112	636	
中学校(音楽)	162	177	111	202	168	47	111	特小中(中学英語)	212	112	636	
中学校(保健)	77	41	24	34	11	22	13	特小中(中学英語)	212	112	636	
中学校(理科)	461	834	329	376	867	341	378	特小中(中学英語)	212	112	636	
中学校(保健)	180	298	103	179	239	236	27	特小中(中学英語)	212	112	636	
中学校(地理歴史)	490	1230	361	792	967	104	454	特小中(中学英語)	212	112	636	
中学校(数学)	320	660	245	181	442	250	113	特小中(中学英語)	212	112	636	
中学校(情報)	118	212	79	183	126	34	80	特小中(中学英語)	212	112	636	
中学校(国語)	279	362	246	1084	246	1084	168	特小中(中学英語)	212	112	636	
中学校(公民)	186	360	82	117	173	173	0	特小中(中学英語)	212	112	636	
中学校(芸術)	565	565	411	638	349	83	997	特小中(中学英語)	212	112	636	
中学校(外国語)	353	228	197	222	296	79	110	15	特小中(中学英語)	212	112	636
中学校(家庭)	162	489	107	251	388	59	168	0	特小中(中学英語)	212	112	636
	7,550	11,437	7,331	9,001	87,144	4,872	5,019		3,296	2,230	22,075	

学習指導要領コードの細部キーワードでの検索結果より図書(本)が紐づく割合が低い結果となった。

階層表、単元表のキーワード検索では検索条件に図書(本)を加えて紐づけを実施。

階層表、単元表のそれぞれで約5千冊の図書(本)が紐づいた。

本実証の対象教科ではないが、自立活動の動画がある。

図:デジタル書籍紐づけ結果

大きく網をかけて紐づけ対象の書籍を抽出した。

発達段階(教科)	コード数	細部キーワード数	画像細部紐つき数	書籍細部紐つき数	発達段階(教科)	コード数	キーワード数	画像細部紐つき数
小学校(理科)	272	381	0	0	特小中(小学体育)	116	113	0
小学校(体育)	223	341	0	0	特小中(小学図画工作)	75	33	0
小学校(生活)	40	65	0	0	特小中(小学算数)	274	317	0
小学校(図画工作)	98	113	0	0	特小中(小学国語)	127	121	0
小学校(社会)	293	598	785	466	特小中(小学音楽)	181	141	0
小学校(算数)	634	421	0	0	特小中(中学理科)	129	42	0
小学校(国語)	239	292	86	61	特小中(中学保健)	92	63	0
小学校(外国語)	158	127	0	0	特小中(中学美術)	59	24	0
小学校(家庭)	98	201	0	0	特小中(中学数学)	222	127	0
小学校(音楽)	263	315	78	87	特小中(中学職業・家庭)	192	119	0
中学校(理科)	267	704	0	0	特小中(中学外国語)	107	113	77
中学校(保健)	221	405	0	0	特小中(中学国語)	74	42	0
中学校(美術)	107	158	147	71	特小中(中学英語)	107	101	0
中学校(数学)	305	226	0	0	特小中(中学英語)	74	42	0
中学校(社会)	301	683	228	228	特小中(中学音楽)	173	141	53
中学校(国語)	210	206	215	202	特小中(中学英語)	161	5	0
中学校(音楽)	175	420	0	0	特小中(中学英語)	106	103	0
中学校(外国語)	193	158	0	0	特小中(中学英語)	65	61	83
中学校(音楽)	162	177	49	217	特小中(中学英語)	314	96	0
中学校(保健)	77	41	0	0	特小中(中学英語)	86	3	0
中学校(理科)	461	834	0	0	特小中(中学英語)	110	129	0
中学校(保健)	180	298	0	0	特小中(中学英語)	107	90	0
中学校(地理歴史)	490	1230	640	22	特小中(中学英語)	131	99	0
中学校(数学)	320	660	0	0	特小中(中学英語)	138	15	0
中学校(情報)	118	212	0	0	特小中(中学英語)	150	132	43
中学校(国語)	279	362	123	64	特小中(中学英語)	3,296	2,230	301
中学校(公民)	186	360	0	0				
中学校(芸術)	565	565	389	132				
中学校(外国語)	353	228	0	0				
中学校(家庭)	162	489	0	0				
	7,550	11,437	3,206	1,111				1,976

予想のとおりで社会で画像紐つき数が多いが他教科は少ない結果となった。

図:デジタル画像紐づけ結果

発達段階(教科)	コード数	NhkForSchool	eboard	StreamLibrary	発達段階(教科)	コード数	NhkForSchool	eboard	StreamLibrary
小学校(理科)	272	14	0	11	特小中(小学体育)	116	0	0	0
小学校(体育)	223	1	0	0	特小中(小学図画工作)	75	0	0	0
小学校(生活)	40	1	0	0	特小中(小学算数)	274	0	0	0
小学校(図画工作)	98	0	0	0	特小中(小学国語)	127	0	0	0
小学校(社会)	293	29	31	6	特小中(小学音楽)	181	0	0	0
小学校(算数)	634	0	114	2	特小中(中学理科)	129	0	0	0
小学校(国語)	239	114	114	2	特小中(中学保健)	92	0	0	0
小学校(外国語)	158	114	114	2	特小中(中学美術)	59	0	0	0
小学校(家庭)	98	0	0	1	特小中(中学数学)	222	0	0	0
小学校(音楽)	263	0	0	0	特小中(中学職業・家庭)	192	0	0	0
中学校(理科)	267	0	0	0	特小中(中学外国語)	107	0	0	0
中学校(保健)	221	0	0	0	特小中(中学国語)	74	0	0	0
中学校(美術)	107	0	0	0	特小中(中学英語)	107	0	0	0
中学校(数学)	305	0	0	0	特小中(中学英語)	74	0	0	0
中学校(社会)	301	0	0	0	特小中(中学英語)	74	0	0	0
中学校(国語)	210	0	0	0	特小中(中学英語)	74	0	0	0
中学校(音楽)	175	0	0	0	特小中(中学英語)	74	0	0	0
中学校(外国語)	193	0	0	0	特小中(中学英語)	74	0	0	0
中学校(音楽)	162	0	0	0	特小中(中学英語)	74	0	0	0
中学校(保健)	77	0	0	0	特小中(中学英語)	74	0	0	0
中学校(理科)	461	0	0	0	特小中(中学英語)	74	0	0	0
中学校(保健)	180	0	0	0	特小中(中学英語)	74	0	0	0
中学校(地理歴史)	490	0	0	0	特小中(中学英語)	74	0	0	0
中学校(数学)	320	0	0	0	特小中(中学英語)	74	0	0	0
中学校(情報)	118	0	0	0	特小中(中学英語)	74	0	0	0
中学校(国語)	279	0	0	0	特小中(中学英語)	74	0	0	0
中学校(公民)	186	0	0	0	特小中(中学英語)	74	0	0	0
中学校(芸術)	565	0	0	0	特小中(中学英語)	74	0	0	0
中学校(外国語)	353	0	0	0	特小中(中学英語)	74	0	0	0
中学校(家庭)	162	0	0	0	特小中(中学英語)	74	0	0	0
	7,550	3,458	3,957	25,124					

NHK for Schoolは、小中学校の幅広い教科の内容が含まれていることが分かる。

eboardは、小中学校の主要教科の内容中心であることが分かる。

STEAM Libraryは、中高の内容中心であることが分かる。

図:デジタル動画紐づけ結果

4-2 . メタデータ自動付与システムプロトタイプの開発と検証 - AIの役割と処理フロー

(1) AIの役割と処理フロー

教育デジタルコンテンツに対する学習指導要領コードなどメタデータの付与は膨大な量の紐づけを可能とする仕組みが必要となる。しかし、学習指導要領の知見のある作業要員が紐づけ作業すべてを手動で行うことはコンテンツの量から現実的ではない。メタデータ自動付与システムとは、そのデジタルコンテンツへの学習指導要領コードの紐づけを自動化する役割を担う。

紐づけの自動化にあたっては、機械がコンテンツや学習指導要領コードの特徴を掴み、コンテンツにどのコードを紐づけるべきか判断する必要がある。これを実現する手段として、データから物事の特徴を掴むことができる技術 = AIを活用することが望ましいと考えられる。

AIを用いたメタデータ自動付与システムのプロトタイプの実証結果について述べる。なお、本システムの基本的な処理フローは以下のとおりである。

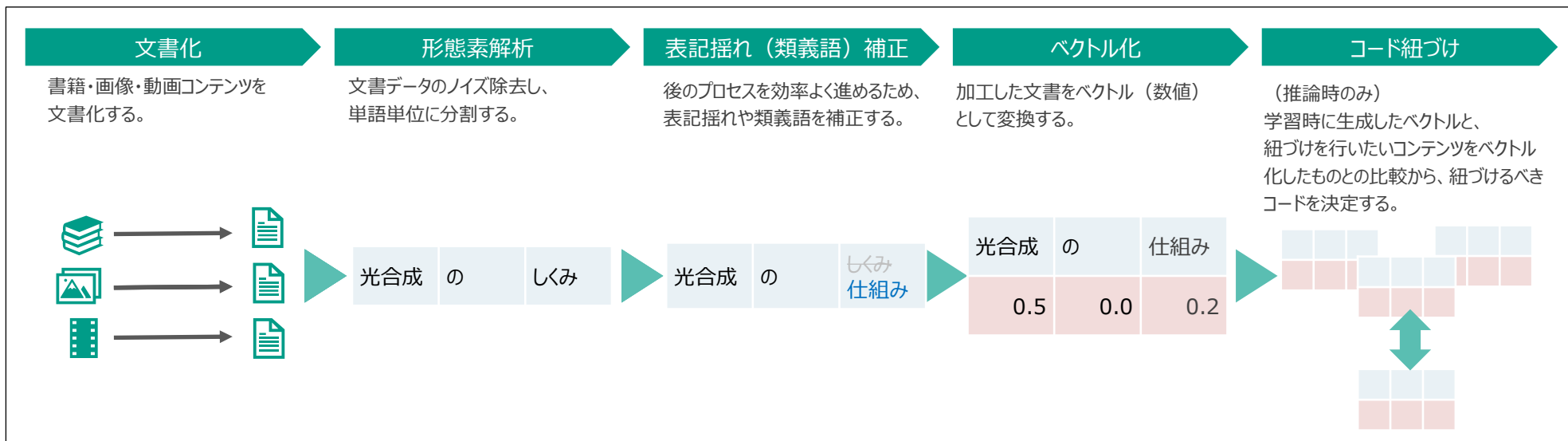


図:メタデータ自動付与システムの基本的な処理フロー

AI利用の必要性

AIを用いる以外にも、予め学習指導要領コードごとにキーワードを設定し、紐づけを行いたいコンテンツにそのキーワードが含まれるか否かで紐づけるコードを決定する方法なども考えられる。しかしこの方法では、キーワードの重要度やキーワード同士の共起性を考慮するのが難しいほか、そもそも何のキーワードを以て学習指導要領コードを表すことにすべきかを、人の知識で漏れなく設定していく手間がかかる。これらの問題を回避する上で、コンテンツのデータから機械が学習指導要領コードの特徴を学んでいく仕組み = AIの利用にはメリットがあると考えられる。

