

第6回データ利活用制度・システム検討会

(開催要領)

1. 開催日時：令和7年3月4日（火）14:00～16:00
2. 場所：デジタル庁中会議室
3. 出席：

安中	良輔	日本製薬工業協会産業政策委員会健康医療データ政策 GL
生貝	直人	一橋大学大学院法学研究科教授
依田	高典	京都大学大学院経済学研究科教授
稲谷	龍彦	京都大学大学院法学研究科教授
中嶋	康	一般社団法人日本経済団体連合会産業技術副本部長（岩村構成員代理）
上野山	勝也	株式会社 PKSHA Technology 代表取締役
岡田	淳	森・濱田松本法律事務所外国法共同事業パートナー弁護士
落合	孝文	渥美坂井法律事務所・外国法共同事業プロトタイプ政策研究所 所長・シニアパートナー弁護士
宍戸	常寿	東京大学大学院法学政治学研究科教授
丹野	美絵子	公益社団法人全国消費生活相談員協会消費者情報研究所消費生活 専門相談員
森田	朗	一般社団法人次世代基盤政策研究所所長・代表理事

<ゲストスピーカー>

- | | | |
|----|----|---|
| 江崎 | 典宏 | 文部科学省大臣官房審議官 |
| 小崎 | 誠二 | 奈良教育大学大学院教育学研究科准教授／奈良市教育委員会ス
クール DX プロジェクトマネージャー |
| 佐藤 | 潤 | 株式会社 COMPASS 代表取締役 CEO |
| 蓮井 | 智哉 | デジタル庁戦略・組織グループ審議官 |
| 藤村 | 裕一 | 鳴門教育大学大学院特命教授／鳴門教育大学教員養成 DX 推進
機構機構長 |

<事務局>

- | | | |
|----|----|----------------------------|
| 小川 | 康則 | デジタル庁財政改革会議事務局長代理 |
| 村上 | 敬亮 | デジタル庁財政改革会議事務局長補佐／デジタル庁統括官 |
| 山澄 | 克 | デジタル庁財政改革会議事務局審議官 |
| 吉田 | 宏平 | デジタル庁財政改革会議事務局審議官 |
| 木尾 | 修文 | デジタル庁財政改革会議事務局参事官 |

飯嶋	威夫	デジタル行財政改革会議事務局参事官
坪井	宏徳	デジタル行財政改革会議事務局参事官
中野	芳崇	デジタル行財政改革会議事務局企画官
楠目	聖	デジタル行財政改革会議事務局企画官

<オブザーバー>

瀬戸口	文博	公正取引委員会事務総局経済取引局調整課長補佐
吉屋	拓之	個人情報保護委員会事務局参事官
松澤	瞭	デジタル庁政策・法務ユニット法務スペシャリスト
黒藪	誠	デジタル庁統括官付参事官
八代	将成	総務省情報流通行政局地域通信振興課デジタル経済推進室長
守谷	学	経済産業省商務情報政策局情報経済課長

(議事次第)

1. 開会
2. 議事
 - (1) 教育分野におけるデータの利活用について
 - (2) AIとデータの利活用について
3. 閉会

(資料)

- 資料1 文部科学省提出資料
- 資料2 デジタル庁提出資料
- 資料3 小崎誠二奈良教育大学大学院教育学研究科准教授／奈良市教育委員会スクールDXプロジェクトマネージャー提出資料
- 資料4 佐藤潤株式会社 COMPASS代表取締役CEO提出資料
- 資料5-1 藤村裕一鳴門教育大学大学院特命教授／鳴門教育大学教員養成DX推進機構機構長提出資料
- 資料5-2 藤村裕一鳴門教育大学大学院特命教授／鳴門教育大学教員養成DX推進機構機構長提出資料
- 資料6 上野山構成員提出資料
- 資料7 参加者名簿

(概要)

○森田座長 皆さん、こんにちは。ただいまより、第6回「データ利活用制度・システム検討会」を開催いたします。

まずは事務局から本日の進行についての御説明をお願いいたします。

○山澄審議官 事務局でございます。

本日は教育分野に分野におけるデータ利活用に関しまして、有識者の皆様からお話をいただいた後、構成員の皆様にご意見を伺いたしたいと思います。

また、後半30分間でございますが、それとは別にAIのデータ利活用をテーマに上野山構成員から御発表いただきまして、こちらにつきましても御議論いただきたく存じます。

資料はタブレットに格納しておりますので、不具合等がございましたら、お申しつけください。

本日、安中構成員、生貝構成員、依田構成員、稲谷構成員、上野山構成員、岡田構成員、落合構成員、宍戸構成員がオンラインで御参加いただいております。また、岩村構成員の代理といたしまして、中嶋日本経団連産業技術副本部長もオンラインで御参加でございます。阿部構成員、越塚構成員、巽構成員は御欠席となっております。

以上でございます。

○森田座長 ありがとうございます。

それでは、早速ですが、「教育分野におけるデータ利活用について」の議事に入りたいと思います。まずは江崎典宏文部科学省大臣官房審議官から、文部科学省の取組状況につきまして御紹介をお願いいたします。よろしくをお願いいたします。

○江崎審議官 文部科学省大臣官房審議官、総合教育政策局担当の江崎と申します。

それでは、資料1に基づきまして御説明したいと思います。

まず、1ページをおめくりいただきまして、教育DXの実現に向けて、児童生徒1人1台端末環境というものが整備されております。誰一人取り残すことなく全ての子供たちの力を最大限引き出せるように、教育データをいかに効果的に活用するかということが必要になってまいります。

そのため、文部科学省におきましては、国が行う取組というところが3つありますが、教育データをツール間で相互に交換・蓄積・分析するため、相互利用性を確保するためのデータ内容、それから技術の標準化、ツール間の接続規格策定などの共通的なルールの整備、それから、真ん中になりますが、児童生徒の学びの保障の観点から、学習やアセスメントができる文部科学省CBTシステム「MEXCBT」、それから、国からの調査などができる文部科学省WEB調査システム「EduSurvey」といったものをツールとして開発し、また、活用をしているということでございます。

それから、右のほうですけれども、教育データの分析・活用手法について、実証研究、また、横展開をしているような状況でございます。また、個人情報保護の観点からの教育委員会や学校における留意点の周知、それから、教育データの活用に向けた取組といったものを行っております。

これらの取組を通じまして、教育現場のきめ細かな指導、行政における政策・施策改善等をより効果的に行い、さらには日本全体の教育水準の向上に役立てるということを考え

ております。

その次の2ページでございます。こちらで1つ事例を御紹介いたしますけれども、様々な教育委員会、あるいは学校で先進的な取組というのが幾つかございます。この2ページ目は渋谷区の教育委員会での取組でありますけれども、教育ダッシュボードというものを構築しております。渋谷区教育委員会におきましては、令和4年度より教員がこのデータに基づいた指導・支援、それから子供たちの学校満足度の向上のため、欠席・遅刻・保健室情報、学校生活アンケート、心の天気情報のようなデータを総合的に教育ダッシュボードに集約いたしまして、先生方が多面的に把握することができるということによりまして、子供の実態をより正確に捉えた指導・支援というものにつなげていっているということでございます。

次に、3ページになります。これはもう一つの事例ですけれども、福岡市教育委員会における教育データの分析システムの構築というものです。福岡市教育委員会におきましては、蓄積された複数の研究データの相関関係といったものから子供の学力の変化や心の状態の変化、これらの要因や傾向を分析するシステムの試行、それからその検証を行っておりまして、効果的な指導や教育政策の立案のために教育データを利活用するという取組を進めてございます。

4ページになりますけれども、このような意欲ある、それから先進的な教育委員会や学校におきまして、教育データを活用して児童生徒への指導や支援等に生かす先進的な取組が進められたという一方で、教育データの利活用についての意義の共有やデータリテラシーなどの課題から、教育データの活用状況というのは自治体ごとにかなり異なっておるという状況でございます。

また、デジタル行財政改革取りまとめ2024におきまして、教育におけるシステム間のデータ連携の課題などによりまして、教育データ利活用が全国的な動きになっていないことも指摘をされています。こうした課題や御指摘を踏まえ、文部科学省といたしましては、2月20日に公表した教育データの利活用に関する有識者会議の取りまとめに基づきまして、政府全体のデジタル基盤の検討状況等も注視しつつ、自治体、民間企業との役割分担の下、教育システムや学習リソース間の円滑な連携を支える標準・指針等の策定・更新を行いまして、各自治体の実態やニーズに応じて学習リソース等を選択・活用できるような支援をさらに進めてまいりたいと考えてございます。

加えて、先進事例の紹介等を行う教育委員会向け研修など、教育データ利活用の意義の周知、必要性や有用性の認識共有などのためにそのような取組も行いまして、教育データの利活用をさらに推進していきたいと考えてございます。

このほかでありますけれども、その後、デジタル庁よりも御説明があると伺っておりますけれども、認証基盤の在り方の検討につきましてもデジタル庁と密に連携して検討していきたいと考えております。

文科省としましては、御紹介いたしましたこれらの取組によりまして、デジタルを活用

した教育について推進してまいります。

説明は以上でございます。

○森田座長 御報告ありがとうございました。

続きまして、蓮井智哉デジタル庁戦略組織グループ審議官から、デジタル庁の取組状況につきまして御紹介をお願いいたします。よろしくお願いいたします。

○蓮井審議官 御紹介いただきました、デジタル庁の蓮井でございます。

それでは、デジタル行財政改革の取組も含めまして、デジタル庁で行っている教育分野の認証基盤に関する検討について御紹介したいと思います。

2 ページ目を御覧ください。デジタル行財政改革会議では、先ほどもございましたけれども、教育分野のデジタル化につきまして、学校教育活動の充実、先生方の負担の軽減につながるように、効率的・効果的な端末やシステムの導入、教育データの利活用などについて検討を行い、段階的に対応を進めているところでございます。

左側の欄にございますように、端末やシステムの導入・活用や通信ネットワーク整備につきましては自治体間の格差が大きいなどの課題がございまして、昨年6月のデジタル行財政改革取りまとめ2024におきまして、共同調達スキームによる効率的な端末システムの整備や通信ネットワークの改善などを文部科学省を中心に進めているところでございます。

今後目指すべき姿が右にございますけれども、それに向けてさらに検討を深めることとしまして、資料中央の列の下段にあります、児童生徒や保護者が安心してデータ連携を行うための認証基盤の在り方や、これは関係省庁と策定しておりますけれども、教育データ利活用ロードマップの改定といった課題があると認識をしております。

このような経緯を踏まえまして、本日は教育分野の認証基盤の整備等に関する検討の状況について御説明いたします。

3 ページ目でございますけれども、こちらは現時点での課題認識について記載しております。教育のデジタル化が地域によって言わばばらばらに進んでいることから、年度末や年度初めにアカウントの更新作業で教職員が多大な作業時間を取られておりまして、デジタル完結やワンスオンリーの徹底が必須と考えてございます。

また、自治体や学校ごとに異なるアカウント管理が行われておりますけれども、自治体の情報システム担当の人員も少ないということでございますので、自治体だけで標準化に取り組むのは非常に難しゅうございます。認証基盤を特定のプラットフォームに依存せず、自治体の枠組みを超えて国全体で共通の方針を策定するということを考えております。

