

教育データの利活用に関する有識者会議
委員のこれまでの主な御意見（※）

※事務局にて整理

教育データの対象

- 教育データは、コンピューターのログとして取れるものに矮小化してはならない。学校の先生の普段の仕事の中での情報と、機械で取れるログをどうつなげていくかが非常に重要。ありとあらゆる事実関係の情報、先生の見取りを、うまく教育データの利活用に入れていく必要がある。
- 数値化されるデータばかりではなく、ポートフォリオなどの、子供たちの製作物や、学習課程の中で作成される成果物、子供たちが対話をしながら自分の意見を変容させたり、他者との対話により思考が広がっていったりという見取りをどのようにデータ化して、これから教育の中に利活用していこうとするか。
- 出欠など従来から集めているデータに加え、表情、行動、PCの活用状況、教室の気温等、新たにとれるようになった教育データも考えていくべき。
- データの分類としては、①学校に持ち込まれるデータと、学校にいる間につくっているデータ、②保持しておく必要がありずっと持っているべきデータ、③卒業後には不要となって削除してもよいデータ。校内だけで利用するデータと校外でも利用するデータなどの分け方で整理していく必要がある。
- スタディ・ログに加え、いじめの早期発見、不登校傾向の早期発見、健康的な指標等 PHR などのライフ・ログ、先生がどのようにその子に合わせて適切な指導をするかという処遇情報を収集したアシスト・ログの3つの教育データを活用して、全ての先生を支援してはどうか。

教育データの目的

- 教育データの利活用は、本人の学習のために使うというのが一次利用で、全学習者が対象。二次利用はビッグデータ分析によって指導方法や教材開発を行うなど、世の中全体をよくするという話だが、二次利用に使うデータは、一部がデータを提供すればよく、全員のデータである必要はない。全体として見れば一次利用の方が二次利用と比べればはるかに価値が高く、かつ一次利用ができれば本人同意でそれを提供することもできるため、二次利用もできる。
- 情報の一次利用と二次利用の整理をすべき。一次利用は、学習者のためのスタディ・ログ。二次利用は学校や社会のためという活用の方向になる。一次利用を定義していかないと、二次利用の形が大きく変わってしまうと懸念している。一次利用まで定義して議論をすべき。一次利用の情報を定義し、学習者にとっての個別最適化が進めるべき。
- 大学教育の場合のラーニングアナリティクスの視点では、教育は学校教育を中心に回り、LMSを導入して、そこで蓄積されたデータをうまく共有、活用する話からスタートしている。教育データは、指導を行っている先生のためでもあって、その先生の教育をいかに支援するかという視点も重要。特に公教育では、どのようなデータをどう活用するかを、まず先生のほうが決めていかないといけない部分もある。まずは先生をサポートするために教育データを使うということが、学校教育においては特に重要。

- 学びは本当に個人ひとりのものか。「一次利用を考えて学習者のために」というときに、学習者は教室内で集団で学んでおり、それをどう個人に返すかが複雑である。学習者が自分の学びを使うときに、どうしても成績など個人化しやすいものだけがデータとなってしまう可能性がある。その折り合いをどうつけていくかが一次利用と二次利用の難しさである。

教育データを活用したこれからの教育・教育政策の在り方

- 学校教育でやるべきことは学びを追うこと、成果を生み出しているはずの教え方、そのモデルを見直していくこと、それを社会全体がやっていくことの3点。今、公教育制度から生涯学習、子供が自分で学びたいことを決めていくという転換期にあり、そうできる子とそうできない子の格差が広がりそうになっている。教育データ標準化により、今後の学校教育の充実、そこからの生涯学習への接続につながっていくべき。
- 教育データを用いて、先生たちは授業改善を行い、子供たちが1人1台を持って自ら学ぶための授業改善に使う。子供にとっては自己を認知するために、データを見て自らの学習を計画していく、そのように子供たちを育てるべき。
- 「経験と勘と気合い」のみに基づいた指導から「客観的な根拠」に基づいた指導への船出をしていき、教員の個人プレーに頼らず、データを活用して教室や授業を科学していくべき。
- 科学捜査、医師の血液検査のようなデータに基づいた教育施策がこれから不可欠になってくる。