4-2 . メタデータ自動付与システムプロトタイプの開発と検証 - 作業前提 (サマリー情報生成)

(2) 作業前提

・デジタルコンテンツからのサマリー情報生成

調達仕様 (抜粋)

表:メタデータ自動付与システムに関する調達仕様

No	機能分類	機能	機能概要
1	デジタル書籍 ・デジタル画像 ・デジタル動画 メタデータ自動付与	デジタル書籍 サマリー検出	入力されたデジタル書籍に対し、ディープラーニングなどの自然言語処理機械学習技術を用い、記録されている内容について、内容のサマリー情報を生成する。 なお、本機能の実現にあたっては既存の文章認識、画像認識、動画認識サービスを活用することを可能とする。
2		デジタル画像 ・デジタル動画 サマリー検出	入力されたデジタル画像・デジタル動画に対し、ディープラーニングなどの機械学習技術を用い、記録されている内容について、内容のサマリー情報を生成する。 なお、本機能の実現にあたっては既存の画像動画認識・分析サービスを活用することを可能とする。

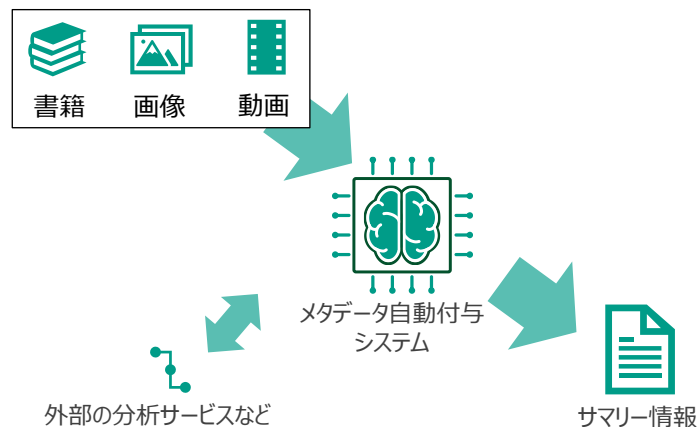


図:調達仕様のイメージ

本実証におけるアプローチ

OSSやベンダーサービスを用いて動画コンテンツなどを解析し、サマリー情報を取得して、それがメタデータの自動付与に有用であるか検証する。

一方で、コンテンツ本体が解析可能な形式で手に入らないケースや、解析の許可が得られないケース、そもそも解析の技術として難しいケースなどの問題も多いことから、コンテンツ本体ではなくコンテンツの掲載サイトのテキスト情報を用いてサマリー情報を生成する方法についても検証する。

4-2 . メタデータ自動付与システムプロトタイプの開発と検証 - 作業内容と課題（サマリー情報生成）（1/2）

（3）作業内容

表:本調達への協力を依頼したコンテンツホルダーと、
アーカイブ（コンテンツ本体）の使用可否

・デジタルコンテンツからのサマリー情報生成

コンテンツのサマリー情報生成に際して、以下の問題が発生。

[想定とブレ]

コンテンツ本体にアクセスできるケースが限定される

デジタル書籍コンテンツについては、本文に直接アクセスできるケースがほとんど無い（公開されていない）。

デジタル動画コンテンツについては、動画本体（アーカイブ）の利用許諾を得る必要があり各コンテンツホルダーとの調整を行ったが、権利関係などの問題から、無償での動画アーカイブ提供の許諾を得られたのは **STEAM Library** のみであり、幅広いコンテンツホルダーに対応する必要がある本システムとしては、書籍も含めてコンテンツ本体に依らないサマリー情報生成の方法を検討する必要が発生した。

No	コンテンツホルダー	コンテンツ掲載サイト	コンテンツのアーカイブ使用可否
1-1	国立国会図書館	国立国会図書館サーチ [書籍]	- : 書籍の本文の情報はほとんどない
1-2		国立国会図書館デジタルコレクション [画像]	可: サイトからダウンロードして二次利用条件下で利用
1-3		JAPAN SEARCH [画像]	可: サイトからダウンロードして二次利用条件下で利用
2	人間文化研究機構	国立歴史民俗博物館	否: 紐づけ自体がNG
3	国立文化財機構	東京国立博物館 [画像]	可: サイトからダウンロードして利用
4	経済産業省	STEAM Library [動画]	可
5	NPO法人eboard	eboard [動画]	可: ただし有償
6-1	NHK	NHK for School [動画]	否
6-2		NHK高校講座 [書籍 (テキスト)]	可: サイトからダウンロードしてAIの分析用に利用は可

問題への対応

コンテンツを提供するサイトにおいて掲載されているコンテンツの概要文やタグの羅列といった情報に着目し、Webクロールしてこれをサマリー情報として確保することを基本方針とした。

なお、STEAM Libraryからは動画アーカイブの無償の利用許諾を得られていたため、STEAM Libraryのコンテンツに対しては動画本体を解析する検証も実施できたが、クロールによる方法以上の紐づけ精度は得られなかった。

4-2 . メタデータ自動付与システムプロトタイプの開発と検証 - 作業内容と課題 (サマリー情報生成) (2/2)

(4) 検証結果と課題

・デジタルコンテンツからのサマリー情報生成

サイトのクローリング (URL内のすべてのテキストを取得する) で実施したところ、ノイズが強く、AIの精度が下がった。
対策として、URLのドメイン単位で個別にスクレイピング処理を実装する事で、再現率・適合率共に精度が向上することを確認した。
今回の実証により、サイトのテキストデータからサマリー情報を取得する場合、スクレイピングを行う必要があることが分かった。

[実証内検知]

運用容易でかつ、均質なコンテンツサマリー情報を得られる方法が必要

本検証では、書籍や動画の本体を入手して解析することに様々な制約があったため、コンテンツの掲載サイトからテキストデータをスクレイピングしてサマリー情報とする手段を主として用いた。しかしこの方法についても、以下のような問題がある。

- ・スクレイピング処理を掲載サイトの構成変更などに追従させる必要がある。
- ・クローリングを行う上で、コンテンツ掲載サイトへのアクセス負荷に配慮する必要がある。
- ・コンテンツごとの情報量が不均一であることが原因で、精度低下に繋がる場合がある。
(本検証の途中には、タイトルのみしか得られない多数のコンテンツによって、タイトル以上の有用な情報を持つはずのコンテンツへのメタデータ付与精度が低下したケースも見られた。)

今後の課題

- ・教育用のコンテンツのオーナーが、コンテンツのサマリー情報を所定のフォーマットに従った**オープンデータ**として公開する一つの解も考えられる。
例えば、タイトル/ジャンル/著者/概要などを含むフォーマットとすることで、AIなどで機械的な処理をし易いコンテンツとなり、メタデータ自動付与の精度向上も見込めると考えられる。
- ・クローリング/スクレイピングでサマリー情報を得る方法が続ける場合には、コンテンツのオーナーがコンテンツの掲載サイトに**ガイドライン**を設け、どのパーツがコンテンツのサマリーに該当するのかを明示することで、上記と同様の効果が期待できる。

4-2 . メタデータ自動付与システムプロトタイプの開発と検証 - 作業前提（メタデータの自動付与）

（2）作業前提

・デジタルコンテンツへのメタデータの自動付与

調達仕様（抜粋）

表:メタデータ自動付与システムに関する調達仕様

No	機能分類	機能	機能概要
1	デジタル書籍 ・デジタル画像 ・デジタル動画 メタデータ自動付与	学習指導要領コード・メタデータ自動付与	前述の機能により生成されたデジタルコンテンツサマリー情報に対し、本調査事業で得られたデジタルコンテンツおよび対応付け情報を教師データとした機械学習プログラムにより、学習指導要領コードおよびコンテンツ検索システムで使用するメタデータを付与する。

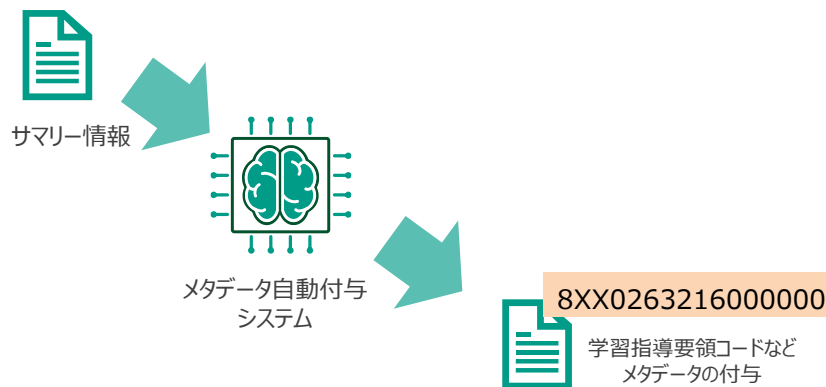


図:調達仕様のイメージ

本実証におけるアプローチ

数千種に上る学習指導要領コードの紐づけを行うにあたり、膨大な特許関連書類の検索システムに応用する研究などで使われた実績のある自然言語処理アルゴリズム「Okapi BM25」をベースとして機械学習システムの構築を試みる。

また、機械学習の中でもディープラーニングを用い、多数の自然言語処理タスクで実績を残している「BERT」についても適用が可能か検証する。

4-2 . メタデータ自動付与システムプロトタイプの開発と検証 - 検証結果と課題 (メタデータの自動付与) (1/2)

(3) 検証結果と課題

・デジタルコンテンツへのメタデータの自動付与

「Okapi BM25」と「BERT」の2つの手法で検証した結果、「Okapi BM25」を使うことで高い精度が得られることが確認できた。
検証中に発生した主な課題としては、以下が挙げられる。

[実証内検知]

発達段階の判定が難しいコンテンツ

コンテンツの中には、中学校の社会と高等学校の地理歴史など、ヒトが見てもどちらに紐づけるべきか（あるいは両方紐づけるべきか）判断に悩むものがあり、AIについても精度低下が見られた。

課題への対応

小学校・中学校・高等学校でAIモデルを分離することにより、大幅な精度向上が見られた。

今後の課題

教師データとなる紐づけ対応表について、将来的に「発達段階をまたいで教育に使えるコンテンツには、関係するすべての発達段階のコードを付与する」などの対応が徹底されれば、モデルを小中高で分ける必要がなくなる可能性はある。

[実証内検知]

普通学校と特別支援学校

AIの教師データとなる紐づけ対応表（ヒトによる紐づけ作業）において、すべてのコンテンツに普通学校と特別支援学校の両方からコードを紐づける検討がされているわけではない。この状態でAIに学習させると、意味的には正解の紐づけであっても、紐づけ対応表と一致しないことから不正解となるなど、実態と精度評価が乖離する懸念がある。

課題への対応

普通学校と特別支援学校はAIモデルを分けて作成することにより、正しく精度を評価できるようにした。

今後の課題

すべてのコンテンツについて普通学校と特別支援学校の両面からコードを付与することが将来徹底された場合には、AIモデルを分ける必要はなくなると考えられる。

[実証内検知]