それに基づきまして具体的な中身に入っておりますが、4 ページ目でございます。以上を踏まえまして、教育分野の認証基盤の在り方につきましては、先ほどもございました文部科学省の有識者会議の提言を踏まえまして、デジタル庁において検討会を立ち上げたところでございまして、その検討結果を教育データ利活用ロードマップに反映するとともに、来年度に調査研究を実施していく予定でございます。

以下、具体的な検討の視点について御説明したいと思いますけれども、5 ページ目は教

育分野のデータ連携のパターンを整理したものでございます。御覧いただきますように、1つ目に、法令などに基づきまして学校同士などの組織起点で連携されるデータと、②にございますが、学習者本人の意思に基づきまして転校先・進学先、あるいは就職先、さらには塾など、学校以外の組織との間で連携されるデータの大きく2つに分類できるのではないかと考えてございます。そのために、学校の組織認証や学習者・保護者の個人認証が必要になると認識してございます。

6ページ目でございますけれども、そのうち組織起点のデータ連携において想定されるデータでございますが、例えば指導要録や健康診断票などが想定されているところでございます。

続きまして、7ページ目でございますが、こちらは個人起点での連携が想定されるデータでございますが、卒業証明などの学修証明、通学・在学証明書などは現状、紙でも発行されてございまして、標準化が整うことによって技術的には早期に連携が可能と考えてございます。一方で、学校健診情報や学習履歴などは提供ルールの在り方なども含めて広く議論することが必要かと考えてございます。

8ページ目以降でございますけれども、今まで議論した教育分野のデータ連携の対応を今後進めるに当たって取るべき3つの段階やステップにつきまして載せているものでございます。

まずステップ1は、3ページ目で御説明したようなばらばらなアカウント管理によって生じる課題を解消するために、これまで実証してきた標準規格などによりまして、児童生徒や教職員のID、すなわちアカウントの管理を適切に行うことが必要だということを書いているものでございます。それぞれではなくて、右に「HUB機能」とございますけれども、こういったところで適切なアクセス制御を実施しながら様々なコンテンツにアクセスしていくという形がまず考えられるということでございます。

続きまして、ステップ2でございますが、9ページにございますように、アカウント管理がばらばらに進みますと、自治体を越えた進学や転校におけるデータの連携が難しくなるとして、教育委員会以外の主体に対する手続のデジタル完結も困難になってしまうということでございます。学習者が学校を超えて自らのデータを活用して自主的に学んだり、教員からの確かなサポートを得たりする上でも認証基盤の共通化というのが課題になると考えてございます。

その際、システムやアプリ自体を統一するということではなくて、将来的にデータ連携を共有する範囲を見越しながら、認証基盤の共通化やデータ自体の標準化を進めていくことが2ステップ目として必要となると考えてございます。

最後の3つ目のステップでございますけれども、10ページ目でございますが、人口減少社会において行政サービスの質を維持・向上していくという大きな課題との関係で考えましても、個人を中心とした教育分野に限らない分野横断のデータ連携の実現に向けた基盤整備などを行うことが現状求められていると考えてございます。医療、防災、雇用、福祉

などの他分野との連携を見越した認証の枠組みの整備が必要になると考えてございます。

最後にまとめになりますけれども、11ページ目は今後の取組の方向性を記載しているところでございます。この認証基盤の構築に当たりましては、教育分野では特に安定的な基盤整備や運用が求められること、また、GIGAスクール構想の進展状況や次世代校務DX環境の整備などを勘案しますと、既に利用実績がある技術を用いましてスピード感を持って整備していくことが重要と考えてございます。

具体的には、組織認証につきましては左側でございますけれども、2020年の運用開始から安定的に運用されていて既に120万を超えるアカウントを払い出しているG Biz IDというものがございます。こういったものの活用、それから、個人認証につきましては右側でございますが、厳格な本人確認手段として9600万を超える国民が保有しているマイナンバーカードといったものを活用した公的個人認証、いわゆるデジタル認証アプリでございますけれども、こういったものの活用が候補として考えられると思っております。

今後、有識者の方々に御助言を賜りながら検討を進めてまいりたいと考えてございます。以降は参考でございますので、説明は割愛いたします。

以上でございます。

○森田座長 ありがとうございます。

続きまして、ゲストスピーカーといたしまして、小崎誠二奈良教育大学大学院教育学研究科准教授・奈良市教育委員会スクールDXプロジェクトマネージャーから御発表をお願いしたいと思います。小崎先生、よろしゅうございますでしょうか。よろしくお願ひいたします。

○小崎准教授 よろしくお願ひいたします。

私は今日は、これまで奈良県域で自治体を越えてアカウントを共同運用してきた成果と中身について簡単に御紹介さしあげたいと思います。よろしくお願ひいたします。

今、現時点でGIGAスクール構想をきっかけに子供たちと先生に1人1アカウントを付与していろいろなことを取り組んできたその成果を踏まえて、国レベルで何をしてほしいかと考えたときに、大きく決断してほしいことと、ナッジングとして自然にそうなるような流れをつくってほしいということがあります。具体的には、国でやることは何か、広域の自治体である都道府県でやることは何か、基礎自治体では何をすべきかということで、どちらでもできることがあったときには、分担が明確化されるとそれぞれでできることがあるなと感じています。それから、やってみないと分からないことというのも結構ありますので、チャレンジする自治体がチャレンジする機会を与えてもらえたらいいなと感じています。あと、横並びで同じ階層にいる人たちで手を結ぶというのはなかなか調整が難しかったり届かなかったりするので、1つ上のレイヤーのところ、下のレイヤーをつないでいくということが大事で、今は、県で市町村をつなぐということができはじめていますので、ぜひ次は、国が都道府県を越えてつなぐということについて取り組んでいただきたい。

それから、先生と子供たちの視点というのがこういう議論においては含まれてはいるのですけれども、結局、先生と子供達へ届かないような議論になってしまうと、せっかく準備して実行しようとしたときに、物事が全く動かないということを経験しましたので、その辺りも、誰にでも分かる、という資料を一つだけ作るという方法もありますし、それぞれの立場で見たら分かりやすい複数の資料なり提案というのができたらいいのかなとも感じています。

ICTというと、学校教育の場合はイコール「パソコン」になってしまっていて、手元にあるパソコン、コンピュータをどうやって使うのか、それで何ができるのか、というところに陥りがちなのですけれども、今度の次期学習指導要領の中のコンセプトにも入ってくるでしょうが、そもそも現在ICTと言われているものがコミュニケーションを取る手段のためのテクノロジーだよねということを超えて、知力であったり、主体性であったり、好奇心・想像力を引き出すものであったり、トランスフォーメーションしてDXだよねという解釈で物を考えていくのがスムーズだよねと、今は、そういう理解で進み始めています。

奈良県において、全員が共通のアカウント・IDとして、1人が確実に1つずつのナンバーを持っているということで、例えば、平城宮跡のような学校外で教育イベントをしようとしたときに、大学生、高校生、小中学校、幼稚園の子供たちも集まって、参加確認やデジタルスタンプラリーなどいろいろなことにそのアカウントを活用することができます。歴史学習の発表会、都づくり、それから産業の支援、地域創生という産地学官連携のイベントや取組というのを、奈良ではチャレンジしようとして、小さい企画から始めているのですけれども、こういうときにもデジタル学習基盤があって共通で利用できることとか、地元の人たちと交流していきたいというときに、子供たちも先生も自分のIDをわざわざ付与しなくても持っているということは、非常に大きいのです。子供たちが学校のデバイスを持ってこういうイベントに参加してくるということはもうでき始めています。データの共有とか教育の学校の中での活用はもちろんなのですけれども、もう少し広い視点で、学校や地域を超えた、場合によっては観光客なども受け入れながら子供たちが学んだり発表したりする場を創ろうとしたときに、そもそもの共通の安心・安全な環境とIDを持っているということはとても重要で、その可能性を感じているということが現状です。

「e-net.nara.jp」という地域ドメインを、学校教育の中でも国立や公立という範囲を超えて私立とも共同管理・利用したり、管轄が違う幼稚園・保育園の先生方や子供たちも共通で使えたらいいよねというイメージで、地域ドメインを取得して、これを「いいねっとなら」と名づけて1人1アカウントでスタートしたわけです。マイナンバーに近いイメージでやっているのですが、学校代表のメールアドレスを利用するという事は、サービスに応じて仕方なくやっていることもあります。基本的には設定しないで、とにかく個人が1つのIDを持つということ、所属に関係なく実践してきました。国公立の教員は、教科の研究会で交流もしていますし、当然子供たちは校種を超えて進学しますし、一時的に私立へ行ったり、県外へ行ったりということに対しても県域全体で対応したい、ということ

をイメージしてスタートしたのですけれども、思った以上に効果があると感じています。この環境が前提だったら、こういうことが考えられるということは、皆さんすごく発想が豊かなので、結局できないから諦めているとか、もともとできる前提がないからしなかったことというのが、県域公用アカウントがあることで、すごく動き出しています。

例えば、免許状の更新講習がなくなって、教員研修の記録を取らなければならないと決まったときでも、奈良県の場合は、1人1つのアカウントを非常勤であろうが常勤であろうがみんな持っていますから、それで研修履歴を取ることができる。担当者の頭の中で、情報をどうやって提供して、申込みをして、それをどのように集約していくということは非常にイメージしやすいし、実際それで物事を動かすことができている、多くの人が労力を注ぐ単純作業の部分はシンプルに対応を考えて、本来力を注ぐべき、このような地図をつけたほうが分かりやすいのではないかという工夫とか、自分で研修記録がどんどんたまっていくことが可視化されたら楽しんで教員研修ができるよねとかを考えることができる。そのほかにも例えば、子供たちの実態を知りたい、心と生活のアンケートを取りたい、いじめのアンケートを取りたい、コロナで中学校の授業が遅れたから高校入試の範囲を限定したい、という発想が浮かんだときに、これも全県で対象者を絞って、ぼんとアンケートが実施できるということなども、実際やってみていますので、当然こういうデータ分析というのも、先ほど文部科学省やデジタル庁さんが示していただいたようなダッシュボードで管理している側や当事者が見るということも可能なのですけれども、先生によって知りたいことが違うので違うフォーマットで見ようとか、クラスの単位で見ようとか、個人の単位で見ようとか、アレンジすることも可能になっている。

また、利用の状況を、学力と結びつけてみたいと思った時には、AIドリルとも連携して活用できるような仕組みも、共通のIDをもつことで運用し、結びつけたりヒントにすることができている。「気づき見守りアプリ」といって、先生たちが子供の様子を見て、何か今日ちょっと顔色がおかしいよねとか、元気がないよねというのをぽっぽっとタッチで記録をためていくことによって、何かがあったときにさかのぼって見ることができる。ダッシュボードどうこうといったら、何か今ある現実から問題を早く発見しようということに取り組んでいることが多いようです。それも確かに有効だと思うのですが、実際、先生たちにとっては何が起こってしまったからその原因が何だったのかとか、どういう経緯でこうなったのかというのをさかのぼれるということが、経験を生かすという点で非常に重要だと感じていますので、そういう意味では日頃からデータを蓄積しておくことで、今は探索的にデータを扱っていく必要があると思います。

実際、国の学力・学習状況調査をCBTで実施しようという流れが見えてきたときも、まずは独自でこのようなことからできるのではないのか、子供たち、先生たちも慣れておいてみようということがチャレンジできます。これは公立だから、私立だからと言っていたら参加できる・できないで混乱してしまいます。県域で1つ、共通のアカウントを持っていましようということ、そのときはこれを使えばいいという準備ができます。普段は独