教育データのユースケース

- 教育データ利活用の目的と期待は、本人の学習効率を上げるという個人の論点と、教材開発、教授法の改善という全体の底上げの両面があり、それぞれ重要。
- 子供たちや教員、保護者の一日の活動の中では、①OS アプリを使った健康観察、AIドリル活用によるスタディ・ログや出欠確認、心理状態の可視化、②朝自習の読書データと学習の相関を取り、県の読書推進事業との連携、③子供たちが朝の会での1分間スピーチなどをポートフォリオとして蓄積、④加えて日課表や連絡帳なども連携すると、保護者支援にもつながる。授業においては1人1台端末を使ってよりよい教育をすることを目指す中で、各教育支援システムの中のフォルダを教科ごとに分けたり、子供たちが個人のデジタルポートフォリオを作成したり、プログラミングのデータやスタディ・ログの蓄積ができる。こうすると一日活動して見えてくる活動データが明らかになる。
- スタディ・ログにより、学習者一人ひとりの個性や特徴、興味関心や学習の到達度に合わせた形で個別最適化を行える。学習者が自身のデータを自ら振り返り、リフレクションを起こすことで、主体的で能動的な学習者になっていく。
- 教育データ利活用でやりたいことは、単に問題を解けるようになるだけでなく、それを結びつけて一人一人が理解していく学習のプロセスをその場で詳しく分析していくこと。特にインプットだけではなく、本人のアウトプット、インタラクションを一緒に取っていくこと。学習のプロセスをロングスパンで追っていくこと。たくさんのビックデータにおける一人一人の学びの道筋と、それをまとめた理論の抽出を行うこと。
- 学習者のデータの単位は、学習者個人で考えるはずだが、あの先生のとときにこの人と学んだというようなことを、たくさんの名簿のテーブルを保存して、本人の学習プロセスを記録していくようなデー

タの取り方もあるのではないか。

- ICT 活用により、評価の在り方がどのように変化するかとの点で検証が必要。データの背景にどのような認知の過程があるのかを明らかにして、評価者の共通認識を作っていく必要がある。薄い個別最適化論、軽い授業論に現場が依存してしまうことを危惧している。改めて学習の本質に立ち戻ったエビデンスに基づく授業デザインが見えてくることを期待している。
- 誰がどこまで理解していたかということをも可視化して、先生や児童生徒に共有することは非常に大事。
- 学習履歴を使ってできることとして、例えば学習者の振る舞いのモデリング、評価やフィードバックの改善などがある。こうした個別最適の学びに加えて、学校教師などの事務作業の簡便化、負荷軽減も目的。
- 「匠の技」の可視化について、例えば発問や声かけといった教師の指導活動を記録して、子供の反応、つまり学習活動との関係を分析することで、優れた教師はどんなタイミングで誰にどんな内容を伝えているのか、それによって子供たちの学びはどう変容しているのかなどを可視化・分析できないか。
- 国全体でデータ共有することで、先生や児童生徒などに対して、エビデンスに基づく教育方法や学習方法を推薦できる。民間の教材会社等の教材の改善に役立てることもできる。また、研究者は、デジタル時代における効果的な学習方法や教育方法が分かっていく。
- 特別支援教育は一部のベテラン教師の経験と勘によって強くリードされる現状があるが、特別支援教育にこそデータ利活用が重要。
- 学習データを利用したい主体は学校や親、学校や教材会社だけとは限らず、奨学金を交付している団体、金融機関等が教育ローンの貸付け、雇用者が雇用をするに当たって利用したい、というニーズもあり得る。