紐づけの柔軟性向上

検証当初は、学習済みのコンテンツと自動付与したいコンテンツとを比較し、類似した学習済みコンテンツに紐づいていたコードを真似る方法を採用した。しかしこの方法では、学習済みコンテンツ1件1件の個性による影響が大きく、コードの選定を誤りやすい。また、学習済みコンテンツになかったコードの組み合わせは紐づけできない。

課題への対応

学習時にコンテンツをベクトル化する際、コード単位でベクトル化結果の平均をとることで、重要度の低い単語や尖ったコンテンツに惑わされずに紐づけの判断ができるようにした。また、学習済みコンテンツのコードを真似るのではなく、各コードについて紐づけの判定ができるため、コードを自在に紐づけられるようになった。

4-2 . メタデータ自動付与システムプロトタイプの開発と検証 - 検証結果と課題 (メタデータの自動付与) (2/2)

(3) 検証結果と課題

・デジタルコンテンツへのメタデータの自動付与

[実証内検知]

不適切な類義語処理による精度低下

今回類義語や表記揺れを補正するために用いたOSSの類義語辞書は、必ずしも教育の分野に最適ではない。例えば「層」という単語はこの辞書によって「レイヤー」という単語に寄せられるが、これによって地学で登場する「層」と、芸術で登場するペイントツールの用語である「レイヤー」が同義となり、正しいコード紐づけができないケースなどが見られた。

今後の課題

教育分野に適した類義語辞書を採用又は育成していくことで、より精度の高い紐づけが可能になる。
紐づけ誤りのコードやコンテンツを個別に分析して悪影響を及ぼしている単語を特定しつつ、修正後は全体への影響も確認する必要があるため、長期的な取り組みになると考えられる。

[実証内検知]

適合率と再現率のどちらを優先するか

適合率と再現率のどちらをどの程度優先するかは、ユースケースによって異なるため、適合率・再現率のバランスの閾値を決めることが難しい。

今後の課題

再現度を高めにしてコードを紐づけておき、かつ紐づけ時の**確信度 (スコア)**を保持しておくことで、確信度に応じた処理を実装することが可能になるが、本実証では対象外とした。
将来的には対応し、確信度の閾値を調整できるようにしておくことで、幅広くコンテンツを検索するか、関連度が高いコンテンツに絞って検索するかをユーザーが選択可能とするUIでサービスが可能となる。

[実証内検知]

教師データの精査

教師データとなる紐づけ対応表は人が作成しており、作業者によってコードの紐づけにブレが生じる場合がある。実際に、AIの紐づけ結果のうち不正解であったものを対応表作成者が確認したところ、正解とみなして良いケースが一定量見つかっている。

課題への対応

AIの紐づけ結果を分析し、教師データとなる紐づけ対応表に**フィードバック**し、修正した対応表でAIを再学習させることで、精度が改善した。

今後の課題

上記対応のように、教師データとなる紐づけ対応表に**フィードバック**する手順を繰り返せるように運用として確立することで、AIの精度も、対応表の品質も向上していくと考えられる。

4-2 . メタデータ自動付与システムプロトタイプの開発と検証 - 作業前提（再学習機能の実装）

（2）作業前提

・再学習機能の実装

調達仕様（抜粋）

表:メタデータ自動付与システムに関する調達仕様

No	機能分類	機能	機能概要
1	デジタル書籍 ・デジタル画像 ・デジタル動画 メタデータ自動付与	学習指導要領コード・ メタデータ自動付与用 モデル再学習機能	教師データが増加した場合に学習指導 要領コード・メタデータ自動付与機能の動 作に必要な学習済みモデルを再生成 する。

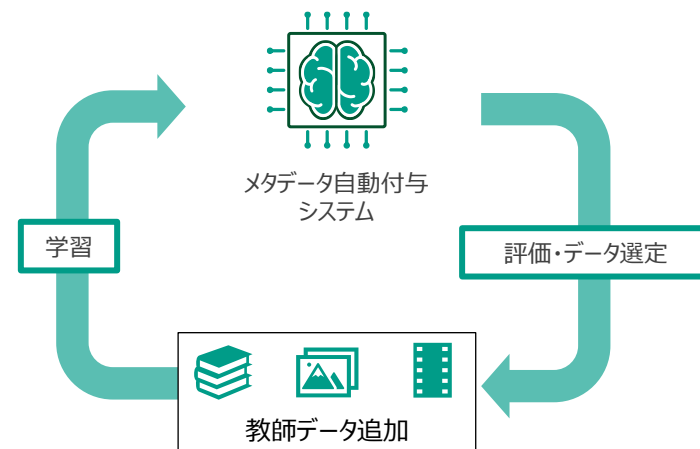


図:調達仕様のイメージ

本実証におけるアプローチ

将来的に実装コードが公開されることを見据え、再学習を行うための実装を行い、手順を整備する。

4-2 . メタデータ自動付与システムプロトタイプの開発と検証 - 検証結果と課題（再学習機能の実装）

（3）検証結果と課題

・再学習機能の実装

AI再学習手順を整備し、手順通りに実施する事で再学習が可能であることを確認した。
ただし、再学習においては以下の問題がある。

[実証内検知]

AIモデルの学習における品質管理の方法を定める必要がある

AIモデルの再学習を行う際、作成したAIの精度を評価し、リリースすべきか否かを判断するにはAIエンジニアやデータサイエンティストの知見が要請される。また、精度評価にあたっては技術面だけでなく、学習指導要領コードが正しく活用されるようガバナンスの観点も取り込む必要があると考えられる。



今後の課題

- ・管理者の負担を減らし、学習による紐づけ精度の改善を容易にするため、メタデータ自動付与システムの評価・改善を半自動的に実行できる機能を追加することが望ましい。
- ・文部科学省の定めた学習指導要領コードが正しく活用されるためのガバナンスの観点から、メタデータ自動付与システムの管理者を定める必要があると考えられる。

4-3 . コンテンツ検索APIプロトタイプの開発と検証 - 紐づけ作業の作業前提

(1) 作業前提

調達仕様 (抜粋)

- ・コンテンツ検索のアーキテクチャ整理
- ・コンテンツ検索システムのプロトタイプ版の開発&検証、限定的な公開

プロトタイプについては以下の機能を実装すること。ただし、その他管理機能やセキュリティ確保や利用実現において必要となる機能については実装を想定すること。
なお、デジタル庁が各機能の動作を確認するため、また、API接続を検討する事業者が動作を確認することを可能とするため、キーワードなどパラメータやデジタルコンテンツを入力し、機能を利用可能とする画面を作成すること。

「誰もが、いつでもどこからでも、誰とでも、自分らしく学べる社会」の実現に向け、一人一人のニーズにあった最適な学びが実現できるよう留意すること。

(コンテンツ検索システムの機能)

- 1) 学習指導要領コード検索
- 2) コンテンツメタデータ検索
- 3) 登録/更新

本実証におけるアプローチ

上記の調達仕様に対し、以下に留意した方式設計・開発&検証、および限定的な公開を実施した。

- ①クラウドサービスを利用したREST-API実装とし、標準規格に準拠したAPIインターフェース仕様書を整備する。
- ②APIが保持するメタデータは、国際的な標準規格を参考としつつ、学習指導要領コードとコンテンツ情報が検索し易い構造とする。

4-3 . コンテンツ検索APIプロタイプの開発と検証 - 方式と基盤構成

(2) 方式と基盤構成

教育デジタルコンテンツ検索基盤のプロタイプ実証検証を行うため、以下の基盤構成で教育デジタルコンテンツ検索APIおよびAPI検証画面の構成を行った。

- クラウドサービスを利用したAPIの実装

将来望まれる教育デジタルコンテンツ検索基盤として、AWSクラウドを利用したサービス構築を行った。安価な運用コストの実現やIaCによる基盤の柔軟な展開などの運用性の高い設計を実施した。

- API仕様書の作成と提示

教育デジタルコンテンツ検索サービスから容易な接続性を担保するためのインターフェース設計を基にAPI実装を行った。教育デジタルコンテンツ検索サービスから接続処理を実装する際に必要とするインターフェース仕様として、APIの標準仕様であるOpen APIの形式でインターフェース仕様書を記載し、実証関係者に提示した。

- API検証画面の準備と検証

メタデータ利活用の実証に関わる関係者が、教育デジタルコンテンツ検索基盤の機能を容易にイメージすることができ、実証作業の効率化を図るためにAPIへのリクエスト送信とその返却結果を簡易的に画面表示可能とした。

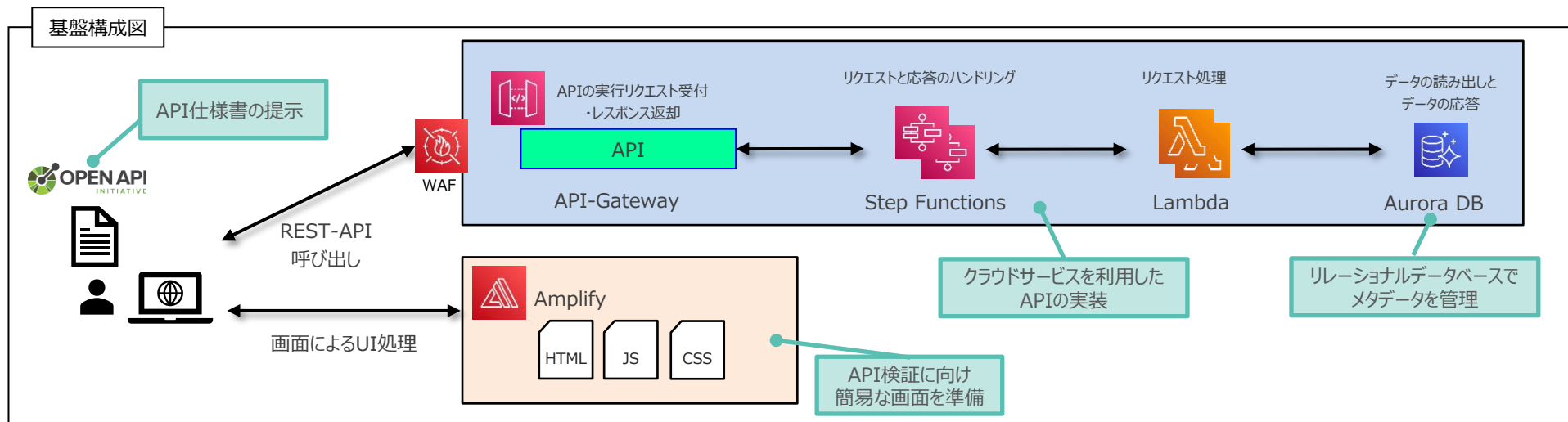


図:基盤構成図

4-3 . コンテンツ検索APIプロトタイプの開発と検証 - 提供機能・画面イメージ

(3) 提供機能

教育デジタルコンテンツ検索システムのプロトタイプとして以下の機能を作成した。

表:提供機能一覧

No	機能	機能概要
1	学習指導要領取得	検索の絞り込み条件に該当する学習指導要領の内容を一覧で取得する。
2	コンテンツ情報取得	検索の絞り込み条件に該当するコンテンツの情報の一覧を取得する。
3	コンテンツ紐づき学習指導要領取得	検索の絞り込み条件に該当するコンテンツと紐づいている学習指導要領の内容を一覧で取得する。
4	学習指導要領紐づきコンテンツ情報取得	検索の絞り込み条件に該当する学習指導要領と紐づいているコンテンツ情報を一覧で取得する。
5	コンテンツ情報登録	コンテンツの登録を行う。
6	コンテンツ情報更新	コンテンツの更新を行う。

(4) 画面イメージ

教育デジタルコンテンツ検索システムのプロトタイプの実証検証用の画面イメージは以下のとおりである。

The image displays two screenshots of the educational content search system's prototype. The left screenshot shows the '学習コンテンツを探す' (Search for Learning Content) interface. It features a search form with filters for school type (e.g., '高等学校'), subject (e.g., '算数/数学'), and keywords. Below the form is a table of search results with columns for 'コンテンツID', 'タイトル', '種別', and 'キーワード'. The right screenshot shows the '学習指導要領を探す' (Search for Learning Guidelines) interface, which has a similar search form and a table of results with columns for '学習指導要領コード', '表示バージョン', '学校種', '教科', '科目', and '学年・段階'.