自分で使っていてもいいのだけれども、こういうときにはこれを使いましょうねということで物事が停滞せずに進みますし、不登校支援としてフレキシスクールを立ち上げたときも自治体を超えて学級をつくることができたのは、共通のアカウントを持っていたからです。ここ数年試行錯誤してきたことが、いよいよ本番がスタートしたという感じになってきました。普段から小学校、中学校と使ってきている学習履歴も紐づいているIDでそのまま高校入試に申し込んで、出願から可否の手続、県外から来た子にはそのときに付与するという形で、手続を全部オンラインでやって、可否の確認までをオンラインでIDでやってみて、これもできているということです。

細かいところの調整などは、まだまだこれからもずっと必要なのでしょうけれども、バックヤードには、県内のすべての市町村が参加している連携協議会という、先生たちや教育委員会が連携した組織と、それを実務としてこなしていく県域運営支援センターというのががつつり組んで取り組んでいる。これらを全部、ちゃんと動かそうと思ったら、まずはアカウント・IDのこと、それから認証基盤としてそのデータは確かなのか、ということの担保が必要です。今はこれを人力でやりながら、多分大丈夫だろうということで目視で確認しながら進んでいる部分も大変多いのですけれども、国のほうで方針を決めていただいたり、きっちりこれでいきましょうという認証基盤がもしできたとしたら、私たちはいろいろなことにチャレンジできるようになるだろうと思っています。

最初の話に戻りますけれども、そういうことをしていくためには、国レベルでは何をやるのかということの明確化と、前例や実績がないとブレーキがかかることが多い教育行政の体質なので、ぜひ概算要求などで、チャレンジする自治体を後押ししてみようというモデルをつくることとか、今、説明したような先生や子供たちをつなぎながらいろいろなことにデータを安心・安全に使えるような基盤というのを、国として作れたらということを感じています。

以上です。

○森田座長 ありがとうございます。

続きまして、同じくゲストスピーカーといたしまして、佐藤潤株式会社COMPASS代表取締役CEOから御発表をお願いいたします。どうぞよろしく願いいたします。

○佐藤代表取締役CEO COMPASS代表の佐藤と申します。よろしく申し上げます。

では、私からは実際に学校現場にデジタル教材を提供している事業者の立場から、具体的に学校現場で何が起きているかというところをお話しさせていただきます。

まずは概況になりますけれども、現状、学校現場においては校務支援システムを中心とする校務系サービスと、あとはデジタル教科書や我々が提供しているデジタルドリル等を含めた学習系サービスといったところが普及をしています。これらは紙からデジタルに移行したばかりというところもありますので、まずデータ活用といった側面で言いますと、第1段階としてはそれぞれのサービス内でデータを活用しているということが進んでいるという状況になっています。

それらが利用されまして、今、データが徐々にたまってきているところでもありますので、第2段階としましては、その下に書いてあります横断系サービスといった複数のサービスからデータを集めて、そこを組み合わせるデータを活用していくようなサービスが出てきておりまして、まだまだ試行錯誤をしている状況かと思えますけれども、本日はその辺りも含めて御説明させていただきたいと思えます。

次をお願いします。まずは学習系サービスの中では一番導入・利用が進んでいるところと言えますと、デジタルドリルというところになっておりまして、現在、ほとんどの学校で導入をされておりまして、あとはドリルという特性もありまして、日々日々利用されていてデータが蓄積されている状況となっています。我々が提供するデジタルドリルのQubenaというものに関しましても、全国100万人以上の子供たちが今使っていて、年間と言いますと、問題の回答数としては20億件以上の規模になっておりまして、ある種大規模なデータがたまっているという状況になっています。

次をお願いします。第1段階として、まずはサービス内でのデータ活用というところで、これは児童生徒向けの画面になっておりますけれども、デジタルとしてはできるのは当たり前なのですけれども、これまでは紙でやってきたところから学習時間や回答数、正答率というところを確認できるようなものになっていて、それが左側になっています。

加えて、こういった形で間違えたかとか、正解したかというデータもありますので、これは弊社独自ではあるのですけれども、右側にありますような單元ごとの習熟度判定がAからDみたいなところで学習成果も把握できるようになっていますので、これらのように学習を見える化することによって子供たちが振り返りをしながら主体的な学びにつなげていくということが日々行われています。

次をお願いします。もちろん教員向けにもダッシュボードを提供させていただいておりまして、これも紙との比較にはなるのですけれども、リアルタイムでクラスでこの辺が分かっているのではないかとデータが取れたり、あとは個別の子供たちごとにこの辺でつまづいている子供たちがいるというのをほぼリアルタイムに把握することができますので、それらに応じて指導をよりスムーズに行うことができるように現状ではなっております。

次をお願いします。今までのところは教育データのサービス内での活用といったお話でしたけれども、これらの学習データはデータとして残っていますので、効果検証等にも利用させていただいています。こちらは22年度から慶應のSFC中室研究室と共同で、Qubenaの学習データと学力調査等のデータを掛け合わせて、実際のデジタルドリルの利用が学力にどれぐらいの影響があるのかといった検証に関してもこういった学習データを使って行うことができるようになっているのが特徴かなと考えております。

次をお願いします。それらを第1段階としまして、第2段階としましては、自治体さんへのデータ提供を行って、冒頭お話ししました横断系サービスでのデータ活用というところが行われています。現在、奈良市さんや鹿児島市さん等、10自治体10万人規模でデー

タ提供を行わせていただいております、ローデータレベルの学習データ、一問一問子供たちが解いたデータと、あとは何の問題なのかというのが分からないとデータ提供しても意味がありませんので、メタデータみたいなものもつけて、鮮度が高い形で日次でデータを連携させていただいているといったところになっています。

次をお願いします。現状ではローデータを提供しているとお話しさせていただきましたが、冒頭、文科省さんのお話等でもありましたが、業界全体でこういった学習データというのをこの形でデータ連携しようというフォーマットが決定していない状況かなと理解しております、我々としてはサービス内に存在しているデータを特に限定することなく、まずフィージビリティが必要かなと思っておりますので、各自治体様に提供させてもらって、その中で今後、データフォーマット等を固めていければいいのかなと考えております。

次をお願いします。具体的な取組として2つ御紹介させていただきますが、1つ目が奈良市様とやらせていただいているものになっています。こちらが児童生徒向けに奈良市さんが作られているダッシュボードになっていますけれども、もちろん最初のほうで御説明した我々が作ったダッシュボードというものもありますけれども、実際には学校さんや自治体さんには個別の教育目標等があり、それに合わせてどのようなフィードバックをするのがその教育目標に合わせたフィードバックになるのかといったところが、我々が作るものとはまた違う形での見せ方というのが必要になってくるというところがありますので、そういったところも踏まえて奈良市さんのほうでは独自にダッシュボードを作られています。

加えて、ドリルのデータだけではなくて、今後、デジタル教科書やMEXCBT等のデータも含めてデータが連携されてくることを考えると、自治体さん独自でこういったものを持つておくというのが重要となってくるかと思えます。

次をお願いします。教員向けにも同様にクラス単位や個人単位で学習状況を把握できるようにやられているところでもありますけれども、右側は弊社サービスではありませんが、子供たちに日々アンケートを取って健康の状況をはかるようなサービスも提供されております、ある種こういった校務系データと我々のような学習データというものを組み合わせることで新しい視座をこの中から見つけるといったこともできる可能性がありますので、その辺りも奈良市さんのほうでは模索されているといったことになっています。

次をお願いします。2つ目が、鹿児島市さんと取組をさせていただいているものになります、こちらに関してはちょっと特殊な学習の手法を取っております、子供たち自身が計画を立てて実際に学習を行って振り返りを行うといった自己調整型の学びを実践されています。自己調整型の学びとは具体的にどういうことかという中の一つが、子供たち自身が複数のドリル教材の中から自分に合った教材を自由に選択できるというまさに個別最適な学びというところにふさわしい取組かなと思っていて、子供たちにとってはすごく良いのですけれども、一方で、先生側にとってはクラスに30人ぐらいいる中で違う教材をば

らばらに使っている子供たちをどう見取っていくのかというところが課題になっていて、その辺りに学習データの利活用ができるのではないかとこの取組の一つになっています。

次をお願いします。具体的には、実際のダッシュボード等はこれから作成されると思うのですが、複数ある我々のQubenaとeboardさんという他社のドリル教材になるのですが、2つのデータをまず並べるところからということで、これ自体が結構教育業界ではまだ行われていなかったことでして、こうすることで一元的に子供たちを先生が見取っていくということができるようになっていて、左側が学習の進捗、右側が達成度という形で、ドリルの取組問題量と正答率をプロットして子供たちの様子をうかがうという形になっています。

今後に関してで言うと、1つ前に戻っていただくと、これらの達成度みたいなところに関してはデジタルドリルを通した学習の評価みたいなところにも活用することが期待されるのではないかなと考えています。

次をお願いします。最後になりますけれども、学習データの利活用に向けてということで、事業者から見た際の課題みたいなところを少しお話しさせていただきます。

1点目が、データ粒度の統一とメタデータの整備ということで、途中でローデータレベルのデータを現状は提供している。これはフィージビリティだからというお話をさせていただきましたが、これはあくまでフィージビリティであるためなのですが、現状で言うと我々のほうにも大量の学習データがあり、自治体さんのほうにも大量の学習データがあり、複数の場所に同じデータが存在するというかなり非効率な状態でして、この状態を早く解消していきたいと考えています。実際には我々の学習データと何かを掛け合わせるということもありますので、それらのやりぶりも含めまして、どの程度の粒度にデータを集約するのがいいのか、有効に使えるのかというのを実際に検証して、教材間でも統一していく必要があるかなと考えています。

統一に向けては、このデータが何を示しているのかということももちろん重要になりますので、メタデータとして現状、学習指導要領コードを文科省さんのほうで用意いただいています。ちょっと粒度が粗いということもありますので、教科書の単元情報等がオープンに使えるような状態になれば、より精度の高いデータにしていけるのかなと考えています。

2つ目が、学習データにおける習熟度等の評価情報における基準統一ということで、1個目のデータの粒度やメタデータの整備といったところは、そもそもこの学習データ自体が確からしいということが前提で必要と思っています。確からしいというのはどういうことかということなのですけれども、途中でお見せしたような習熟度のスコアを我々も出してはいるのですけれども、それが実際に適切な評価になっているのかというのは各事業者に任されているところもあります。

加えて、途中で見せた複数のドリルのデータを一緒に並べているということもあります。

したけれども、あれも実際問題の難易度が違うものとなっています。難易度の違う問題の正答率を一緒に並べていいのかといった問題が実際にはある中で、まずはトライとしてやっていたところかと思っています。

そういった意味で言いますと、これらの評価情報というのを今後、どう基準を統一していくかというところが2つ目に必要かなと思っています。

最後になりますけれども、それらのデータの粒度がちゃんとできて、習熟度もちゃんと評価できているというところを言うと、データにかなり価値が出てくるのかなと思っていますので、デジタル庁さんの話でもありましたけれども、現状、転校すると有効になったデータが持ち運びできないとか、あとは中学を卒業してしまうと我々も学習データを消さなくてはならないというところがあって、子供たちにとってはかなり不利益かなと考えています。