教育現場で教育データを利活用・解釈する際に必要な視点

- データの利活用の際は、どんな資質・能力をつけたいかを明確にし、このデータから育成したい資質能力をどう図るかといった目的意識を明確にしていくべき。データ利活用をする上では、授業をどんな思考で子どもたちが取り組むべきなのか、場面ごとに見ていくべき。
- 効果的・効率的な授業の構築、そして、先生の発問、授業中の焦点化、どんな既存の知識があつて、どこに不足している情報があるので、どんな発問でそこに紐づけていこう、もしくは授業の内容を焦点化していこうか、それから下位層の子供たちにはどんな支援を用意しておくか等、そこが重要になってくる。そして、授業後の理解度を確認することで、子供たちは自分の理解度を知り、家庭学習に反映する、そして保護者はその様子を可視化された情報で確認することができる、そういったフィードバックをしていくべき。
- 子供たちにデータをどう返していくかということと、ステークホルダーである教員に対してリコメンドを出して、対面の授業も改善していく視点も大事。
- データに振り回されないようにするよう必要がある。得られたデータをどのように日々の授業等に落とし込んでいくかが重要。量的なエビデンスは優れていて、質的なエビデンスは曖昧だからあまり優れていないという考え方は適当ではない。現場で、目の前の子供と日々対峙して、授業をやっている

る中で生まれてくるものは、教師の非常に優れた認知機能の結果であり、現場主体のデータを蓄積して活用する必要がある。

- AIドリルで問題を解けるといことが本当に理解したことになるのか、子供のためにAIドリルで弱点や、自動的に問題のリコメンドが出るが、それを指導に生かそうと思うと、そこが精緻化されればされるほどブラックボックスになり、教え方に反映しづらい。感情は顔のマークで反映できるか、グルーピングで話量をうまく上げるようにしたことが本当に理解につながるのか、こういう疑問をテクノロジーの導入と一緒に解消していく、そこにデータが利活用できることが大事ではないか。
- データの利活用に関して、数値的な目標設定が重要。例えば保護者から見た時に学力テストの成果が今よりも10点あがる20点あがる、教師の負担が減って教師の残業時間が20パーセント減るなどが考えられる。

教育データを利活用する際の法的・制度的論点や留意点

- 何を目的とするか、誰がステークホルダーか、事実関係がどのようなものかによって、法的な論点は全く異なってくる。仮のユースケースをまず作って、そこから法的な論点を検討し、必要な対策を打つ、というプロセスを早く打ち出すことが大切。場合によっては新たな立法や既存の制度の改正などが必要になり得る。
- 個人情報や学校などから別の学校などに移していく場合を想定すると、現時点では、個人情報保護条例2000個問題や、国立、公立、私立の違いを考えていかなければならない。
- 匿名化をうまくやらないと、ビッグデータに加工できない部分がある。個人情報は、非常にデリケートで、これまでも学校の先生たちは、子供たちの個人情報をうまく把握することによって適切な指導をしてきたという歴史がある。有効利用と匿名化、個人情報の部分、データの所有権の問題がある。これらをはっきりさせない限り、自分のデータを開示し、塾や民間の教育産業で更なる手厚い指導をうけるということも実現しなくなる。
- 学校の児童生徒は未成年のアカウントであるため、様々な課題が出てきている。進級や学校が変わるときにデータがどうやってひもづくか、パスワードも子供が管理するのではなく、保護者や先生が管理したりする。データそのものの扱い方も、インプットの時点でたくさんの課題が出ています。データの扱いの合意なども、手間取っているのが現状。
- 個人情報について検討しなければならない点が多く、児童生徒の学習データが教師や学校、教科書・教材の出版社等に対して出され、学校や企業の方がその情報を使って児童生徒にアプローチをしていく場合、情報銀行やPDSのような仕組みを取るかどうか論点。
- 個人情報として蓄積するものと匿名加工できるものの切り分けをどのように行うかも論点になる。
- 個人情報として利活用する場合には、教師、学校や企業側としては、生徒に対して、実質的には親権者に対してかもしれないが、利用目的をどのように説明してどのように同意を取るか、あるいは同意を取らなくてよい道をどう探るか、という点も論点になる。
- 公教育と民間教育の分断が、学習者にとって不利益を与えており、今回の教育データの利活用を機に、データのオープン化を図り、民間教育の知恵を取り込む、公教育に取り込む、公教育を開き、そして民間教育の知恵を取り込むという視点で、データ設計が必要。民間データの囲い込みを禁止するために、どうやって相互データを利活用するかが重要。