図:学習コンテンツ検索画面、学習指導要領検索一覧画面

4-3 . コンテンツ検索APIプロトタイプの開発と検証 - メタデータ設計と構造 (1/2)

(1) メタデータ設計の成果

- 教育デジタルコンテンツ検索システム (API) による検索対象メタデータ構造の設計**

コンテンツ検索のユースケースを整理し、すべての検索パターンに耐え得るメタデータ構造を設計した。
W3Cが提唱するメタデータ規格のDCATを参考に、本システム独自の汎用的な構造とした。

- メタデータ構造**

教育デジタルコンテンツ検索基盤のプロトタイプ実証検証を行うため、以下のメタデータにより
教育デジタルコンテンツ検索APIおよびAPI検証画面の検証を行った。

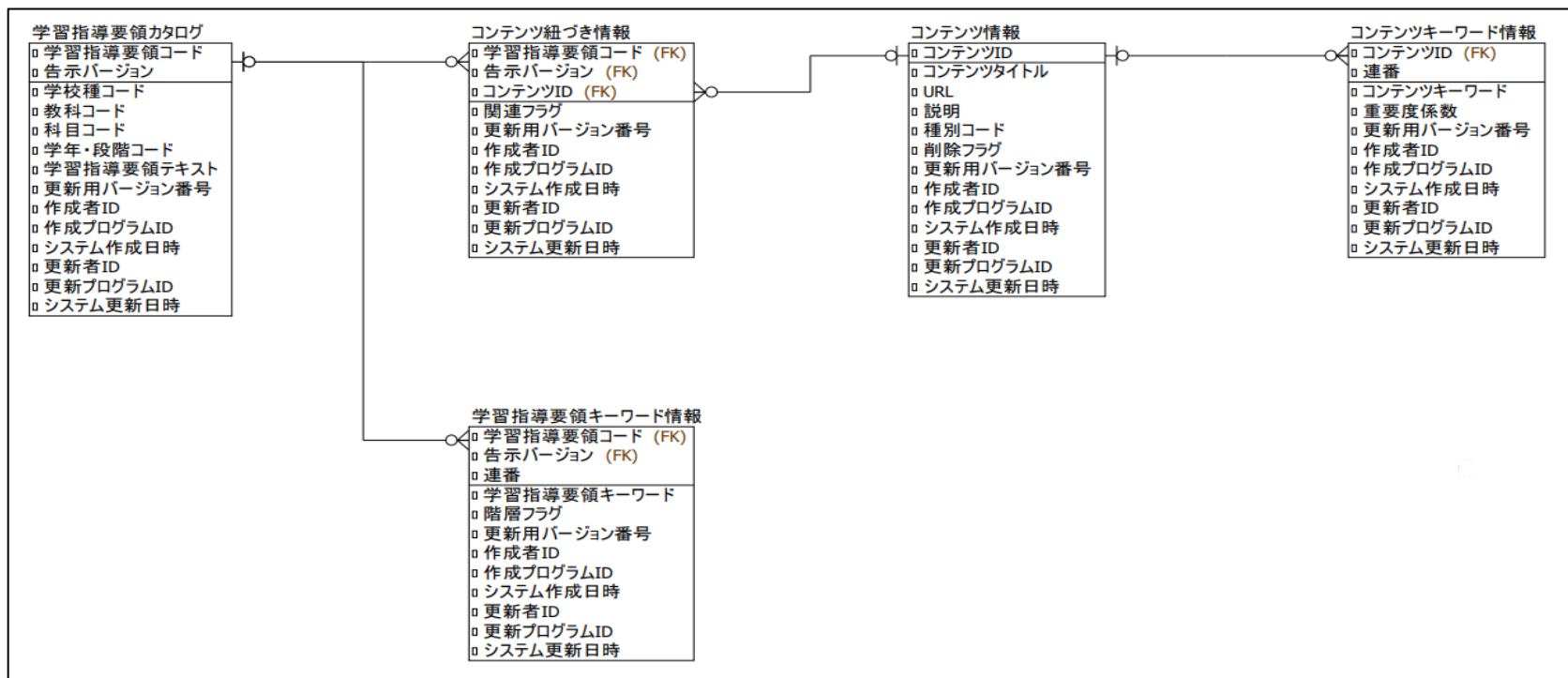


図:メタデータ構造モデル

4-3 . コンテンツ検索APIプロトタイプの開発と検証 - メタデータ設計と構造 (2/2)

(1) メタデータ設計の成果

- **実証を通じた気付き**

調達要件である「コンテンツ検索のアーキテクチャ整理」、および「コンテンツ検索システム（API）のプロトタイプ版の開発&検証、限定的な公開」をすることで、ユースケースに基づく利活用ができるか実証した。その結果、以下のような気付きに至った。

- ① 実証に参加いただいた他システム連携実証事業者、および教育関係者により様々なメタデータ項目の追加希望があったが、個人ではなく公共財としての蓄積やデータ流通を考えるため、品質の均一化が必要である。
- ② 作成方法や品質基準を規定することにより、誰もが同じ品質でメタデータを作成可能とする必要がある。

4-3 . コンテンツ検索APIプロトタイプの開発と検証 - データ連携実証の概要

(1) データ連携実証の概要

実証検証の中で作成したAPIをデータ連携事業者（※）からアクセスを行えるようにし、画面などの実装と検証を行えるように整備した。学習eポータルでのユースケースに基づく検証観点において検証を実施し、データ連携事業者ならではの視点で教育デジタルコンテンツ検索システムの課題や要望を抽出した。

・ データ連携実証による検証

データ連携事業者にて、APIへの実行リクエストとレスポンス受付を行う画面を他システムとして作成した。他システムとのデータ連携実証を行うことで、外部の他システムからも正常にAPIの動作を確認することができた。

本データ連携実証では、学習eポータルから教育デジタルコンテンツ検索基盤を活用する際の利活用方法を検討した。各データ連携事業者内でAPIの利活用方を考えた上で画面を作成しているため、各事業者の考えるAPIを活用した。画面の見え方や使い勝手を検証することができた。

※ … 学習支援システム（学習eポータル）の運営・開発事業者
（エヌ・ティ・ティ・コミュニケーションズ株式会社・ラーナーズ株式会社）

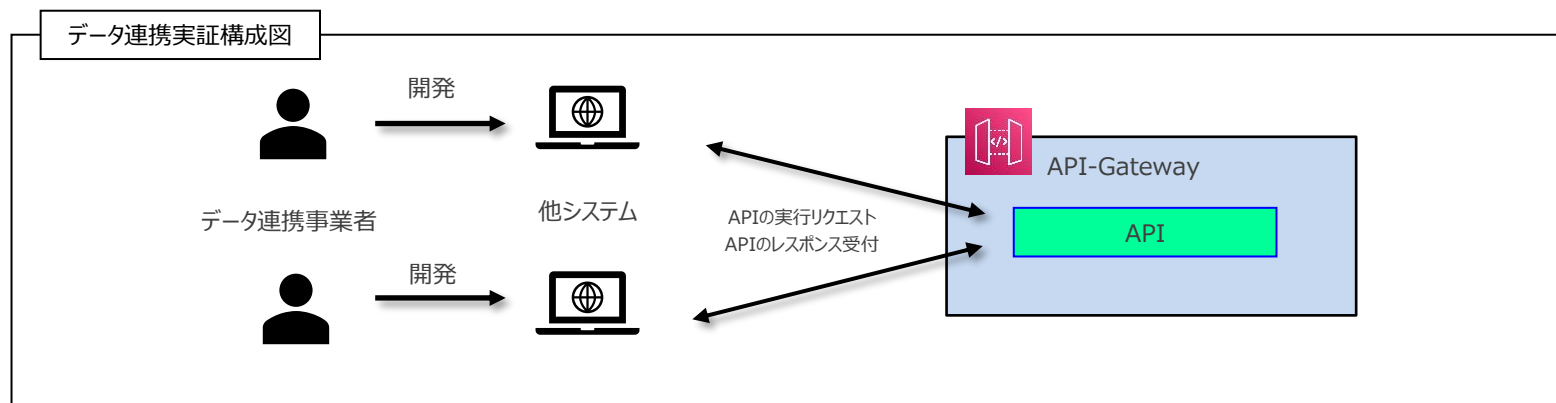


図:データ連携実証構成図

4-3 . コンテンツ検索APIプロトタイプの開発と検証 - データ連携実証の成果 (1/2)

(2) データ連携実証の成果

- **問題なく接続しデータ連携が可能であることが確認できた**

インターネット経由で、所定のAPIキー付与によるHTTPトランザクションにより、認証処理およびREST-API接続による検索処理の呼び出しができていたことを確認できた。

- **学習支援システム上で実際にREST-APIにアクセスする実装を行い、指定した条件にマッチするコンテンツ情報を取得して画面に結果を表示することができた**

データ連携実証事業者におけるAPI利用の実装を行い、ジャンルに基づく探索のイメージやコンテンツタイトルおよびキーワードによる絞り込みの操作によるコンテンツ検索サービスの1つのイメージが確認できた。

- **実サービスで使うためには処理性能改善が必要 (レスポンス時間が10秒を超えることがあった)**

実際の運用サービスでは、2-3秒程度のレスポンスが必要。

- **アクセス分析は本実証では未検証**

実際の運用サービスに向けて、改めて検討・検証が必要。

4-3 . コンテンツ検索APIプロトタイプの開発と検証 - データ連携実証の成果 (2/2)