加えて、公教育でつくられたデータを塾等の民間教育で活用することも現状ではできませんので、そういった活用も本来では視野に入れるべきなのではないかと考えています。

これらを踏まえますと、途中にも出てきましたが、自治体間でのデータのやり取りをどうしていくのかであったり、子供たち個人にデータを提供していくというところを、多分PDSみたいなものを通してなのかなと思っていますが、そういったところの検討も並行して必要になると考えております。

私からは以上になります。

○森田座長 ありがとうございます。

続いて、最後になりますが、ゲストスピーカーといたしまして藤村裕一鳴門教育大学大学院特命教授・鳴門教育大学教員養成DX推進機構機構長から発表をお願いしたいと思います。藤村先生、よろしくお願ひいたします。

○藤村教授 鳴門教育大学大学院の藤村でございます。どうぞよろしくお願ひいたします。

私からは、今までは様々な実践事例等のお話でしたが、そこは軽く扱いながら、教育分野の特殊性、領域固有性のあるところについてのご説明と、それを受けての課題を中心に話ししたいと思います。

まず、教育データの利活用ですけれども、ここ数年積極的に行ってきまして、子供のため、教師のため、そして保護者のため、そして学校設置者のためを中心にしながら研究目的での利用も可能にするということも考えています。

次のスライドをお願いします。そのために、教育データを幾つかの区分で考えていまして、公教育データは学校で1次利用として子供たちのために活用する、それから、先ほどお話のあった個人活用データみたいなものを考えると同時に、先ほどCOMPASSさんも話をされていましたけれども、大学やその他の機関で個人属性を外した上で2次利用することも視野に入れております。

次をお願いします。従来は学校の中だけでクローズドなデータ活用を考えていたのが、教育委員会や文部科学省、それから教育委員会だけではなく首長部局ですね、子供支

援のためには福祉系のデータも必要だし、医療系のデータの連携も含めて非常に幅広く教育データの利活用を考えていますし、保護者や地域との連携も考えています。

ただ、現状で言うと、私は教育データの利活用に関する有識者会議の座長代理もしていますけれども、当面のスコープが初等中等教育限定であり、高等教育と幼児教育がスコープ外となっているのはもったいないなど。この辺をちゃんとつなげるようにするのが今後の課題かなとも考えています。

次のスライドをお願いします。これはよく教育DXについて説明して、私どもとしては第3段階の学習モデルの構造等が教師主導の一斉事業から個別最適な学びと協働的な学びの一体的な充実のようなこともできるようにと考えていますが、例のところを御覧いただくと、従来は教育データも「部分的・静的」な把握しかできなかったのが、今は「全体的」に把握したり、「動的」にリアルタイムで把握したりもできるようになってきました。そして、従来先生の仕事というのは経験と勘による「属人知」だったのですけれども、それを「集合知」を形成してやる。これらを「アシストログ」と呼んでいますけれども、そういうものを生かして「集合知」を形成するというのも考えて、今、実証研究を始めています。そして、従来は「標準モデルアプローチ」、みんなが同じことをするだったので、先ほど鹿児島県の例もありましたけれども、個別最適な学び、「個別最適アプローチ」にするためにデータを活用する。

また、いじめ、不登校などの問題が今は重視されていますが、従来は「後手後手」対応だったものを、教育データを活用していじめや児童虐待を「早期発見、早期解決」。それから、不登校などについては予兆を捉えて不登校になる前に解決を図る、「未然防止」ということも可能にしようと考えています。

次のスライドをお願いします。現在、デジタル庁が2022年に作ったこの教育データ利活用ロードマップというのを作っていますが、これは先ほどデジタル庁さんから説明があったように、来年度中にこれを改正するというので検討を進めています。ここで「公教育データプラットフォーム」みたいなもの考えるということが挙げられています。

次のスライドをお願いいたします。やはり大事なことは、教育課題に対して役立つ仕組みであるべきだということです。そのためにはそれぞれ「データ標準」と「データ連携標準」をちゃんと決めてデータを流通できるようにするという、ただし、セキュリティが大事ですよということ、その辺を考えていくということだと思います。

次のスライドをお願いします。従来の教育現場の特徴としては、「校務系」システムと「学習系」システムという大きな固まりがあり、校務というのは先生方、教職員の仕事に使うもので、成績入力や保健管理などですね。学習系は児童生徒が学習に使うもので、回線も分離していました。2回線で「2層分離」しているところもあれば、自治体と同じように校務系と校務外部接続系、いわゆるメールやホームページみたいなものと、そして学習系という「3層分離」のところもありました。

仕事は職員室でできる、教室でできない。それが、次に進めてください。「GIGAスクー

ル」といいまして、1人1台端末を子供たちに配付しましたけれども、そこでは教室のデータ、そして職員室のデータが分断しては何の意味もありません。両方で生成された情報をうまく使うということで、次に進めてください。次に、どんどん学習系のデータもたまり、それから発生源入力、今は回線が分離しているから教室で出席入力を校務支援システムに入れられないのです。非常にナンセンスで、職員室に持って行って手入力するということをしているのです。それも変えるために、次をお願いします。どういう情報種を全体最適して、どういう設計をして、セキュリティーをどうするか、情報連携を前提とした新たな標準化をデータ標準とデータ連携標準をどうすべきかというのを、今、試行しているところです。

次をお願いします。そのために、文部科学省の方でこのような教育データの標準化を行っていただいています。「主体情報」、「内容情報」、「活動情報」です。「活動情報」では、先ほどCOMPASSさんから説明があったように、「スタディエログ」の解析については世界的にも非常に研究が進んでいます。また、「ライフログ」と「アシストログ」に関しては日本発祥の概念ですので、まだまだ検討が必要な段階ですが、「ライフログ」に関しては、日本は諸外国・欧米と違って学習だけするのが学校ではなく、「知・徳・体」いずれも学ぶという特徴があるために、生活健康指導もあれば、体育の授業といったものもあるので、「ライフログ」の研究は、現在、IEdeTech Globalという国際標準化組織から日本が「ライフログ」の標準化をやってほしいと言われて、今、検討を始めたところです。

「アシストログ」は、先ほどCOMPASSさんの説明もありましたけれども、これからが研究のしどころで、「集合知」をどうつくるかということだと思っています。

上のほうに赤で書いています「児童生徒ID」については結構課題がありまして、どうしたらこれをうまく連携させることができるのかというのは課題だと考えています。

次をお願いいたします。そのために、従来は「校務系」と「学習系」を全く分断して考えていたのが、「学習系」は子供たちと向き合う「フロントエンド」で、「校務系」は子供たちの学びを支える、人格形成を支える「バックエンド」だと考えて一体のものとして動作し、データ連携も進めるべきだという概念を、今、進めていることになります。そのときに、全く独自のものをつくるのは大変ですし、既に普及している国際標準仕様があるのであれば、その日本適合版をつくる、つまりJapan Profileという形で、例えば「振り仮名」という概念は日本しかありませんので、そういったものも取り入れてやっていく。

子供の個人に関する情報はOne Rosterという国際標準仕様がありますし、いろいろなアプリケーションが連携動作するためのLTI (Learning Tools Interoperability)、そして学習系の内容の連携に関してはxAPI等を使ってやりましょうということになっています。

次のスライドをお願いいたします。それを円滑にするために、従来、「校務系」と「学習系」を回線も分離し、データも完全に分離していたのを、「2層分離」・「3層分離」から今度はインターネット回線一本でセキュアにやる、いわゆる「ゼロトラストセキュリティー」の基本概念に従いながら、強固なアクセス制御によって回線が一本であっても

安全に活用できる、そうすることによってデータの流通も円滑にしつつ、セキュリティーも担保するという事を現在考えています。どこからでも先生方は仕事できるし、最近、教師の働き方改革が話題になっていますけれども、ロケーションフリーで、希望する人は例えば自宅で親の介護をしてから仕事するとか、それもちゃんと勤務時間にカウントできたりとか、小学校などは女性の教員が多いのですけれども、そういったときにはお子さんを保育所に早く迎えに行かなくてはいけないのに、成績処理のときに学校でいつまでも残業しなくてはならないということを解消したりするという事もできるようにしようということをやっています。

ただ、このとき、「ゼロトラストセキュリティー」は今の「2層分離」・「3層分離」でやったときとコストがあまり変わりません。そのために、何かインセンティブとして補助金が欲しいですし、それから、南海トラフ地震や首都直下型地震などの大規模災害が予想されても、子供の学びも絶対止めない、教師の仕事、子供を守ることも止めないという意味で、「強固なアクセス制限によるセキュリティー」を担保した業務が円滑にできる仕組みをやるために補助金等が欲しいなと考えている次第です。

次のスライドをお願いします。そして、大きな問題点、これはぜひ皆様のお力をお借りしたいところなのですが、これは世界で初めて「教育クラウド」についてロンドンの国際学会で提案したときの資料の日本語訳なのですが、「IaaSベース」と「SaaSベース」での整備は進んできましたが、実は「PaaSベース (Platform as a Service)」の部分について、先ほども話があるように「共通認証基盤」が実は確立されなければならないはずだと最初から考えたのですが、これがまだできていない。

それから、「真正性」の証明のための内容固定みたいなものも必要なのですが、これもまだ確定していないので、データの「真正性」の保証がないままデータの利活用をする。大学の先生方だとよく分かると思いますが、海外のある国からの留学生が入学したいと言ってきて、「成績証明と卒業証明書がほぼ偽物だ」みたいなものもあったりするぐらいですので、これは早期にぜひ解決して、転校処理や進級処理なども安心してできるような仕組みをぜひ国としてつくっていただきたいと考えています。

次のスライドをお願いします。そのために、デジタルデータの「真正性」を確保するための「認証局」というものができるだけ早期に欲しいと考えています。今後、進学、転校等の手続をデジタル化する、今は紙ベースで保護者に発給してそれを向こうの学校に持って行ってもらうという非常に原始的なことをやっていますし、そういったことは医療や防災などの他分野も含めてデータ連携が行われるような社会になっていかなくてはならないと考えていますので、その際、誰がつくったデータか、誰に関するデータなのかといった「真正性」の確保は必須だと思います。卒業証明書や成績証明書などの学習成果証明がデジタル化していく中で、今のままでは「真正性」を確認できないということを非常に憂慮しています。

次をお願いします。そこで、これは自治体やベンダーがそれぞれに対応できる範囲を超

えている話と認識していますので、認証の仕組みを検討・整備していく必要がありますが、ニーズが顕在化してから検討するのでは遅いと考えています。既に国で整備されているデジタル基盤との整合性もあることから、ぜひお願いしたいのは、デジタル庁の管理下でこれらの検討をしてほしい。例えば有力な候補として、先ほどのデジタル庁さんからの「GビズID」みたいなものもいいのかなと候補として考えていますけれども、ただ、教育分野でビジネスのものを使わせるのという心理的抵抗も想定されますので、「GビズID」という名前は大好きなのですけれども、そこは「パブリック」なもので例えば「GパブID」などと名前を変えて、余計な懸念をなくしながら全ての分野で横断的に利用できるようなことを配慮していただければと思います。

次のスライドをお願いします。あと、児童生徒情報連携の課題なのですけれども、現在、私どもの検討を進めまして、「住基台帳システム」から「学齢簿システム」に情報を流し込んで、そこから「就学援助システム」へと流し込まれます。「学齢簿システム」と「就学援助システム」は「ガバメントクラウド20業務」のうちの2つに当たります。「就学援助システム」の座長、「学齢簿システム」の委員をしていますけれども、ここで課題が明らかになっています。現在、学校で手入力するという原始的なことをやっているわけですが、学齢簿システムから校務支援システムに児童生徒情報を流し込めば、それで名簿情報を引き継ぎますし、今、既に校務支援システムから「学習eポータル」というハブになるものから各種「学習系システム」に児童生徒情報を流し込めるようになっています。