教育データ利活用のためのポリシー・原則の必要性

- ①学習者のためにデータを集めること、②学習者のために使えるよう提供する義務があるということ、③集めるデータの質を上げておく必要があること、という3つの原則が重要。
- コロナ禍の影響でオンライン授業が大学を中心に行われており、1人1台の情報端末の環境で、大量の教育データが蓄積されつつあるが、教育データの取扱いについて明確な指針がない。研究動向としては、教育データを分析して教育改善・学習支援の効果を出し始めている中、教育改善や学習支援に有効活用するための指針を提示することが急務。
- 教育データが標準化して流通したときに、どこまで、誰まで流通させて良いのか、どういうデータを誰まで開示して、どういう責任で使っているのかというルールを決める必要がある。
- 今の公教育を小学校1年生と中学校3年生で同じく語るのは少し乱暴で、データに関しても発達段階に応じた利活用が重要。アメリカには、COPPA（児童オンラインプライバシー保護法）というものがああり、13歳未満とそれ以上で分けている。
- 保護者の立場では学校で収集した教育データを学校外で活用したいというニーズがあるため、そのためのデータのやり取りに関するルールづくりが必要。

教育データを活用するためのサポート体制

- 教育にデータを利活用するには、これらが簡便に分析できるようなシステム化、取得したデータを分析するための教員および児童生徒のデータリテラシー教育も並行して考えるべき。
- 教師に授業デザインを定着化させていくことが、これからの取組の一つの目玉になる。データリテラシー研修も併せて行う必要がある。
- 得られたデータは過去のものだという認識のもと、それを現在や未来に生かすために解釈して策を講じられるデータと現場のつなぎ役が必要。1つの自治体でそのような人材を採用するのは難しく、複数の自治体で共有できる仕組みが必要。

教育データの標準化の意義・目的

- すでに国際的に様々な標準化が行われている中、教育データの標準化が必要であることは間違いがない。
- ドリル等の学習活動データだけでなく、複合的で総合的な学習活動に関するデータや資質・能力の育成に間接的に寄与する活動に関するデータもあり、膨大なデータをどう扱っていくか課題。システムごとに色々なデータの形式が不統一で活用する段階も不統一だと感じており、データの標準化が検討課題。
- 教科書がデジタル化し、他のデジタル教材とうまくリンクできるようにするには、データ形式が標準化され、参照テーブルのようなものがないといけない。そうした観点から、学習指導要領がコード化されることには意味がある。
- 働き方改革に資する負担軽減の取組として、学校基本調査などの各調査の回答の準備期間を長くすることや、ワンソース・ワンマスタの考えで重複の調査を避けること、オンライン化を推進し、全体

的なボリューム感を減少させるべき。

- 学習者の履歴をコンピューター間で理解するために、例えば学校 ID、学習内容 ID、学習要素、学習指導要領コード、学習者 ID がデータとして整理されてくると、これらを紐づけた形で、様々な履歴を集めて、リコメンデーションをすることが可能になる。
- 講師や非常勤講師、また、本務者や兼務者などの定義が、国の調査と県の調査とで食い違いがある。また、教職員数の数値のカウントの仕方が調査によって異なるため、用語、定義等の明確化の検討をするべき。
- 学習の評価を使いながら、一緒にみんなで改善していくために教育データ標準化がうまく働けばいい。
- 教育データを集めるための手順として、まずは教育データの標準化を行って制度設計し、次に、公教育を対象に教育データを各学校で収集し利活用していく。そのためには LMS や LRS の各学校への導入が必要。その次に国でデータを共有し、その後で民間教育のデータを扱うことを考えてはどうか。データの共有を考えると、教育データの標準化が重要。