(2) データ連携実証の成果

他システム連携実証結果を、機能・非機能観点で整理した結果は以下のとおり。

表:他システム連携実証結果

No	大項目	中項目	小項目	本事業スコープにおける前提・確認結果		将来像（検索サービス提供）に向けた課題	
1-1	機能	API接続機能	接続経路 ・認証方式 ・データ連携方式	・インターネット経由でアクセス ・APIキーによる簡易認証 ・HTTP（REST-APIに接続）	問題なく接続しデータ連携が可能であることが確認できた。	・学習プラットフォームにおけるコンテンツ検索基盤（API）への接続経路・認証方法・データ連携方式の課題抽出	・REST-APIを活用した接続・連携方式を踏まえつつ、サービス提供シーンの具体化と、実際の適用システムにおける課題抽出が必要。
1-2		接続管理機能	ライフサイクル管理・バージョン管理	プロトタイプによる試行のため、特に考慮無し	-	・接続管理方法に向けた課題抽出	・管理・運用者向けに、登録/参照/変更/削除などの機能要件の整理が課題となる。 ・接続先の非互換をどのように吸収するか、などの課題がある。
1-3		検索機能	コンテンツ検索機能	REST-APIを利用したコンテンツ情報取得を検証	REST-APIにアクセスし、検索結果を学習支援システム上で表示できた。	・学習プラットフォームにおけるユースケースの検討。 ・検索サービス提供の課題抽出	・コンテンツ検索サービスのユースケース、およびコンテンツ検索基盤（API）要件の深化が必要。以下、具体的な課題。 ①学習指導要領コードのマッチング条件の設計見直し ②利用者が検索時の足がかりとして利用するジャンル・基礎分類としてのキーワード整理 ③検索条件入力ワードの自然言語処理の高度化
2-1	非機能	性能	サービス性能	機能確認を通じたAPIレスポンスの定性評価のみ実施	検索結果が多くなる時にAPIレスポンスが10sec超となった。	性能要件の検討	・コンテンツ検索基盤（API）の性能要件の明確化が必要。 ⇒実運用に向けては、改めて「利用者・利用頻度・ピーク利用」などの前提条件を明確化した上で、利用負荷想定に基づく性能設計が必要。
2-2		セキュリティ	アクセス制限 ・レート制限	プロトタイプによる試行のため、特に考慮無し	-	セキュリティ要件の検討	・アクセス可能なコンテンツ検索基盤（API）やデータを制限できる要件の検討が必要。 ・コンテンツ検索基盤（API）の過負荷に対する、サービス制限方法の検討が必要。
2-3			監視			閉塞・障害考慮	・閉塞応答や障害応答時のハンドリング検討が必要。
2-4		分析 ・モニタリング	アクセス分析	プロトタイプによる試行のため、簡易な履歴記録のみ実施	特筆事項無し。	主体情報や利用履歴の課題抽出	・活動情報のスタディログへの統合検討を含めたサービス要件と、コンテンツ検索基盤（API）側におけるシステム要件の検討が必要。 ⇒コンテンツ検索などを含む活動情報の管理・利活用方法と2次利用の在り方は、今後の課題。

4-3 . コンテンツ検索APIプロトタイプの開発と検証 - 各種ヒアリング概要

(1) 各種ヒアリングの概要

関係事業者・教育関係者・省庁関係者のそれぞれの視点での気づきや要望、および課題を抽出するために、作成したAPI・画面を限定公開して、ヒアリングを実施した。

参加者

表:ヒアリング参加者

No	参加者	ご協力いただいた関係者	備考
1	事業者	学習プラットフォーム事業者 : 6社	デジタル庁における「教育関連データのデータ連携の実現に向けた実証調査研究」、および文科省の「学習指導要領コードの利活用に関する調査研究事業」の参加事業者。
2	教育関係者	・国立付属小中校の教員 : 3名 ・国立大学の教員志望学生 : 3名	
3	省庁関係者	文科省・経産省のご担当者	

● ヒアリング結果

説明会にて主旨と概要を説明後に画面試用いただいた上で、ユースケースや画面表示結果についてヒアリングした。

結果

表:ヒアリング結果

No	参加者	ご意見・気づき	備考
1	事業者	<ul style="list-style-type: none"> 実際に利用者が教育デジタルコンテンツを利用する場合には、学習指導要領が持つ曖昧さや構造を踏まえた紐づけが必要であり、品質面の課題がある。 メタデータの項目について、コンテンツの要旨や利用条件などの追加が必要である。 	学習プラットフォーム上の利用シーンについては、様々な意見があり、方向性は今後議論が必要。
2	教育関係者	<ul style="list-style-type: none"> デジタル教科書や指導書から関連コンテンツ（資料やワークシート素材）が引けるようになることが望ましい。 GIGA端末を生徒にどこまで使わせているか学校ごとに差異があるため、学校内・学校外・家庭などにおける画一的ではない環境の実態を踏まえて利用シーンを考慮することが必要。 児童にとっての関連コンテンツは、特に興味関心を引き出し自身の気づきに基づいて考えさせるための導入としての役割が大切であり、コンテンツの質が非常に重要となる。 	「東京ベーシックドリル」などの有益なコンテンツの充実が必要といった意見もあった。
3	省庁関係者	<ul style="list-style-type: none"> 基礎となる分野・分類・カテゴリなどのキーワード体系があれば、教育デジタルコンテンツの分類や紐づけメタデータの社会利用（オープンデータ化）がし易くなると考えられる。 現場の先生方が納得感を持って使っていただけるシステムにしないと民間利用も進まないため、随時全国の教員の先生方のご意見を踏まえて修正できるようなスキームを設定すべき。また、産官学民連携によるコンソーシアム的な議論も必要。 有効なユースケースの明確化が必要であり、教育関係者による実証やデータ連携（学習プラットフォーム）の実証によるユースケースの具体化が必要なのではないか。 	

4-3 . コンテンツ検索APIプロトタイプの開発と検証 - 課題 (サービスの観点 1/2)

(1) 課題

教育現場で実際に教育デジタルコンテンツサービスを活用するためには、生徒や教員が直接触れるシステムである「学習プラットフォーム」上でのサービスの在り方を考える必要がある。例えば、デジタル教科書や指導書などから関連コンテンツを引けるなど具体化が必要であり、実際の学習eポータルなどのシステムが求める機能を深堀して、教育デジタルコンテンツ検索基盤として具備すべきシステム要件の明確化を進める必要がある。

①「ユースケースの具体化」に関する課題

[ヒアリング指摘]

教育デジタルコンテンツ検索サービスのユースケース具体化

今回の実証検証においては国立国会図書館やNHK for Schoolなどの限定されたコンテンツを使用し、教育デジタルコンテンツ検索基盤(API)のプロトタイプを作成し限定的なユースケースの想定のもとでサービスに向けた検証を行った。しかし、実運用に向けては十分にユースケースが具体化されていない。

今後の課題

今後は、有識者や教育関係者を交えより多くの意見を集約し、デジタル教科書・指導書から関連コンテンツを引けるなどのサービスレベルのユースケースを仮定し、**教育関係者による検証や学習プラットフォーム事業者を巻き込んだ実証を行う**ことにより、ユースケースの具体化することが必要ではないかと考えられる。

[実証内検知]

教育デジタルコンテンツ検索基盤のシステム要件具体化 (機能面)

サービスレベルのユースケースを踏まえ、基盤として具備すべきシステム要件の明確化が必要である。また、教育・学習の観点や利用者が検索時の足がかりとして利用するジャンル・基礎分類としての側面でキーワード整理が十分に行えていない。

今後の課題

教育に特化した用語の考慮や、より自然で柔軟な入力を許容できるようにするなど**検索条件設計の改良**の検討が必要。また、検索結果が巨大なデータサイズとなる場合の**別経路や非同期連携などの考慮**が必要な可能性がある。

[本来やるべき]

教育デジタルコンテンツ検索基盤のシステム要件具体化 (非機能面)

サービスレベルのユースケースを踏まえ、基盤として具備すべき非機能面のシステム要件の明確化も必要である。

今後の課題

基盤に必要な機能をも明らかにするほか、実運用に向けては、改めて「利用者・利用頻度・ピーク利用」などの前提条件を明確化した上で、**利用負荷想定に基づく性能設計**が必要。フィルタリングによる**アクセス制約**への配慮や、**セキュリティ考慮**が必要となる。**分析モニタリング**によるコンテンツ検索などを含む活動情報の管理・利活用方法と2次利用の在り方の検討が必要。

4-3 . コンテンツ検索APIプロトタイプの開発と検証 - 課題 (サービスの観点 2/2)

② ユースケースの具体化に伴って付帯的に考慮すべき課題

[ヒアリング指摘]

ユーザビリティ・アクセシビリティの明確化

実証検証の中で省庁関係者や教育関係者向けに作成した教育デジタルコンテンツサービスを使用し、画面試用結果のヒアリングを実施した。

実証検証において用意した簡易な画面は、検索API仕様に基づいて呼び出しを行うのみのリファレンス実装であるため、ユーザビリティの面において使い勝手や多様な利用者を想定した設計が必要である。

[本来やるべき]

検索基盤の利用規約策定

本実証では、実際の教育デジタルコンテンツ検索基盤を利活用する上での利用規約の策定を行う必要があることを確認した。サービスレベルのユースケースを踏まえ、基盤として具備すべきシステム要件の明確化が必要である。

今後の課題

ヒアリングした中では、キーワード検索など、検索画面を通じたサービスでは小学生にとってコンテンツを探すことが難しいという点が挙げられたことについて、**指導者用・学習者用のデジタル教科書に埋め込む**形で検索APIを利用することや、**急上昇キーワード**などその時々々の傾向や時期に合わせて結果を返却できるようにするなどの検討が必要である。

今後の課題

取得できるコンテンツに関しても提供するのみでなく、そのコンテンツを基に**生徒自身で考えさせ教員の支援による実践を伴った指導が行えるようなサービス環境を整備**する必要があると考えられる。このためには、有益なコンテンツの充実と併せて検索結果の共有やコンテンツ評価の在り方の検討も必要と考えられる。

今後の課題

利用規約の策定としては利用時の**前提事項や問合せ先の整備、SLAの策定**などが挙げられる。また、システム運用に向けては、**運用設計・保守設計**を行い、運用手順および保守手順を明確化し、適切な運用保守ができるようにする必要がある。

4-3 . コンテンツ検索APIプロトタイプの開発と検証 - 課題（データ連携観点）

③「公共財としてのメタデータ」に関する課題

[ヒアリング指摘]

産学官民の線引き

実証検証にご協力いただいた関係者より、教育デジタルコンテンツ検索基盤およびメタデータ利活用における産学官民の線引きを行うための議論がステークホルダー全体で必要ではないかという評価を頂いた。

今後の課題

メタデータ利活用に向け、産学官民の線引きのためには、**ステークホルダーの合意形成**が必要であると考えられる。そのためには**ワーキンググループなどによるルール検討や役割分担の明確化**が必要となる。

官：ガイドラインを制定し、ガバナンスを効かせる事で、利活用の促進を行う。
産：ガイドラインに従ってコンテンツメタデータのオープンデータ化を進める。
学：実運用を基にした官へのガイドライン制定に対するコメントを行う。
民：教員や生徒がサービス利用を通じて評価やフィードバックを行う。

[ヒアリング指摘]