ただ、そのときの課題としては、ガバメントクラウド側で「MJ+」が文字コードとして採用されていますが、ここに昔でいうところの外字相当のもの、多様な漢字があるのですが、実は学校現場の端末は、教育委員会も含め一般的な端末には「JISX0213」という一般的な文字コードしかありません。そこで流し込むと文字が化けてしまうわけです。だから、そういったことがないように、「学齢簿システム」から「校務支援システム」に行くときに「文字コードの縮退」が必要ですので、これは各事業者がやるとか、各自治体がやるのはナンセンスですので、ぜひ国のほうでこの「文字コードの縮退」をどうするのかというのをご検討いただきたいと思います。

次をお願いします。次のところは軽く飛ばします。これは渋谷区に先駆けて日本で初めて教育データを活用し、「スタディーログ」だけではなく「ライフログ」（生活健康情報）や、教師の支援がこういうタイプの子供に有効だ、こういうタイプに有効ではなかったみたいな情報（「アシストログ」）を処理して、次をお願いします。ここにあるような様々なシステムから3種類のデータを読み込んで、次をお願いします。先ほどの渋谷区に先駆けて、その元になった大阪市のものなのですけれども、赤い部分が「学習系」のデータで、青い部分が「校務系」のデータからもってきて、真ん中に「心の天気」というものがありますけれども、心理情報ですね。雨なら悲しい、雷は怒っている、晴れなら気持ちが良い。それと、真ん中の下の部分は保健室の利用記録なのですけれども、これが非常に相関が高くて、

これでいじめの早期発見や不登校の予兆の発見ができるようになる。こんなこともしていました。

次をお願いします。これは具体的にシステム連携がこうなっていますよといういろいろなダッシュボードを持っていて、次をお願いします。問題は、「いろいろなデータをどこにためるのか」ということで、ちょっと戻っていただいていいですか。1つ前に戻ってください。「データレイク」にため込んでやると、わざわざダッシュボードを何千万もかけて作り込まなくても、生成AIを使って分析が可能なのですね。今、その試行研究も始めていますけれども、この「データレイク」は誰が費用負担するのとか、事業者なの、教育委員会なのとか、どういう設計でやるのみたいなものが問題になっていますので、この辺の検討も必要かと思っています。

次をお願いします。ここは先ほどの「COCOLOプラン」等の話も出ていますが、実際にやっている事例としては、次をお願いします。実際にタブレットのインカメラを使って感情データを集めて、次をお願いします。こういうデータで帰りの会と朝の会等でデータを取って、「いじめの早期発見」、「不登校の予兆発見し未然防止」することなどもやっているということです。

文書資料の方をお願いします。ちょっと駆け足で恐縮です。文書資料の方でこれからご検討いただきたいことの概要についてお話しさせていただきます。

一つは、「児童生徒の識別ID」を含む名簿情報の連携のAPIによる自動化が必要だと。今、CSVなどで連携しているのでリアルタイム連携できない、転校してきてもしばらくかかるみたいなことがありますので、これの「名簿情報のAPI連携」が必要なわけですが、これは現在、今年度のデジタル庁事業の実証で取り上げられる予定で既に準備が進んでいますので、ぜひその年次更新を先ほど話があったように過度の負担がないように。名簿やアカウントの管理は、実は教育委員会がアカウント管理していないところが2割あるということが分かっています。とんでもない話ですね。その辺もちゃんと明らかにしたい。

次は、「校務支援システムに記録される情報のさらなる標準化」と、「標準仕様準拠商品」といったものの連携等をぜひやってほしいということです。

次のページに行っていただいて、ここでお願いしたいのは、自治体間連携が「校務支援システム」でできない。要するに、転校処理が自治体がまたがるとできないのですね。自治体内は転校処理できるのですけれども、それをぜひできるようなさらなる標準化が必要です。

それから、文字セットの話は今、お話ししましたので、ガバメントクラウドの「学齢簿システム」からエクスポートする際に縮退を御検討ください。

それから、「指導要録」という子供たちの学習や在籍の記録の「真正性」の保証と原本の完全電子化の推進ということ。今は紙ベースのものを電子化し、さっきの「認証局」の検討をお願いしたい。

「校務支援システム」間のデータ連携による転校時のデータの引継ぎは先ほどお話しし

たので飛ばします。

次をお願いします。また、データに対するアクセス権限の管理と技術仕様の策定、いわゆるRoll-based Access Controlロールベースドアクセスコントロールについてまだまだあやふやなところがございますので、そこについて検討を進めていただきたいということ。

また、[学習ツール]「LTI (Lerning Tools Inter-operatability)」対応によるアカウント管理が不要になる運用への移行。今はそれぞれのシステムで名簿情報を入れているのですね。非常にナンセンスです。そういったものを「LTI連携」してやれば、全てが連携動作する。ハブのような形で、もしくは全部が網目状に連携するという形を検討いただきたいと思います。

また、先ほどのCOMPASSさんからのように、単元を示す「教科書見出し」というのは現在、著作権使用料が必要です。これのためにいろいろなデータ連携がうまくいかないというのがあるので、これの「オープンデータ化」をぜひ御検討いただいて、法律でその辺を検討するなり、何らかの仕組みをつくるなりしてほしいと思います。

次をお願いします。そして、教科書・教材やウェブサイトのアクセシビリティ。欧米等では視覚障害のある子供とか、ディスレクシアみたいな特定の文字が読めないみたいなものをこのようにするといよいよとか、色弱のためにというのはちゃんと基準があるのに日本はないので、それも検討していただきたい。

それから、「パーソナルデータストア (PDS)」による個人へのデータの還元ということも先ほどCOMPASSさんも提案していただきましたが、全く同感です。この仕組みもご検討いただきたい。

そして、先ほど説明したので省略しますが、「データレイク」、「ゼロトラストセキュリティ」のため、その辺をどうしたらいいのかという点などもぜひ御検討いただきたいと思います。

私からは以上でございます。

○森田座長 ありがとうございます。

それでは、5名の方からプレゼンいただきましたけれども、これから討議に移りたいと思います。いつものように御発言のあるときには、会場にいらっしゃる方はネームプレートを立てていただきたいと思いますし、オンラインの方は挙手ボタンで手を挙げていただければと思います。

なお、もう一つ議題がございます、これが15時30分からを予定しておりますので、それまでの間、ディスカッションいただきたいと思いますけれども、御発言はなるべく簡潔にお願いしたいと思います。

それでは、どなたからでもどうぞ。

では、丹野構成員。

○丹野構成員 様々なプレゼンターの方からお話を伺って、大変啓蒙されたといえますか、ここまで進んでいるのだなということを実感として聞かせていただきました。

教育分野の全国共通データベースを設けて、児童生徒だけではないというお話はありましたけれども、児童生徒自身の転校等の様々なお話がありましたけれども、そういう事務的な事柄について教員の皆様の負担軽減をするというのは、先生方は非常に多忙でいらっしゃるし、人手不足だというお話も聞こえてきておりますので、デジタル化に非常にふさわしいものだと考えます。それが一つです。

それから、もう一つだけ申し上げれば、再三言われていますように児童生徒一人一人にIDというお話なので、それでデータを取るのしょうから、そもそもそういうことの漏えいも不正取得も一切許されないと固く思っておりますので、セキュリティーに万全を期すのも当然のことだとまず思っています。

それに加えて、少し懸念がございます。今、御説明がございましたが、児童生徒の健全育成のところでは言わば不登校やいじめなどの非常にセンシティブな心の問題について、早期把握及び適切な対処を実際にできているというお話もございましたが、それを目的にしていると思われませんが、実態としてはもう少し複雑なのではないのかなと様々な事例を見聞きする中では思っております、そういう意味では現実と乖離をしていないかという部分で多少の懸念を感じました。非常に重要な問題だと思いますので、関係機関、ステークホルダーをはじめ、より精緻な議論を行ってより良いものにしていただきたいと思いますと思っております。

以上でございます。

○森田座長 ありがとうございます。

それでは、ほかにいかがでしょうか。

続きまして、依田構成員、お願いいたします。

○依田構成員 ありがとうございます。依田高典でございます。

私個人は今日、プレゼンがあった全ての皆様の意見に対して賛同の立場を取るものでありますので、基本的に賛同しつつ、かつ、そのときに議論しておいたほうがよいのではないかということについて2点言及し、最後にそれに対してあらかじめ私の考えを示したいと思います。

資料としては今日の資料2番に大きく関わるものでございますが、教育データを組織間連携していくときに、特定の要件を満たす場合において、現行で法令が認める以上に個人同意がなくても組織起点の連携を今後拡大していくということを、多くの今日いる皆さんは場合によっては法令改正も含めて考えられると考えております。私自身も公益性に十分に配慮しながら、そうした可能性を排除するものではない賛成の立場でございますが、やはりその際は国民目線から見て2つの論点を議論すべきではないかと思っております。

第1においては、個人情報データの組織間連携を望まない生徒さんや親御さん、親権者がおられるであろうから、そうした方々をどうしても配慮する必要があるので、オプトアウトの権利は認めるべきかどうかというのが第1の論点でございます。

第2の論点は、とはいっても絶対に事故が起こらないかということ、将来、何らかの善意

の過失に基づいた情報漏えい等はゼロではないと思います。その場合において、今日、議論しているように組織起点で本人同意がない場合に、ひょっとしたら学校、あるいは先生の管理責任が重くなる危険性はないか、そこについても十分に議論してほしいと考えております。

結論から言いますと、私は今の個人情報保護法でオプトイン・オプトアウトの区別が十分についていないところではございますが、もしも仮にオプトアウトの権利が認められるのであれば、社会的な抵抗は多少なりとも減るだろうし、データ連携の正当性を強化できるのではないかと、かつ、オプトアウトの仕組みがあることによって組織の管理責任も過重に学校並びに先生に対して重くなるリスクも避けられるのではないかとこのことを私の意見として述べさせていただきます。

また皆様方において御検討いただけましたら幸いです。

私からは以上です。

○森田座長 ありがとうございます。

それでは、安中構成員、どうぞ。

○安中構成員 製薬協の安中でございます。PCの調子が悪いので、カメラはオフのまま失礼いたします。

本日、御説明いただきましてありがとうございます。前回の検討会で医療分野の議論をした際に、私の製薬協のプレゼンで、出生前から老年期までのライフコースデータの基盤の重要性について述べさせていただきますけれども、その中で学童児健診のデータの必要性についても資料中に盛り込んでおりました。本日の御説明の中にも学校健診の情報や医療データの連携の必要性について言及がございましたので、本日の本論とは若干離れるかもしれませんが、2点ほどコメントさせていただきます。

1点目は学校健診のデータと全国医療情報プラットフォームとの連携についてで、これが実現しますと、ライフコースデータがより充実しまして、より精度の高い長期的な医薬品の効果、あるいは副作用の把握が可能になります。また、同様の指摘が医学会や学校健診データの利活用に対する政府の検討でもあったと伺っております。健康診断情報のデータが医療データと連携できれば、子供さんが受ける医療の質の向上にもつながると期待されますので、この連携についても視野に検討すべきかと思っております。