教育データの標準化の際の留意点

- データの標準化と合わせ、グローバルスタンダードを目指していくべき。教育に特化しすぎたデータの標準化は適切ではない。標準化は様々な動きを把握して行っていくべき。
- One Roster や Caliper 等の国際標準仕様に沿うことで、日本がガラパゴス化することなく、様々な知見を取り込める。
- 教育データ利活用の本質的な目的とそれをどのように実装するかという具体的なユースケースの検討に加えて、教育データの標準化は非常に重要なため、実施可能なところから実施していくべき。
- IEEE の LOM については、国際的な場での議論により、データ項目が増えていったが、実際に活用される項目はそんなに多くない。例えば単元の ID などは非常にコアな部分なので、それを含めた、ミニマムのユースケースを洗い出して優先順位をつけて、それに従ってメタデータをつけていくのが、現実的な路線ではないか。
- 教育のデータ標準化は一回策定して完成ではない。特に活動情報のところは、どういう情報を取ることが実践の現場にどういう影響を及ぼすのかを注意深く見ながら、何度も標準化をやり直していくことが必要。
- 教育データの標準化として、xAPI という形式で、LRS にデータを蓄積するべき。様々なシステムを使っても LRS に 1 つのデータとして保存することができ、何十年とデータが蓄積されても後から分析できる。
- 現在、様々な学習のリソース、クイズに対して、メタデータを付与して整備していくことが推奨されているが、それを使って学習者にどういうリコメンドしていくかという議論が、まだ成熟していない。

学習者の ID について

- GIGA スクール構想が急速に進み 1 人 1 台が導入される中で、様々な教育委員会やベンダーから個人を特定するユニーク ID の必要性について言及されている。最終的にはユニバーサルなものが望ましいと考えており、マイナンバーそのものではなく、マイナンバーに伴うマイキーID 等、そのよう

な ID を将来的に検討するべきではないか。

- 学習履歴の分析は学習者 ID が整っていないとできないのではないかな。
- 従来の学籍番号などは進学のプロセスで変わるものだが、学校に入ってから社会に出ていくまでの間に一貫した ID で管理できるのか、また個人情報保護の観点との両立をどのように考えていくか。
- 各サービスから本人にデータを提供すればそれが名寄せになるので、全てのサービスで統一した ID を持たなくてもよいのではないかな。

教育情報システムの在り方について

- 世界各国の教育データの利活用や、国内のこれまでの利活用では、学習系は学習系で、校務系は校務系で分離され、部分部分での利活用となっているが、今後の教育データの利活用を考えると、全体を見渡した教育データの利活用を考えるべき。
- データが各所に散在しており、データを利活用するためのサイクルの中に定着しておらず、学習履歴は取得しているが、効果的に活用していないことが課題。このため、データを集約するプラットフォームが必要で、1人1台の新しい学校の生活様式の中に組み込まれているということが大事。
- 一人一台の端末利用を前提に、初等教育から高等教育まで、校務系データだけでなく、細かな学習系データも含めて、教育データを収集するのは世界初であり、また全国で共通の LMS を用いずに国際技術標準のデータ書式でデータを収集するのも世界初であるため、2つの世界初を目指してはどうか。
- まずセキュアでポータビリティのある仕組みであること、データは保護者も含めて学習者のメリットのために使うこと、学習者のデータを第三者が使う場合にパーミッションを取ること、他業種や世界を含めてなるべくプラットフォームや技術の面について教育特化型にしないことが必要。

教育情報システムの普及に向けた課題

- 小学校、中学校、高校は大学と違ってシステム管理者がいないので、できるだけ人的等の運用コストを低くする必要がある。
- 教育がひとりにつき十数年という長期間である中、システムのライフサイクルは5年～7年であり、そのギャップを埋めることが大きな課題。
- 様々な自治体が教育支援システムを導入していて、自治体が目指す教育を行っている。それらを標準化して、一つの国の仕様になっていくことが可能なのか。それを待っている間に既存の教育支援システムを用いて先に進む中で、どう整合性を取っていくのか。
- 多くの学習のリソースがある中、デジタル化がまだ追いついていない部分もある。有効活用するためにはメタデータと情報、データの流通、レイヤー間のデータの相互利用・乗り入れが必要で、メタデータをどう付与するのかや、付与に対する人件費や金銭的成本、時間、手間などを誰が負担するのかが大きな問題。
- 現在のリコメンデーションは、例えば1つの教科書、1つのドリル、あるいは1つのベンダーのコンテンツの中で閉じているケースが多いが、よりオープンなリコメンデーションを考えた場合には、マイクロコンテンツ化して、個別の個々のものとして扱い、それをアダプティブに提示していくフレームワークが必要になる。そのマイクロコンテンツに、データインフラを紐づけていく必要がある。こ