メタデータ属性の改良

実証参加者から、教育デジタルコンテンツ検索サービスとしてのメタデータ項目としては、検索結果に表示させる内容としては不十分であり、様々な情報を追加した方が良いという意見が多数あった。メタデータ属性の改良と、メタデータの公共性に基づく項目の均質化が課題となる。

今後の課題

省庁関係者、教育関係者向けのヒアリング、およびデータ連携実証事業の関係事業者からのアンケートなどから、**付加価値を高める項目の要望取り込み**が必要になると考えられる。本実証でも以下のような項目の追加が必要でないかという意見が挙げられている。

- ・コンテンツホルダー情報
- ・著作権情報/利用条件
- ・コンテンツ内容についての要旨
- ・分類情報（NDCコード、博物館分類）
- ・コンテンツ評価情報
- ・関連度
- ・教科書掲載情報

また、データベースの**RDF化など共通的なデータフォーマット**を使用し、**項目定義の標準化**を行うことが必要となる。

[本来やるべき]

オープンデータ化

実証検証にご協力いただいた関係者より、相互運用性のあるデータ運用のためのオープンデータ化が必要なのではないかとの意見を頂いている。

今後の課題

メタデータの流通・柔軟な運用に向けて、**LOD化などのオープンデータ規格に準拠**させることが必要となる。オープンデータとして運用する場合には、適切に柔軟な運用を可能とするための検討が必要となる。例として以下のような検討が必要になると考えられる。
・個人情報を含めない形で履歴情報を適切に管理し、外部提供可能とするための仕組みの検討。
・公示バージョンの変更点や保持しているデータの移行対応方法の検討。

4-3 . コンテンツ検索APIプロトタイプの開発と検証 - 課題 (コンテンツ紐づけ観点)

④「紐づけ」に関する課題

[ヒアリング指摘]

コンテンツ紐づけ品質

実際に利用者が教育デジタルコンテンツを利用する場合には、学習指導要領が持つ曖昧さや構造を踏まえた紐づけ品質面の課題がある。

[ヒアリング指摘]

コンテンツ紐づけ評価

民間利用を促進させるためには、教育現場の先生方が納得感を持って使っていただけるシステムにする必要がある。

[ヒアリング指摘]

有益なコンテンツ拡充

「東京ベーシックドリル」などの有益なコンテンツの充実が必要といった意見があった。

今後の課題

実証を通じたヒアリングでは、**学習指導要領の特性（抽象性・階層構造）を考慮した紐づけ方法の改良**が必要との声があった。また、メタデータ自動付与の実証において認識された科目別の特徴（国語・図工・道徳に代表される履修主義科目、算数・英語・理科・社会に代表される修得主義科目、などの学習指導要領コード体系の差異）の考慮を含め、今後紐づけ方法の改良が必要である。

今後の課題

学習指導要領コードと教育デジタルコンテンツの紐づけ結果に対する妥当性検証・評価が必要であり、随時全国の教員の先生方のご意見を踏まえて修正できるようなスキームを設定すべき。
有効なユースケースの明確化が必要であり、教育関係者による実証やデータ連携（学習プラットフォーム）の実証が求められる。

今後の課題

教育デジタルコンテンツとして教育の現場で**有益とされる素材の増補**が必要である。

05

将来に向けた解決すべき課題

5. 将来に向けた解決すべき課題 - 「コンテンツ検索・メタデータ自動付与システムの将来像整理」

5-1 . 課題の整理

ここまでの実証結果報告を踏まえ、改めて実証を通じて認識した課題を以下のとおりグルーピングした。

表:課題一覧

No	カテゴリ	課題	対応策	内容
1-1	サービス	ユースケースの深堀（デジタル教科書・指導書からの教材連携など）	データ連携・データ利活用実証（ユースケースに基づくシステム化の実証）	教育デジタルコンテンツ検索サービスのユースケースと、教育デジタルコンテンツ検索基盤のシステム要件の具体化
1-2			データ連携・データ利活用実証（サービス層のユーザビリティ・アクセシビリティ）	教育デジタルコンテンツ検索サービスのユーザビリティ・アクセシビリティ要件の明確化
1-3			データ連携・データ利活用実証（基盤層の利用規約）	教育デジタルコンテンツ検索基盤の利用規約策定
2-1	データ流通	社会実装に向けた合意形成	ワーキンググループなどによるルール検討	教育デジタルコンテンツ検索基盤およびメタデータ利活用における産学官民の線引き
2-2		メタデータ属性の改良	標準化	メタデータ属性の改良（コンテンツホルダー・要旨・著作権・分類情報・評価情報・関連度、など）
2-3		オープンデータ化	RDF化・LOD化	相互運用性のあるデータ運用のためのオープンデータ化
3-1	コンテンツ紐づけ	コンテンツ紐づけ改善	有識者・教育関係者によるコンテンツ紐づけ評価	学習指導要領コードと教育デジタルコンテンツの紐づけ結果に対する妥当性検証・評価
3-2			コンテンツ紐づけ評価に基づく改善	学習指導要領コードの特性（抽象性・階層構造・科目別の特徴）を考慮した紐づけ方法の改良
3-3			有益コンテンツの増補	教育デジタルコンテンツとして有益な素材の増補
3-4		メタデータ自動付与におけるコンテンツ情報入手に係る課題対応	公開情報からのコンテンツ要旨情報収集の検討	コンテンツホルダーにおけるWeb公開情報からのコンテンツ要旨情報スクレイピング仕様の検討
3-5			コンテンツホルダーによるコンテンツ要旨情報設定に向けた検討	コンテンツホルダーがメタデータ自動付与の際に設定するコンテンツ要旨情報の検討
3-6		メタデータ自動付与の精度改善	メタデータ自動付与の在り方の検討	メタデータ自動付与で提示された学習指導要領コードの候補を、人手（コンテンツホルダー）が最終確認する利用形態における課題の深堀
3-7			コンテンツ要旨情報の改善	コンテンツ要旨情報の改善によるメタデータ自動付与の精度向上
3-8			メタデータ自動付与の改良	発達段階や科目の特徴考慮、コンテンツ種別の特性考慮、類義語や表記揺れ考慮、などによるメタデータ自動付与の改良
3-9		コンテンツ紐づけ改善	メタデータのメンテナンス運用	紐づけの誤り訂正や学習指導要領改訂に伴うメンテナンス手段と運用方法の確立

5-2 . ロードマップ案

5-1で整理した課題をロードマップ案とし、社会実装に向けて解決すべきことを可視化した。
 また、利活用促進に向けて、連携が必要となる様々なステークホルダーや関係省庁との情報共有や方向性の確認なども実施していく必要があると考える。

(凡例)
 課題No : 1-1 ~ 3-9



図:ロードマップ案

5-3 . 社会実装に向けて - サービスの観点

(1) サービス

1. 短期的から中期的な検討項目

・データ連携・データ利活用実証

実証の推進にあたっては、問題点として以下の点があったため、本来のヒアリング項目を十分に深掘することができなかった。

①教育関係者が有益性を感じられるシーンがイメージできていない。

②学習プラットフォーム事業者においてコンテンツ検索サービス提供にメリットが生まれる基盤要件が明確化できていない。

具体的なシステムを提示した上で関係者への意見聴取を行ってきたが、プロトタイプ作り込みの少ない状態であり、他のコンテンツとの連動や一般検索サービスとの差別化を提示することが難しく、今後本事業を展開していく上では、より具体的なユースケースを示した上で業界全体の方向づけをしていく必要がある。

業界全体の方向性の一例としては、協調領域（※）として、国主導で推進すべき領域の識別を明確化した上で、競争領域として民間が自由にサービス化を行い、付加価値を競えるようなプラットフォームを形成できるように基盤仕様の検討を進めるべきである。例えば、教育デジタルコンテンツ検索における基盤の相互運用性を目的とした標準の機能要件や連携仕様の規定を定め、各システムベンダーがそれに準拠しつつ競争するような環境を形作ることも可能と考えられる。

※ … 協調領域・競争領域の考え方は「5-3 . 社会実装に向けて - アーキテクチャ」に示す。

2. 中長期的な検討項目

・利活用基盤整備

学習プラットフォームにおいて、システムベンダーが実際のビジネス化を行う場合には、それを支えるための利活用基盤が前提となる。

このためには、コンテンツ検索・メタデータ自動付与システムを一般向けサービスとして公開していく場合には、基盤としてのサービスレベルを規定することが必要となる。

教育デジタルコンテンツ検索基盤としての標準規定などの共通のルールに従いつつ、国民向けサービスとして確実な運用を行うことが必要であるが、前述の短期的から中期的な検討項目内でも触れているとおり、求められるサービスレベルはユースケースを踏まえ、利用頻度や普及度合いによって段階的にあらたな要求事項に対応することが必要となるため、システム構成だけでなく、将来的な運営母体の検討も含めて、安定運用のための基盤運用と体制、加えて保守や機能改善のための開発体制を実現することが不可欠であると考えられる。

5-3 . 社会実装に向けて - データ流通の観点

(2) データ流通

1. 短期および中期的な検討項目

・ワーキンググループなどによるルール検討、および標準化・オープンデータ化

本調査研究においてはデータ構造は一部標準（メタデータ規格DCAT）に準拠するものとしているが、標準化を進められている規格がなく、基本的に独自の仕様を策定した。ただし、今後新たなコンテンツが提供される際に、コンテンツ提供者によって検索用メタデータが付与され、コードの紐づけ結果の検証が行われるような自己成長型の自律的なエコシステムを形成していくにあたり、学習プラットフォーム関係事業者およびコンテンツ関連事業者・教育関係者・関係省庁など、様々なステークホルダーの参画が不可欠となる。

現時点で構築した各種仕様をデファクトスタンダードとして採用するべきかも含め、これらの関係者に共通的なルール検討・ルール策定が必要であり、以下の点について考慮することが必要である。

- ①合意形成のためのワーキンググループが必要。
- ②メタデータの標準化・オープンデータ化に専門知識が必要。

このために、次のような対応が必要と考えられる。

- ①教育関連の技術標準などの策定・普及を行っており、広く業界や事業者との繋がりがある業界団体と一緒にワーキンググループを検討する。
- ②将来的にはオープンデータ化により学習eポータル、一般検索サービス、ChatGPTといったAIシステムなどによりUIが提供される事により利用用途やシーンが広がる可能性を考慮する。
- ③研究機関などに所属する情報整理学の研究者や、既にメタデータの標準化やオープンデータ化で事例のある組織の有識者を巻き込んで、メタデータの標準化とオープンデータ化（**RDF化・LOD化**）を進める。

2. 中長期的な検討項目

・コンテンツ紐づけメタデータガイドライン

エコシステム形成に向けて、自律的なメタデータ普及を促進するためには、コンテンツ紐づけメタデータのためのガイドラインを作成し、メタデータを作成する事業者に公開することが必要である。

様々な知見や現場の声を反映した内容とする必要があるため、上記で構築したワーキンググループを活用しながら策定を進めることが望ましい。

なお、自発的な登録のためのモチベーションを高めるためには、本システムへの登録がコンテンツ提供者などにとって十分なメリットが感じられる必要があり、一定程度の事業規模や民間でのサービス普及に至るまでは標準策定のための調整に注力せず、次ページで言及するとおり、有益なコンテンツの増補を含めてサービス提供の実例を創出するのの一案である。