その際、分野横断的なデータ連携を実現するためには、本日、議論がありましたけれども、特にデータを連携するためのIDをどうするかが大きな課題であると認識しております。2次利用はともかく、特に一時利用の際には確実なデータ連携が必要であるため、IDについてはこの際しっかりと議論をお願いいたします。

2点目でございます。データに関する教育についてです。この検討会で議論しているような様々なセクターのデータ利活用の動きは、グローバルな潮流になっております。そして、偽情報対策も喫緊の課題と理解しております。日本が信頼できるデータ基盤を構築すること、さらにはこれらのデータを利活用することで受けられるメリットについて、将来

を担う子供さんが学校教育の中でしっかり触れることが重要かと思っておりますので、その教育の観点についてもぜひ御検討いただければと思っております。

以上でございます。

○森田座長 ありがとうございます。

それでは、続きまして、岡田構成員、その後で落合構成員、お願いいたします。それから稲谷構成員、お願いします。

○岡田構成員 岡田でございます。私からは、データのアクセス権限についてコメントさせていただきますと思います。

今、様々な皆様からの御発表を伺って、やはり教育データというのは本当に様々なデータが含まれるということを実感しました。スタディログ、ライフログ、アシストログといった区分けもありますし、また、校務系データ、学習系データといった区分けもあります。その内容のセンシティブ度合いも含め、純粋に学習の度合いの達成についての情報から、もう少し踏み込んで心身の健康、家庭環境、いじめの問題など、カテゴリーに応じた留意点も異なってくると思っております。

そうすると、誰にどのような形でどのような情報にアクセス権限を与えていくのが適切なのかなというのが情報や文脈によっても大分違ってくるのかなと思っております。職員と生徒の関係、また、生徒・保護者間でも利害が対立する局面というのも当然あるわけです。また、学校を転校する場面や、学校を超えた横断的な分析など、学校の外に情報を出していく場面というのもあります。そういった様々な場面ごとに情報の機微性なども鑑みて、それぞれどういう形で誰にアクセス権限をいつまで付与するのか、また、そのときに同意を必要とするのか、しないのか、誰のどのような同意を必要とするのか、オプトインなのか、オプトアウトなのかといった考え方についても、一定程度共通の認識のようなものがあれば、もちろん事案ごとに異なるので100%統一的な考え方を打ち出すことは難しいかもしれませんが、法律でどうこうという問題に限りませんが、何か基準となるような大きな共通認識がもう少し醸成されてくると、より安心して使えるようになってくるのかなと思っております。生徒のプライバシーはもとより、職員のプライバシーの問題もありますし、また生徒と保護者の利害が対立する場面もあって、三者三様、さらにプラスで様々なステークホルダーが関わってきます。教育での利活用というのは一歩間違えるとすぐ社会的に炎上しやすい傾向にあり、どうしても萎縮効果が生まれてしまうということがありますので、本来もっと有効に利活用できるデータが利活用できなくなってしまうと非常に勿体無いことなので、そのような安心して利用するための環境整備をさらに考えていくということも重要なのかなと思われました。

以上です。

○森田座長 ありがとうございます。

それでは、落合構成員、どうぞ。

○落合構成員 どうもありがとうございます。

そうでしたら、私も3点ほど申し上げたいと思います。

1つ目が、まずどういう形でロードマップないし議論の範囲というのを整理していくのかがあろうかと思っております。本日、御紹介があった中でも校務系と学習系のというお話もございましたが、そのほかにもいわゆる初等、中等、高等、大学、また、就職などもあります。もしくは先ほど安中構成員からは、医療との関係についても言及がございましたが、前回、議論をしておりました医療に比べますと、個別個別の議論になっている部分がかかなり多いのではないかと思っております。全体としてどういう教育というか、それが最終的にこれからの時代においては就職だけではなくて、例えば場合によってはリスキリングなどまでつながっていく中で利用できるかということだと思えます。例えば、小学校や中学校の中でその時代にしっかり使えるようにすることも必要とは思いますが、一方で、医療においても情報が一生に残るような財産であるという議論はあると思いますが、教育においてもまた違った意味で生涯の財産になるものでありますし、社会のためになる場合もあるということだと思えます。そういった視点をもう少し入れて頂きたいと思えます。医療においても、なかなか、1次利用と2次利用、つまり地域医療情報連携と匿名加工、仮名加工での利用を結びつけるには、5～10年以上の期間が実際にはかかっていたわけではあり容易なことではありません。教育においても、比較的直近での変化の差分を見ると、議論の進展が非常に速い印象は改めて本日聞いていて伺ったものの、全般的な視点は、長期議論がされてきていた医療に比べると、まだ全体の総合的なパーツがつながるといった姿が見えていないと思えます。そういったところは、ぜひ今後、こういったデジタル行財などでの議論もきっかけに、議論が整理されていくといいのではないかと思いました。

第2点目としましては、今、目標について申し上げましたが、データベースなどであったり、IDなどのシステムであったり、基盤側の整備についてです。

まずIDについては、非常に直近で使われているプラットフォームのサービスもかなり中心的に利用されているところもございます。一方で教育という分野については一つの立派な公的ないし準公的な役割を持っている領域だと思えますので、そういった中で情報が継続的に連結して利用できるようにしていくこと自体は、社会的なインフラになるようなものであると思えます。そういった社会インフラとしてのデジタル基盤の安定性や、国としてしっかりどう確保していけるのか、という観点でぜひ議論をすることは大事ではないかと思えます。

また、データベースなどの整備という観点でも、前回の医療とも比べていくべきなのではないかと思いました。医療においては、地域医療情報連携の関係で、もともと地域ごとに進んでいたこともございましたが、電子カルテの情報共有サービスなどの一次利用側についても直近の医療法改正などでかなり基盤整備がされてきて、さらにそのほかのデータベース、公的DBもそうですし、さらに民間・学会のデータベースなどについても議論がされ、さらにこれらの情報の2次利用について、個人情報保護法の3年見直しにおける改正や、次世代医療基盤法、また、直近の医療法における仮名加工情報の利用といったものと

組み合わせて議論がされてきております。

データのガバナンス自体については、第3点で述べようと思っておりましたが、そちらはそれで向上させていけません、一方で、データベース自体をどう整備していけるかという観点においては、先ほど依田先生もおっしゃられたような同意を省略してという場合も含めてどういう形で整備をしていくのかを考える必要があります。しっかりデータを集められる形はまずしっかり整備していけるように、データベースやデータがある場所がどこなのか、どこを標準化をする必要があるのか、また、新たにつくっていくべき連携基盤などはどこにあるのかを特定してさらに議論していくことが重要ではないかと思えます。

最後に第3点としまして、ここまでもコメントがあった構成員の先生方からも御議論がありました、こういった情報の利活用を進めていくことと同時に、ガバナンス面での整備も非常に重要だと思います。岡田先生がID管理、アクセス管理とおっしゃられていた点は、組織としてまず情報を利用できるようにしていくための基盤を整備していくということは議論がありましたが、一方で、個々人との関係で極めてセンシティブなデータも多いということが学校の特徴だと思いますし、そこにある種の利害相反が入りやすいことも特徴なのではないかと思っております。そういった意味では、特に具体的な利用場面では、個人ベースでの同意、アクセス権限の管理ということを組み合わせることも重要であるように思います。また、例えばこういったシステムについても、個別個別でなかなか自治体ベースや学校単位でできないという話もある一方で、恐らく情報の使い方のガバナンスの整備も、現場に何となく任せておくこともなかなか難しいところもあると思います。情報を利用できるようにする基盤を整備しつつ、ただ一方で、非常に問題になりやすい性質がある教育分野ということも考えながら、全体の情報の利活用自体は進められるようにしつつも、ガバナンスのレベルも上げていけるようにする、ここは強く意識しながら議論していくことが重要ではないかと思いました。

私の意見は以上でございます。

○森田座長 ありがとうございます。

続きまして、稲谷構成員、どうぞ。

○稲谷構成員 稲谷でございます。

私も既に何人かの構成員の方がおっしゃられたことの繰り返しになってしまう部分もあるのですけれども、まず今回、御提題いただいています教育データを利活用して個人に最適化された教育というものを実現していくということは、少子高齢化で人材が核になってくるこれからの時代を考えると極めて重要な事業であるということはいくぶん分かりますし、そのために様々なデータを連携できるように、医療との連携という話もありましたし、就職や高等教育の連携という話もありましたし、そもそも地方自治体間での連携という話もありましたが、そういったことができるように基盤を整えていくということは極めて重要であるとは思っています。

ただ一方で、今、落合先生も最後に御指摘になっていましたけれども、私個人といたし

ましては、教育において扱われるデータの多様性というものもかなりきちんと考えて進めていくということが必要になってくるのかなと思います。データが多様であるということは、使われることに関してリスクの大きなデータとリスクの小さなデータというものがそもそもあるということだと思えるのです。そのリスクの大きめに見えるようなデータについて、例えば岡田先生がおっしゃられたように誰にどこまでどういう形で見せるのですかというところを厳格に制限していくというところはとても重要になってくると思いますし、あるいは、その部分については少なくともオプトアウトできるようにするという依田先生からの御提案というものも重要であると思います。

特に個人的に思ったのは、データの性質でいきますと、メリットや技能ベースと申しますか、割と客観的に能力などを評価しやすい部分についてはどんどん連携を進めていく方向で考えていくことによって、その人がより社会で活躍できる場を見つけていくみたいな、あるいはそういった方向につながるようになるべく教育のプログラムを組んでいってあげることについて、もちろんこの点につきましても、そういった評価基準や技能認定の方法については、これも御指摘があったようにどういったものを統一的にエビデンスベースでつくっていくのか、また、そのサイクルというのをどういう形で考えていくのかといったところについて議論を深める必要はあるところですが、それを踏まえましても、この部分というものは今、進めていくというところに非常に重要性も感じますし、リスクというものもある程度きちんとコントロールすることができるのかなと考えられます。

一方で、これまで御報告の中で日本の学校教育というのは単純に技能だけではなくて、徳・体も含めてという話があったのですが、やはり知・徳・体を含めた全てのライフログを活用するというのはやや性質が違うところがあるのかなとも思います。必要性を否定するものでは全くないのですが、そもそも知・徳・体についての総合的な評価基準をつくるということ自体がそもそもかなり論争的な考え方でありまして、これまでの構成員の先生方の御意見にもありましたように、ここを拙速に進めてしまうと、せっかく活用できる可能性があるところまで止まってしまうというリスクを生じてしまうのかなということもありますし、この辺りの少しリスクの大きめだと考えられるような、それは分けても人格の形成に関わる以上は避けて通れない問題だと思うのですが、そういったところにつきましても、そこで生じるリスクの在り方の分析でありますとか、あるいはそれを低減するためのガバナンスの仕組みについても併せて慎重に議論を進めていって、その結果としてどこからどのような形で進めることができるのかといったように論点を絞って積み上げていくというやり方が必要になるのではないのかなと思いました。

すみません、あまりまとまりがないかもしれませんが、私の意見は以上です。ありがとうございました。

○森田座長 ありがとうございます。

それでは、もう一人手を挙げていらっしゃるが、生貝構成員、どうぞ。

○生貝構成員 ありがとうございます。

私からは、今日のお話からはとてもスペシフィックなのですが、藤村先生のお話の最後のところでございました、教科書の見出しのオープンデータ化ですね。Wordの資料のほうで最後のほうにございました、あれはぜひ進めていただきたいなと思います。よく見出し、すなわちメタデータというのは料理としてのコンテンツ、データそのものを流通させるためのメニューとも称されるところがございすけれども、ここがオープンにならないと、やはりこういった様々なコンテンツ、データの活用というのも進んでいかない。