これは、マンパワーを含むコストをかけなければ実現できない。

- 学校が取得したログを塾でも使おうとすると、学校にはログを取得させないという自治体も多いのではないかと。設置者が自治体である以上、国がどこまでこれを決めて動かせるのかという現実もある。
- 学校健康診断の電子化について、LGWAN 外の校務支援システムは、ネットワークの切替えなど大変煩雑な作業が必要になる可能性があり、学校や教育委員会の負担にならないようにすべき。

教育データの流通について

- 教育データの収集方法は2つあり、1つ目が分散方式。各学校に LRS を置き、学校単位で管理・活用しそれを教育委員会などの自治体で共有して、外部に出すときは匿名加工情報に変換して、国全体で共有する方式。2つ目がクラウド集中方式で、教育用クラウドを作り、そこに LMS や LRS を1校につき1つ提供し、データを収集していくアプローチ。この方法で実現すれば、共同調達で安く済むのではないかと。
- データの管理運用のためのコストや利便性を考えても、全員の学習データを集中管理するよりも、各学習者ないし保護者にデータを集約して分散管理する方がセキュリティ面で安全。また、本人あるいは保護者がデータを管理することで、本人同意、管理者の同意だけでデータを扱えるので便利。
- 学習に関連するデータを集めても、例えば勉強できるようになるかどうかは生活習慣の問題である場合も多く、学校のデータだけでは足りないため、様々なデータを名寄せして使う必要がある。本人にデータを集約すれば自然に名寄せされてデータの価値が高まる。

教育データの分析（ラーニング・アナリティクス）について

- 学習履歴分析は、2010 年頃から国際会議も立ち上がり、xAPI や Caliper などの技術標準ができてきた。英国では学校評価に近い目的があったが、2010 年代後半になると、諸外国では、取得した履歴をどう使うかというナショナルプロジェクトが走り始めているという状況である。
- データとして使いやすいのは、LMS や CBT に蓄積してあるログである。粒度の細かいものになるにつれて、データを蓄積することはできるが、利用方法の知見がたまっていない。
- 教育データを基本情報とオプションな付加情報に分け、基本情報を集めていくべき。
- データを集める際には、まず教科書・教材のデジタル化、それから試験やノート、連絡ノートなど、基本的なところもデジタル化して、できるところから情報を集めていくことが大事。
- 全てのデータを集めるのではなく、OS や Office などの一般ツールのデータは集めず、学習活動や教育活動時のデータを対象にして、それらを使ってエビデンスを蓄積して共有するべき。
- 出欠くらいの粒度の粗いデータではなく、登校時刻のように詳細化しないと分析が難しい。
- ビックデータを活用した量的エビデンスが果たす役割は大きいと、教室レベルでの教育的な営みとしての成立の度合いや、よい授業の秘訣を観察また分析することが大切。そのために、日々の授業で得られる「スモールデータ」や、より質的な教師の行動分析等も重要。
- 各種のデータを集めていくときに、学びの深まりをどうやって捉えていくかが大きな問題。学びの深まりを学校教育の役割と照らして考えることが全体像を考えるのに役立つ。

教育データを利活用するための環境・体制整備

- 海外ではエビデンスの共有検索サイトがあるが、日本に同様なものがないため研究者と実践者、政策立案者がエビデンス、データに基づいて、関連、リンク、連携をうまく取れるようにできるサイトを構築すべきではないか。
- 全国各地で自治体や学校、大学、企業等が連携して互いの意志や納得感を共有しながら研究をしていく必要がある。
- 教育データを収集・利活用するため、研究者、学校企業などがうまく連携できるように、協議会をつくってはどうか。企業や研究者、現場の先生等、様々なプレーヤーをうまくコンソーシアム化して、情報を共有するという仕組みが必要。
- 基礎自治体のデータは基本的に悉皆、経年のデータで、再生性が大きいという特徴がある。現在は学校現場での実践を重視しているが、今後、自治体の持っている行政データの活用が課題。

(以上)