5-3 . 社会実装に向けて - コンテンツ紐づけの観点

(3) コンテンツ紐づけ

1. 短期および中期的な検討項目

・有益なコンテンツの増補

利活用を進めるに辺り、利便性をユーザーに感じてもらう必要があるため、**有益なコンテンツの増補**が重要である。教育関係者からの意見として、授業準備の際に利用する「東京ベーシックドリル」などが学習指導要領と紐づいていると便利と言う意見があり、素材となり得るコンテンツホルダーのデータを紐づけることで、授業準備における教員の負荷軽減につながると考える。また、フリーの画像など、教育の現場で有益と認められている素材についても情報を集めた上で、充実を図ることが望ましい。

・有識者・教育関係者による紐づけ評価と改善

本事業で試用するための土台が整ったため、今後は**有識者・教育関係者による紐づけ評価と改善**が重要である。今回は一定のルールを基に紐づけを実施したが、本来は文科省や業界団体を含めた形で議論を行い、決めていくべきとの意見もあった。データ利活用の観点から、検索APIで扱うコンテンツの量や質も重要となるため、引き続き議論とメタデータへのフィードバックが必要になる。また、学習指導要領コードの改訂により、変更点の差分をどのように取り込むかなども検討が必要になる。

さらに、コンテンツ情報の入手に係る課題として、コンテンツ要旨情報の確実な入手をどうやって担保するかという点や、コンテンツホルダーによるコンテンツ要旨設定に係る課題が抽出できていないという点についても考慮が必要であり、次のような検討が必要である。

- ①コンテンツ要旨情報の配置や変化を許容する運用をどう行うべきか
- ②コンテンツホルダーを巻き込んだコンテンツ要旨設定に係る課題の抽出

・メタデータ自動付与実証

4-2節で記載した精度向上や運用設計などの課題がある。代表的なものとして、スクレイピングを行うにあたって、コンテンツホルダーに対してガイドラインの公開をするか、AI自体のインプットに要旨情報を入力できるようにし、AIの使用方法を公開する必要がある。

3. 中長期的な検討項目

・学習指導要領コード付与支援ツール整備

学習指導要領コード付与支援ツールを整備することで、学習指導要領コードとコンテンツとの紐づけをAIにより実施可能とし、人手よりもコストが掛からず、ガバナンスを効かせた状態での紐づけが可能となるため、コンテンツとの紐づけにコストが掛かるから参画を躊躇するコンテンツホルダーの参画が期待され、よりコンテンツの利活用が促進される未来に繋げることが可能となる。

5-3 . 社会実装に向けて - アーキテクチャ

(4) 望ましい将来に向けたアーキテクチャ

・本事業で得られた知見を踏まえ、「仕組み」・「ユースケース」・「社会実装」の3つの側面から望ましい将来に向けたアーキテクチャのイメージを示す。

- 1 ユースケースの在り方**
 - ・**教員、児童生徒のメリット**
「誰でもいつでも、個々の学びに役立つコンテンツが探せる。」
 - ・**教育デジタルコンテンツ事業者のメリット**
「公共オープンデータや民間のデジタル教科書や学習教材などへ、学習指導要領コードが付与し易くなり、利用者による検索によりコンテンツ活用およびビジネスが進展。」
- 2 アーキテクチャの在り方**
有益な教育デジタルコンテンツを探すためのアーキテクチャを3つの階層「デジタルサービス・サービス基盤・データ」として表現。
- 3 準公共分野（教育）の社会実装の在り方**
国の関与により民間分野への波及効果を引き出し事業活性化させることで、**利用者ニーズに応じた最適サービス提供**に繋げる狙い。
・そのための**協調領域・競争領域の識別と、産官学民の関係者の合意形成**が重要。

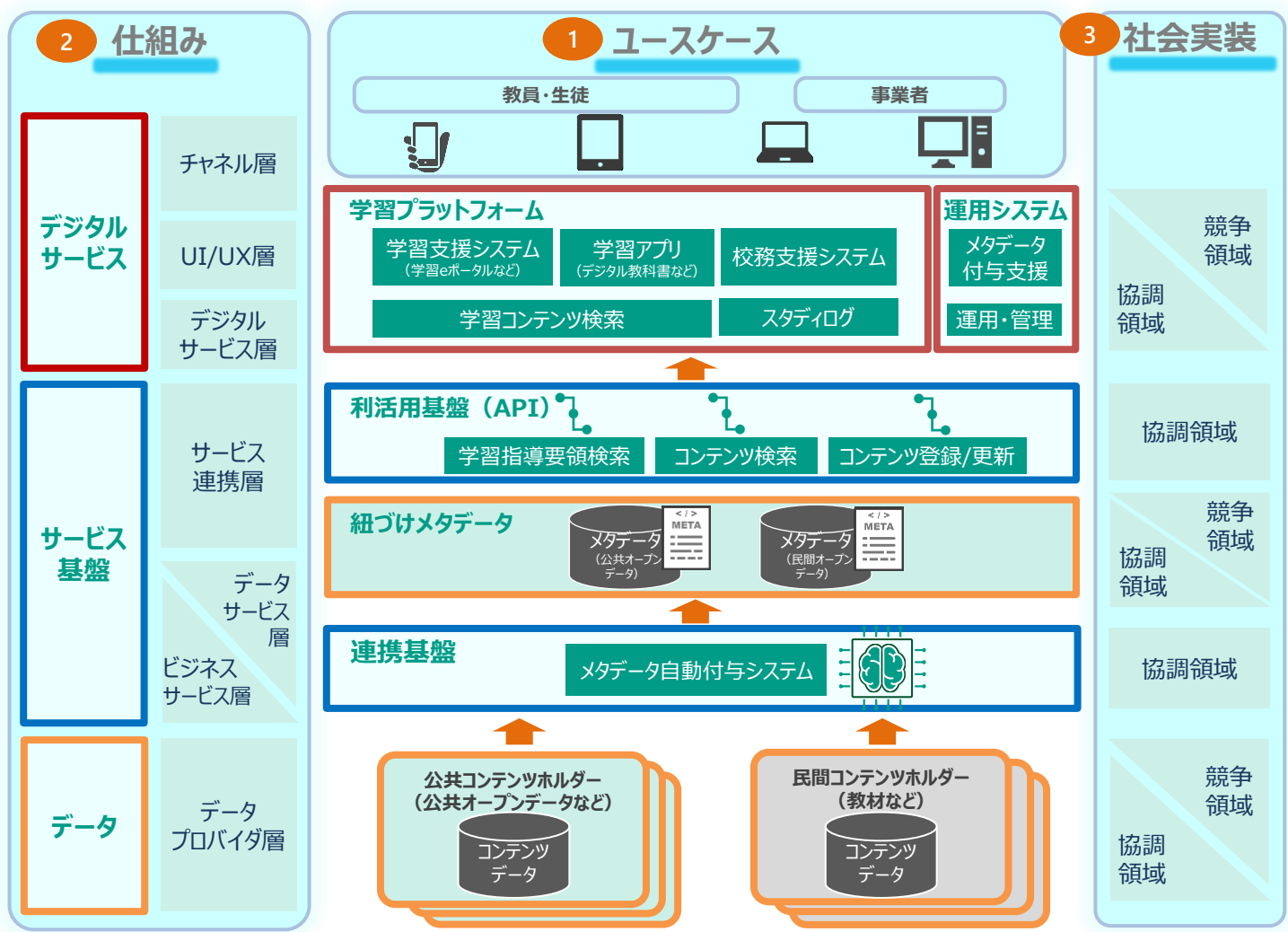


図:アーキテクチャのイメージ

用語 (1/4)

報告書に記載する用語の定義を「表:用語の定義」に示す。

表:用語の定義

No	用語/略称	定義 (又は正式名称)
1	AI	人工知能 (Artificial Intelligence) の略称。コンピュータの性能が大きく向上したことにより、機械であるコンピュータが「学ぶ」ことができるようになり、それが現在のAIの中心技術、機械学習である。
2	API	Application Programming Interfaceの略称であり、「ある1つの機能に特化したプログラムで共有可能なもの」や「ソフトウェアの機能を共有する仕組み」のことである。
3	AWS	Amazon Web Servicesの略で、Amazonが提供しているクラウドコンピューティングサービスの総称。コンピューティング、ストレージ、データベースなどのインフラストラクチャテクノロジーから機械学習、AI、データレイクと分析、IoTなどの最新鋭のテクノロジーに至るまで多くのサービスを提供しており、それぞれのサービスが豊富な機能を備えている。
4	DCAT	データセットとデータ・サービスを記述してカタログに含めることができるように、RDFのクラスとプロパティを提供する仕様であり、W3Cから勧告されWeb技術の標準となっている。
5	eboard	学習サイトの運営やICTを活用した学習の支援などを行う東京都渋谷区の特定非営利活動法人である。
6	GIGAスクール構想	児童生徒向けの1人1台端末と、高速大容量の通信ネットワークを一体的に整備し、多様な子どもたちを誰一人取り残すことなく、公正に個別最適化された創造性を育む教育を、全国の学校現場で持続的に実現させる構想。GIGAとはGlobal and Innovation Gateway for Allの略である。
7	HTTP	「Hypertext Transfer Protocol」の略で、WebサーバーとWebブラウザの間で、Web情報をやり取りするためのプロトコル (通信規則) である。
8	IaC	Infrastructure as Codeの略称 (サーバーなどのインフラ構築を、コードを用いて行うこと。人的ミスやコストの削減が可能) である。
9	ICT	Information and Communication Technologyの略称で、通信技術を活用したコミュニケーションを指す。
10	JAPAN SEARCH	書籍・公文書・文化財・美術・人文学・自然史/理工学・学術資産・舗装番組・映画など、我が国が保有する様々な分野のコンテンツのメタデータを検索・閲覧・活用できるプラットフォーム (基盤となる装置やソフトウェア、サービス、あるいはそれらを組み合わせた動作環境のこと) である。
11	LOD	Linked Open Dataの略称。Web上でコンピュータ処理に適したデータを公開・共有するための技術の総称であり、Web上に存在する他のデータとリンクしていることと、誰でも自由に利用できるオープンなライセンスで公開されているということを兼ね備えているデータのこと。
12	NHK for School	学校放送番組やWebサイト、アーカイブス、イベントなど、学校向けの様々なサービスを展開している。
13	OSS	オープンソースソフトウェア (Open Source Software) の略称 ソースコードが公開されており、無償で誰でも自由に改変、再配布が可能なソフトウェアである。
14	RDF	一定の書式で情報についての情報 (メタデータ) を記述することで、情報の持つ属性や意味に基づいた収集や検索、整理、分類などの論理的な操作をソフトウェアによって自動的に行うことができるようにすることを目指したデータフォーマットの1つであり、W3Cから勧告されWeb技術の標準となっている。
15	REST	REpresentational State Transferの略称で、 下記の原則を満たす設計思想。 ①統一インターフェース (通信方法ややり取りするデータ形式などが統一されていること) ②アドレス可読性 (すべての情報が一意なURIを持つこと) ③接続性 (やり取りされる情報にハイパーリンクを含めることができること) ④ステートレス性 (やり取りが前後に影響しないこと)