と申しますのも、これは実は教育だけではなくて、オープンデータ政策一般として非常に重要な論点にも関わるかなと思っておりまして、以前、異構成員からも今回の検討会の議論の射程にオープンデータの部分というのもしっかり含んでいくべきではないかという御発言もあったかと記憶しておりますけれども、伝統的な、あるいはこれまでのオープンデータ政策というのは基本的に純粋に公的機関が保有するデータのオープン化というものを対象としてきたところ、特に例えば欧州ですと、2019年のオープンデータ指令の中で、国が、あるいは公的機関がお金を出している度合いとか、公的な性質の強い民間保有データについてもオープンデータ政策の枠組みの中にしっかり位置づけていくというある種の非常に大きなバージョンアップというものが行われているところでございます。恐らく教育、あるいは教科書に関わることというのはそういった枠組みの中にも様々な形で関わってくるところでございますし、恐らくはほかの領域においても公益起点、個人起点というところは両方とも重要でありつつ、まさに果たして民間保有まで含めてどういったオープンデータというものを考えていくのかということの皮切りにもしていただけたらよいかと思っております。

それから、ちょっと関連しつつ、今日のお話からはちょっと枠としてははみ出るのでございますけれども、本日、教育データについて取り上げていただいて、しかし、恐らくこの後の展開として、研究データについてはこれから扱う回がないと思いますので、少しここで言及をさせていただくと、まさに今、オープンサイエンスの世界的な進展の文脈の中で、今年募集が始まる公的な研究資金の成果から、その根拠データというものをしっかりとオープンにしていかなければならないという義務づけがなされるというのは、研究業界の人間であれば御存じの方も多いと思います。そういった研究データの活用というのもまさにさきの欧州で言えばオープンデータ指令の中に明確に位置づけられているところでございます。教育に加えて、研究に関わるデータというのも、国がお金を出した公益的な性質が強いデータの活用の在り方というところで、この検討会として少し視野に入れていただいてもよいかと感じているところでございます。

私からは以上です。

○森田座長 ありがとうございます。

それでは、もうほぼ時間がなくなりましたけれども、今日御発表になりました先生方から今までの構成員の発言、意見について何かコメントか御発言がございましたら、お願いいたします。

では、藤村先生、どうぞ。

○藤村教授 様々なコメントをありがとうございました。私も大変納得しながら聞いていた部分です。

検討すべきこと、ガバナンスの話、その他、また、「教科書の単元名のデータのオープンデータ化」の話もどうもありがとうございました。それぞれ生かしながら進めさせていただきたいと思いますが、私どもだけで解決できない問題というのは先ほどの「共通認証基盤」の部分でございます。それについては皆様の御支援をぜひともお願いしたいと思います。

本日は誠にありがとうございました。

以上でございます。

○森田座長 ありがとうございました。

ほかにいかがでしょうか。よろしいでしょうか。

ありがとうございました。

今の部分につきましては、今後、デジタル庁で行われます認証基盤の検討の場におきましては、本日御意見が出ました制度面、システム面の課題も含めましてぜひ議論を深めていただき、その経過などにつきましては、またこの検討会にも報告して共有させていただければと思いますので、よろしく願いいたします。

それでは、次のアジェンダに移ります。次はAIとデータの利活用につきまして、上野山構成員から発表をお願いしたいと思います。上野山構成員、どうぞ。

○上野山構成員 よろしく願いします。

ちょっと資料をお映しします。すごくワーディーなので、皆さんお手元で御自身の資料を見ていただくほうがいいかもしれないですが、これは5～10分ぐらいでしゃべれば大丈夫でしょうか。

○森田座長 はい、どうぞ。

○上野山構成員 では、次のアジェンダでございまして、私からは「データ利活用とAI」ということで、民間の立場からデータ利活用に関して、特にAIの進化という観点からのデジタルの進化の捉え方の話をできればなと思っております。

といいますのも、先ほどのGIGAスクールであるとか、デジタル活用を教育においてという話においても、今、2025年でございますが、デジタルやデータ、AIというものは非常にムービングターゲットというか、日々動いておりまして、2030年ぐらいにAIやデジタルがどうなるかというのは、AIの業界ではボトムラインにおいては大枠コンセンサスが形成されていると思っております。もちろんさらに進化してしまうというシナリオもありますけれども、少なくともここまでは確実に行くというものを前提にこの資料を作っております。

まず、前提となる環境認識ということですが、2030年においてはAIというものが教育に限らずいろいろな領域で重要性をかなり強く帯びてくると思っています。現状認識としまして、現在2025年においても既に語学や数学、科学において、特に中等教育におい

てと書いていますけれども、これは教育の例ですけれども、AIを通じた知識獲得効率が既に爆発的に高まっている。これは非連続に2年前ぐらいに起きた変化ですけれども、例えば私も今、英語の学習というのを移動中、常にAIと会話しながらフィードバックをもらうということをやっておりますが、本当にこれが中高時代にあったら私の英語能力は本当に飛躍的に伸びたという実感の中でやっておるのですけれども、AIでは知識を獲得する効率においては既に人間から学ぶよりもはるかに効率的な分野が出現しているということでございます。このような分野においては近未来において、教員の方というのはただでさえ忙しいので、イメージで言うと人は隣でモチベーションや好奇心というリアルなモチベーションのところのサポートをして、知識は生徒さんがクラウドからAIで引き出していくという設計図に移行していくということがほぼ間違いない流れになっておると思っています。こういう時代において、AIによって人間が隅に追いやられるのではなくて、人間のより良い生活のためにAIを使い倒せるような社会の仕組みにアップデートしていかなければいけないという前提がまず環境認識の出発点だと思っています。この前提でAIやデータ利活用の方策を考えていく必要がある。

2章以降にその考え方を載せていますけれども、1章のところは2030年を見据えた環境認識の前提を書いています。先ほど申し上げたのは教育の分野の一例でございますけれども、これはフィジカルな世界においても同様で、AI掛けるロボットの領域で、農業や自動車などのかなり広範な範囲で影響が出る分野だと思います。実際に中国、アメリカ、一部フランスで非常に滑らかに動くヒューマノイドが増えてきております。知能化された新たなエレクトロニクスという分野が非常に花開いていく新たな巨大な新産業となりますけれども、この分野は日本の強みが非常に生きる分野であるので、データ利活用という文脈もここにつながってくると思っております。この新たな新産業においてデータがうまく循環し、人材や資金もスムーズに流れていくために、データ流通の方策を考えていくべきであるということかと思っています。なので、データ利活用における新たな制度設計においてはこのような視界が重要だとまず思っております。

続きまして、2章ですけれども、データというものを考えたときに、AIというものが非連続に出現してしまったがゆえに非常に議論が難しくなっていると思っています。AI時代のデータの利活用というのは、もともと個人情報というのはデータがあっただけなのですけれども、データレイヤーの上にアルゴリズムレイヤーとアウトカムレイヤーという3層構造、1層のものが3層に急に増えたので、新しい論点が複数出ていると考えています。この3層構造で考えるのがいいと思っています。

まず、データレイヤーのところでございますけれども、データというのは公共財としてのデータという新たな価値がAIによって出現して、2次利用に向けた制度変更が必然化しているというのが足元の捉え方だと思います。詳細は書いてあるとおりでございますけれども、これをもう少し申し上げますと、2層目のアルゴリズムレイヤーが、AIというのはデータを学習し、抽象化し、バロメーターとして取得する、つまり、例えば私の医療データ

というのはAIというものがない世界においてはあくまで私の個人情報であり、医療の統計データとして活用するぐらいの2次利用しかできなかったわけですが、現在においては私のデータやいろいろな方の医療データを集合知として集めて抽象化して医療AIとして学習させることで、誰もが高度な医療を受けられる、つまり自分のデータをクラウドに上げることによって全員がむしろ安全になるという逆向きのベクトルが出現し始めている。なので、アルゴリズムレイヤーという領域にデータが抽象化され、保存される新たな2次利用が出現していると思います。

この2次利用方法にも複数の種類があると思います。見える化、これはEBPMであるとか、統計データとして使うであるとか、今までもあった2次利用の方法です。医療AIというのは、ある種集合知的にみんなで育てることで個人にベネフィットが返ってくるという新たな2次利用方法ですし、あと、与信もセンシティブな2次利用ですが、似たようなバックグラウンドを持っている人であればこれぐらいの与信が妥当であろうという推論ができるようになってしまう。欧州であれば、ここにいわゆる民族や男女といった変数を入れてはいけないという議論になっておりますけれども、そういった与信という2次利用の方法も出てきております。なので、データレイヤーの上にこのアルゴリズムレイヤーが出現している。なので、ここに訓練データの透明性などの新たな論点が出ています。

3つ目がアウトカムレイヤーですが、これは実際に2次利用もどのように使うのかという出口の部分でリスクが段階的に変化するので、そのリスクの対応を明示化することが非常に重要になると思います。

ただ、これも活用にブレーキを踏むということではなくて、2次利用方法を類型化し、リスク対応を明示化することでむしろデータ利活用を促進していくという考え方かと思いますが、こういったことをきちんとすることで、新たに生まれた2層目、3層目の2次利用が加速していくと思います。

一旦以上でございます。

○森田座長 ありがとうございます。

それでは、今の発表につきまして、ディスカッションに入りたいと思いますが、また例のように手を挙げていただければと思います。

依田構成員、挙がっておりますのでどうぞ。

○依田構成員 すみません、先ほどと連続で意見になります。途中でちょっと早めに退室するので、早めに手を挙げさせていただきました。

上野山構成員におかれましては、いつも私が関わったデジタル市場競争会議で委員として活発に御意見いただいて、そのとき以来感謝しております。ありがとうございます。

上野山構成員には僕もデータサイエンティストなので非常に思うところがあって、上野山委構成員が今日言及してくださったのでそれに合わせて意見を述べると、2次利用というときに2通りの意味があると思っていて、今日の前半の会議の話題にもありましたが、データの2次利用というのは1次利用された個人情報を匿名化・仮名化して第三者に対し

て再利用するという2次利用があると思います。でも、上野山構成員が今日ここでおっしゃってくれた2次利用は、どちらかというと機械学習、あるいは統計的データの処理をした後に、バロメーターを出力させた後に、それをもっと社会的に利用する方向における2次利用だったので、似ているのだけれども違うところがあって、例えば私の医療データや教育データは、教師データとしてモデルの中に訓練するときには使われなかったのだけれども、誰かのデータを使ってバロメーターを推定した後に、その後から私のデータ等が使われて、私の知らないところでモデルとして、予測値として、私がどういう行動を取るとか、どのような振る舞いをするかというのをターゲティング的に使われる場合があって、意外にそちらのほうの機械学習をした後の出力のほうの2次的な利用について、このAIブームが来る前に実は個人情報においてもまだそんなに厳密に議論されていなくて、何となくいいだろうという感じでとどまっている場合があったので、そこについて、実は後者のほうのリスクはすごく今、高まっているので、せっかく上野山構成員が問題提起してくださったので、1点私の意見としてコメントいたしました。

以上でございます。ありがとうございました。

○森田座長 ありがとうございます。

ほかに御発言はいかがでしょうか。

では、今の依田構成員の御発言につきまして、上野山構成員、何かコメントはございますか。

○上野山構成員 まさに依田さんがおっしゃられたのは、私が申し上げている3層のアウトカムレイヤーにおける類型の一つに当たると思っております、なので、2次利用にも複数の種類があるということと、アウトカムのところも個人に対してフィードバック、何かしらアクションを働きかけてくるものとそうでないものがあるので、それぞれのリスク対応を明示化するというのは非常に重要になってくるということかと思っております。なので、ここで問題提起していることをアラインしておりますし、重要なポイントかと思っております。

○森田座長 ありがとうございます。

それでは、ほかに御発言はいかがでしょうか。

落合構成員、どうぞ。

○落合構成員 どうも御説明ありがとうございます。

AIに関するアップデートということで、デジタル臨調などの会議のときから含めている新しい議論をアップデートしていただいている、本当に改めて感謝申し上げます。

私のほうも本日、上野山構成員の御発表をお伺いして幾つか述べたいことがございます。

一つが、まずデータのアーキテクチャーの考え方が、本日の上野山構成員のお話もそうですし、前回の東京大学の大江先生の御発表もそういう側面があったかなと思っております、特に1次利用そのものというよりは2次利用のほうに着目した場合に、それを踏まえてデータの整備や用途を考察していく必要性が、AI時代に関するデータの整備として明らか

になってきているかと思っております。

今回、御説明いただいた中でも、どこまで構造化をしておく必要があるのかという部分は前回の大江先生のお話ともつながるところとも思いますが、もちろん1次利用の場合などを特に想定した場合にある程度標準化されていたほうがいい、AIで完全に執行されるわけではない場合も多いでしょうから、そういう場合も含めるとやはり標準化がされているほうがよいような領域は依然としてあるのだらうと思います。

一方で、AIの開発を考えていった場合に、もう少し非構造的なデータにもアクセスをして処理していけると考えた場合に、今日の先ほどの議論でも生貝先生がメタデータのお話もされておりましたけれども、どういうデータがどういうところにあってアクセスできるのかが明らかになっていて、アクセスをできるための手法まで用意しておけば、標準化という、それ自体でかなりコストや手間がかかる作業を必ずしも実施しなくても、データの基盤として有用になるような場合も見えてくる可能性があるのではないかとも思いました。改めて従来どおり標準化などもしっかりしていくべき領域があると同時に、有用な利活用に向けて考えられるような領域については、そもそも標準化まで行かないにしても、アクセスをしっかりできる、どこに何があるかを分かるようにする、まずはこれができるようにするだけでも、もしかすると一定の価値を発揮できるようなデータのレイヤーになってくるとも思いました。

ただ、この辺りは恐らくAIによる解析の進化の程度もあるでしょうし、さらに爆発的に速くなって行って、このように言っていることもどうだったのかというときも来るかもしれませんが、その進捗の程度によって本当に標準化の必要性が、全体のデータ構造の中でどう変わってくるのかは、少し考えながら議論をしていく必要はあると思いました。

また、今、申し上げたのがデータの構造についてですが、第2点としては、データの価値について、上野山構成員からは公共財的な価値が出てきたのではないかというお話がございました。私としては、さらに別の観点からも、公共的意義を見出していくべきと思っております。データ自体については、特にAIを使っていくことで、より高度な情報も得られるであろうということは、私もふだん例えばリサーチなどをするとき、人に聞くよりAIに聞いてしまったほうがいい場面は、最近非常に多くなってきているように思います。一方で、そこまで人を明らかに上回るようなところまで行かない使い方にしても、省人化や労力の削減につながる仕組みは、人口減少社会の中において極めて重要なリソースにもなると思います。このAIに利用され得るようなデータは、その意味での公共的な価値が、諸外国以上に日本の中ではより認識されていくべきなのではないかと思いましたので、そういったデータ自体に価値があり、様々な社会活動をデジタルによって代替していくための源泉になるところは改めて認識して、今後の議論を行えればと思いました。

私のほうは2点コメントさせていただきました。

○森田座長 それでは、稲谷構成員も手を挙げていらっしゃると思いますので、どうぞ。

○稲谷構成員 ありがとうございます。いつもいろいろと教えていただいてありがとうございます。

ざいます。

私も今の話と若干関係があることを少しお尋ねしたいなと思っているのですが、ちょいちょい仕事などで生成AIを開発されたりされている方などと話をすると、今日の話でいくと、ある種探索的な学習をさせる必要みたいなものがどんどん増えてきているという話をされていて、それが今日出てきている公共財的にデータを考えていくことで探索的に学習できる範囲を増やしていくみたいな話などともつながってきているのかなと思っているのです。データの取得範囲について過度に保守的にならないように考えるというところと密接に関係すると思うのですが、この辺りのトレンドというのは、今後、一層ある意味探索的にデータをどんどん学習させることで生成AIのアウトプットを改善していくみたいなことがもっと必要になってくるというトレンドになっているのかどうかということをお尋ねしたいなと思っています。というのも、仮にそうなのだとすると、今のデータの集め方に関する規制、特に個人情報に関してですが、そこでは利用目的制限が結構重要な役割を果たしているわけです。そのところをどういう形で考えていくべきなのか。すぐどうしたいという話ではないのですが、将来的にそういった学習の方向性ももし進んでくるとすると、リスクマネジメントの考え方、データのガバナンスの考え方にしても、少し利用目的制限みたいなところとは違う次元でまた考えていかなければいけない問題が出てくるのかなというところもあるのかなと思いはじめているので、少しお伺いをさせていただきたいという次第です。

よろしくをお願いします。

○森田座長 それでは、上野山構成員、今のはいかがでしょうか。

○上野山構成員 ぱしっと答えられるかあれなのですが、結局これはパーソナライズした何かを人間にフィードバックするという行為とそうではない行為によっても大分議論が分かれるはずですねというのがまず一つ前提で、AIのトレンドを先にお答えすると、これは諸説あるのであくまで私のあれですが、データを学習するというのは人間でいうと勉強をしているという感じで、机に座って勉強し続けていますというこのフェーズが徐々に終わり始めているのではないかという議論があって、もう一個の次のフェーズが、勉強するのではなくて自分で考えるというフェーズがあって、最近のオープンAIのモデルなどはリーズニングモデルと呼ばれて、勉強しているのではなくて考えているのです。考えるアルゴリズムをいかにうまくつくるかというゲームに移ってきているというのが2個目で、その次が、AIがいろいろとアクションして試してみるというフェーズに移ってきていて、重心で言うと徐々に勉強するから考える、試すに移ってきています。

これは人間も結構一緒に、ずっと勉強だけしていても会社で仕事できるようにはならないみたいな、考えたり、試したりするというバランスの議論というのがまずあって、徐々にもうこちらに移ってきているのですが、ただ、御質問の趣旨というのは恐らく個人へのパーソナライズにおいて何かしら侵害されるリスクがあるかないかという議論だと思っているので、それはユースケースを2つに分けて議論したほうが良いとも思います。

そこにおいてはあまり学習するデータがどうこうというよりは、結局アウトカムにおいて個人に何かしらのフィードバックを戻す行為ということをどう捉えるかという問いだけが常にあるということなのだと思います。

○稲谷構成員 分かりました。

データを集めるときに、今だと目的がこうでとか、限定して同意を取ってくるというやり方があるわけですね。その目的の書き方みたいなものとして何か工夫をすることによってクリアできるようになるのか、それとも、何かそもそもそのやり方とは別のやり方を考え始めたほうが、特に試してみたいなことについて様々なことが必要なのだということになってくると、あるのかといったところがお尋ねしたかったという趣旨ですね。

○上野山構成員 それで言うと、目的の一つに御自身に何かしらフィードバックを与えてくる行為に関してどう扱うかというのがすごく重要な目的としてあって、それ以外の目的というのはその人個人の権利を侵害することがあり得るのですかね。

○稲谷構成員 いえ、まさにその論点と関係して、個人情報保護法が何をどういう形でリスクマネジメントしてきたのですかという問題がまさにそこと関係して構造化されてきている話だと思うのですが、立てつけ的に今まではそういった具体的に、例えば情報を渡す人、データを渡す人にどういうリスクが具体化するのか、それをどういう形でコントロールするのかという観点から必ずしも議論されていないところがもともとあったと思っています。

○上野山構成員 なるほど。

○稲谷構成員 そのところが多分こういう議論をやってくることによって、逆に言うところの文脈で実際に、今、まさにおっしゃったように個人に対してアウトプットをどう返すかということ以外に何か問題が起きる可能性があるのかという形で議論をクリアにしていくと、もともとそこで言っていた目的特定や同意の取り方というもののそのものの在り方というものも、まさにAIの時代になってくるとそこで問題としていた事象がさらに可視化されてくることによって、ガバナンスのやり方としてどういう形でアップデートするならアップデートする、あるいは何かより良いやり方があるなら考えるみたいな形で議論していくことになるだろうということなのではないのかなということも含めてです。ありがとうございます。

○上野山構成員 なるほど、ありがとうございます。

ちょっと質問の答えではない補足ですが、こうなるだろうということでは、結局自分のデータを出さない人は便益を享受できないということに基本的にはなっていくと思います。具体的に言うと、例えば英語学習を私はすごくやっているのですが、自分が覚えている単語のリストを入力しているんですね。自分が覚えていない単語をレコメンドしてくるわけですよ。これは一番分かりやすい例ですが、それを開示したくないという人は、自分が覚えるべき単語をサジェストしてもらえないということなので、学習が遅くなるということになるわけです。そういう感じには基本的になってくると思

います。

○稲谷構成員 ありがとうございます。

○森田座長 ありがとうございます。

時間が参りましたけれども、落合構成員、また手を挙げていらっしゃるけれども、簡潔に最後ということをお願いします。

○落合構成員 簡潔にですが、今、稲谷先生の議論を聞いていて、データのほうでも医療のほうで出口規制という話がありましたが、プロセスでどのようにデータを集めていくかというより、それを最終的にどう使うか、そのときに直接人に使うのか、自動的に利用するのかといったリスクを見ていったほうが合理的になるのかな、とも思ったのですが、その点、どう思われますでしょうかということが、上野山構成員に最後の御質問です。

○上野山構成員 ちょっと個人的な意見ですがけれども、まさにそうだと思っておりますということです。今日の資料もそこが結構ポイントの一つだと思っています。

○落合構成員 分かりました。ありがとうございます。

○稲谷構成員 私もよく分かりました。ありがとうございます。

○森田座長 ありがとうございます。

それでは、闊達な御議論をいただきましたけれども、これくらいにさせていただきたいと思えます。最後は少し難しい議論になりましたけれども、私自身はAIはいろいろ使われる場合に目的で制限をするのではなくて、むしろリスクベースアプローチといった形で考えるべきではないかという議論になってくるのではないかと理解いたしました。

それでは、最後に事務局から何かございましたら、お願いいたします。

○山澄審議官 日程でございますけれども、次回、第7回の検討会は、3月12日の4時から、産業分野のデータ利活用ということをテーマに開催させていただきます。引き続きよろしくお願いいたします。

○森田座長 それでは、今日も長時間にわたって御参加いただきましてありがとうございました。

以上をもちまして、本日の「データ利活用制度・システム検討会」を終了いたします。ありがとうございました。