用語 (2/4)

No	用語/略称	定義（又は正式名称）
16	STEAM Library	経済産業省がSTEAM教育を推進・確立していくために始めたコンテンツ（教材）を検索・閲覧・視聴、資料ダウンロード可能なオンライン図書館。STEAMとは、Science（科学）、Technology（技術）、Engineering（工学）、Art（芸術）、Mathematics（数学）の5つの英単語の頭文字を組み合わせた造語で、5つの領域を対象とした理数教育に創造性教育を加えた教育理念のこと。
17	UI	ユーザーインターフェースの略称。インターフェースには、本来「接点、接触面」という意味があり、ユーザーとコンピュータ間で情報をやり取りする様々な機器や入力装置を指す。
18	UX	ユーザーエクスペリエンスの略称。製品やシステム、サービスなどの利用を通じてユーザーが得る体験を表す。
19	アーキテクチャ	コンピュータやソフトウェア、システム、あるいはそれらの構成要素などにおける、基本設計や共通仕様、設計思想などを指す。
20	アクセシビリティ	アクセスのし易さのこと。情報やサービスなどがどれくらい利用し易いか、特に障害者や高齢者などが不自由なく利用可能かどうかの度合いを示すもの。
21	アルゴリズム	ある特定の問題を解く手順を単純な計算や操作の組み合わせとして明確に定義したもの。
22	エコシステム	元来、生態系を表す用語。同じ領域で暮らす生物や植物が、お互いに依存しながら生態系を維持している仕組みがエコシステム。エコシステムという自然界の仕組みをビジネス業界に当てはめて、業界同士や製品、サービス、プロダクトなどが連携し、大きな利益構造を構成することを指す。
23	クラウド	ストレージデータを格納するためインターネット上に設置されたスペース。
24	クラウド・バイ・デフォルト	政府の「ガバメント・クラウド実行計画」に定める、「政府情報システム化においてクラウド活用を第一候補」とする基本方針である。
25	コンテンツホルダー	コンテンツの所有者、即ち、文章や映像などの情報（コンテンツ）に関する著作権などの諸権利を持つ者のこと。
26	スクレイピング	Webサイトから情報を抽出するコンピュータソフトウェア技術のこと。
27	ステークホルダー	「企業の活動に対して、直接的・間接的な利害関係を有するグループ又は個人」を指すビジネス用語である。
28	対応表	<p>本件の業務内容である「学習指導要領コードと個々のデジタル書籍・デジタル画像・デジタル動画との対応付け調査」の結果として整理した成果文書であり、2017年から2019年にかけて告示された学習指導要領のうち、学習指導要領コードに対応させる形で、指導用又は学習用に参考となるデジタル書籍（電子書籍）を、国立国会図書館サーチ、国立国会図書館デジタルコレクション又はSTEAM LibraryやJAPAN SEARCHなどにより検索可能なデジタル書籍・デジタル画像・デジタル動画の中から抽出し、学習指導要領コードと対応するデジタル書籍・デジタル画像・デジタル動画との対応関係を人手による手作業で一覧表にまとめたもの。対応表で採用した学習指導要領コードの公示バージョンと、対応表のファイル構成は以下のとおりである。</p> <p>[紐づけ作業対象学習指導要領コード公示バージョン]</p> <ul style="list-style-type: none"> ・小学校学習指導要領コード：バージョン1.2 ・中学校学習指導要領コード：バージョン1.1 ・高等学校学習指導要領コード：バージョン1.0 ・特別支援学校小学部・中学部学習指導要領コード：バージョン1.2 ・特別支援学校高等部学習指導要領コード：バージョン1.2 <p>[構成]（対応表ファイル数）</p> <ul style="list-style-type: none"> ・デジタル書籍：54 ・デジタル画像：54 ・デジタル動画：3

用語 (3/4)

No	用語/略称	定義 (又は正式名称)
29	データ連携実証事業者	デジタル庁様の別プロジェクトである「教育における広範なデジタルコンテンツを含む教育関連データの連携実証」の受託事業者である。
30	デジタル教科書	学習者用デジタル教科書は学校教育法第34条第2項に規定する教材。紙の教科書と同一内容。児童生徒が使用するものであり、指導者用デジタル教科書 (デジタル教材) とは異なる。
31	デジタル庁	デジタル庁 国民向けサービスグループ 教育班である。
32	バックキャストिंग	中長期的な視点で未来の「望ましい将来像」を仮定し、それを実現するための方法を考えていくアプローチであり、不確実性が高く正解が存在しない課題やテーマに対する解決策の検討に適した手法。本調達では、将来像の整理や取り組む課題の抽出にバックキャストिंग (バックキャスト手法) を採用する。
33	プロトタイプ	原型、試作品のこと。ITの分野では、ハードウェア開発の際の量産前の試作品や、動作や機能を検証するために最小限の規模で試作されたソフトウェアなどのことを指す。 本紙では、特に調達仕様書に示されている「コンテンツ検索システム/メタデータ自動付与システムのプロトタイプ版」を指す。本調達にて作成するアプリケーション、AIモデルを組み合わせたものが、プロトタイプ版のシステムとなる。
34	メタデータ	群本体であるデータに関する付帯情報が記載されたデータの群である。
35	メタデータ自動付与システム	指定されたデジタルコンテンツ (書籍、画像、動画) をテキストデータ化し、機械学習プログラムによって学習指導要領コードなどのメタデータを付与するシステムである。機械学習の教師データには、本調査事業で得られたデジタルコンテンツと学習指導要領コードとの対応付け情報 (対応表) を用いている。
36	学習eポータル	GIGAスクール構想で整備された1人1台環境と高速ネットワークを活かし、ソフトウェア間の相互運用性を確立してユーザーにとっての操作性を向上させるとともに、教育データをより良く活用するために構想された、日本の初等中等教育向けのデジタル学習環境のコンセプトである。
37	学習アプリ	スマートフォンやタブレット型端末向けに提供される学習用アプリケーションソフトの総称である。
38	学習指導要領コード	学習指導要領コード (Curriculum Standards)。「学習指導要領」とは、全国どここの学校でも一定の水準が保てるよう、文部科学省が定めている教育課程 (カリキュラム) の基準。学習指導要領の冒頭から順番に16桁のコードを割り振る。その際、学校種、教科、学年などの検索が容易となるように桁に一定のルールを設ける。
39	学習支援システム	学習支援システムは、学習eポータルなどのICTを活用してより効果的な学習を行うためのシステムである。
40	形態素解析	自然言語で書かれた文を言語上の最小単位である形態素に分割し、それぞれの品詞や変化などを割り出すこと。ITの分野ではコンピュータによる自然言語処理の1つとして、かな漢字変換や全文検索、機械翻訳などで用いられる。
41	校務支援システム	教務系 (成績処理、出欠管理、時数など) ・保健系 (健康診断票、保健室管理など) 、指導要録などの学籍関係、学校事務系など統合した機能を持つシステムである。
42	国立国会図書館サーチ	国立国会図書館をはじめ、全国の公共・大学・専門図書館や学術研究機関などが提供する資料、デジタルコンテンツを統合的に検索できる。
43	国立国会図書館デジタルコレクション	国立国会図書館で収集・保存しているデジタル資料を検索・閲覧できるサービスである。
44	国立歴史民俗博物館	日本の歴史と文化について総合的に研究・展示する歴史民俗博物館である。

用語 (4/4)

No	用語/略称	定義 (又は正式名称)
45	単元	学習によって得られる教科内容又は経験のひとまとまり。単に教材内容の一分節をいうこともある。
46	東京国立博物館	日本と東洋の文化財の収集保管、展示公開、調査研究、普及などを目的として独立行政法人国立文化財機構が運営する、日本の国立博物館である。
47	紐づけメタデータ	「学習指導要領コード」と「学習コンテンツ」の紐づけ情報のデータであり、対応表の情報からシステムが扱えるようにデータ構造を考慮してデータ化したもの。
48	ユーザビリティ	ユーザーがそのシステムの使用方法などを理解して、快適かつ効率的に使えること。一般的に「使いやすさ」「使い勝手」を意味する。
49	ユースケース	利用者があるシステムを用いて特定の目的を達するまでの、双方の間のやり取りを明確に定義したもの。利用者は機器を操作する人間以外にも外部の他のシステムなどを想定する場合もある。

運営委員会の開催状況 (1/2)

報告書に記載した運営委員会の開催状況を示す。

表:運営委員会の開催状況

凡例：◎：主催者、○：出席者

No	会議体	議題	開催時期	会議形式	出席者			
					デジタル庁	富士ソフト	有識者	オブザーバー
1	第1回 運営委員会	・将来像の検討 ・学習指導要領コード紐づけ調査方法について	2022年9月9日	Teams会議	○	◎	○ (2名)	○ (5名)
2	第2回 運営委員会	・学習指導要領コード紐づけ準備作業の中間結果報告	2022年10月7日		○	◎	○ (2名)	○ (8名)
3	第3回 運営委員会	・学習指導要領コード紐づけ準備作業の結果報告 ・学習指導要領コードの紐づけ結果の中間報告 ・実証実験に係わるアンケートについて	2022年11月11日		○	◎	○ (2名)	○ (10名)
4	第4回 運営委員会	・学習指導要領コードの紐づけ結果の報告 ・メタデータ自動付与結果の共有と意見聴取 ・検索APIの利用イメージを画面ベースで共有	2023年1月20日		○	◎	○ (2名)	○ (4名)
5	第5回 運営委員会	・メタデータ自動付与結果の報告 ・実証結果の報告	2023年3月16日		○	◎	○ (2名)	○ (8名)

運営委員会の出席者構成を以下に示す。

表:運営委員会の出席者構成

No	所属
1	デジタル庁
2	富士ソフト株式会社
3	有識者 ※詳細は、「表:運営委員会の有識者構成員」に示す。
4	オブザーバー ※詳細は、「表:運営委員会のオブザーバー」に示す。

運営委員会の開催状況 (2/2)

運営委員会の有識者構成員を以下に示す。

表:運営委員会の有識者構成員

No	所属	役職	氏名	専門分野
1	東京学芸大学	教授	加藤 直樹	情報教育
2	東京学芸大学	教授	北澤 武	教科教育

運営委員会のオブザーバーとして、教育分野に携わる各省庁の担当者と

「CBTシステム(MEXCBT)の拡充・活用推進、教育データの利活用推進事業(学習指導要領コードの利活用に関する調査研究事業)」の受託事業者に陪席していただいた。

参加したオブザーバーを以下に示す。

表:運営委員会のオブザーバー

No	所属
1	総務省
2	経済産業省
3	文部科学省
4	「CBTシステム(MEXCBT)の拡充・活用推進、教育データの利活用推進事業(学習指導要領コードの利活用に関する調査研究事業)」の受託事業者