

## 提 言

# 教育のデジタル化を踏まえた 学習データの利活用に関する提言 —エビデンスに基づく教育に向けて—



令和2年（2020年）9月30日

日本学術会議

心理学・教育学委員会・情報学委員会合同

教育データ利活用分科会

この提言は、日本学術会議心理学・教育学委員会・情報学委員会合同教育データ利活用分科会の審議結果を取りまとめ公表するものである。

日本学術会議心理学・教育学委員会・情報学委員会合同  
教育データ利活用分科会

委員長	美濃 導彦	(第三部会員)	研究開発法人理化学研究所理事
副委員長	楠見 孝	(連携会員)	京都大学大学院教育学研究科教授
幹事	谷口倫一郎	(第三部会員)	九州大学大学院システム情報科学研究院教授
幹事	緒方 広明	(特任連携会員)	京都大学学術情報メディアセンター教授
	遠藤 利彦	(第一部会員)	東京大学大学院教育学研究科教授
	西田 眞也	(第一部会員)	京都大学大学院情報学研究科教授
	松下 佳代	(第一部会員)	京都大学高等教育研究開発推進センター教授
	柴山 悦哉	(第三部会員)	東京大学情報基盤センター教授
	宮地 充子	(第三部会員)	大阪大学大学院工学研究科教授
	乾 健太郎	(連携会員)	東北大学大学院情報科学研究科教授
	菅原ますみ	(連携会員)	お茶の水女子大学基幹研究院人間科学系教授
	原田 悦子	(連携会員)	筑波大学大学院人間総合科学研究科教授
	藤村 宣之	(連携会員)	東京大学大学院教育学研究科教授
	前田 香織	(連携会員)	広島市立大学大学院情報科学研究科教授
	美馬のゆり	(連携会員)	公立はこだて未来大学システム情報科学部教授

本提言の執筆にあたり、以下の方にご協力頂いた。

曾我部真裕	京都大学大学院法学研究科教授
小林 正啓	花水木法律事務所弁護士
久富 望	京都大学大学院教育学研究科助教

本提言の作成にあたり、以下の職員が事務及び調査を担当した。

事務局	松室 寛治	参事官(審議第二担当)
	五十嵐久留美	参事官(審議第二担当)付参事官補佐
	加藤 雅之	参事官(審議第二担当)付審議専門職付

# 要 旨

## 1 作成の背景

現在、教育機関においては GIGA スクール構想の推進や新型コロナウイルス対策でのオンライン授業の実施など、「教育のデジタル化」が行われている。これによって、各教育機関や民間企業には、デジタル教材の閲覧履歴やオンライン試験の答案など大量の学習データが蓄積されているが、これを教育改善のために活用することはあまり実施されていない。例えばデジタル教科書の閲覧履歴やデジタルドリルの答案データを解析することで、いつどの問題で分からなくなったか、その概念がどのようにして理解できるようになったか等が把握できるようになる。さらに個人情報保護に配慮して適切に処理して国全体で学習データを共有することによって児童・生徒・学生（以下学習者）の多面的なデータの分析が促進され、エビデンスに基づく教育が実現できる。そこで本提言は心理学・教育学、情報学の立場から学習データの種類やその活用の必要性について論じ、学習データを活用するための制度設計や支援体制、基盤情報システム等のあり方について提言するものである。なお、学習データの活用はラーニングアナリティクス（以下 LA）に限定せず、教学 IR (Institutional Research) を含む。

## 2 提言

本提言は、文科省・経産省・総務省などの国の政策に向けたものであり、以下の4点にまとめることができる。なお、本提言は、一人一台の情報端末や学習支援システム (LMS: Learning Management System) 等を用いて蓄積される学習データに限定したものであり、教育活動全般で利用される広範な教育データの取り扱いについては、慎重に議論を深め、継続して検討する必要がある。

### (1) 学習データの種類と教育改善のための利用

本提言では、学習データとは、教育のデジタル化によって、一人一台の情報端末を使って LMS や校務支援システム等を用いて蓄積されるデジタル情報と定義する。これは校務系データと授業・学習系データに分類できる（表1）。この学習データは教育・学習活動の一部を切り取ったものであり、教育の効果や達成度などを全てはかることはできない、ある意味限定されたデータある点に注意が必要である。また、学習データの利用は、個人では教員や学習者の教育・学習活動の効果向上をさせるため、教育機関ではカリキュラム等の最適化をはかるため、国全体では政策立案者や研究者・市民などが教育効果を高めるために利用するため、に分類できる（表2）。学習データの収集は、それが比較的行きやすい小学校から大学・大学院までの公教育を対象として集め、各教育機関等内で、表2の教育の改善に活用すべきである。そして各教育機関内部で学習データを個人情報保護に配慮して適切に処理して、国全体で収集する方法を提起する。国全体で共有すべき学習データの基本項目としては、これまでの LA の研究で共通

して用いられ、比較的共有が行いやすい項目（表 1 (1)～(9)）を提案する。

## (2) 学習データを収集・利活用するための制度設計

学習データの利活用の実現は、まずは教育機関内から始めて、次に、国全体で利活用するよう慎重に議論しつつ段階的に進められるべきである。その際、国が制度設計を行う場合に注意すべきこととしては、①教育機関が民間企業等のシステムや試験等を利用する場合、企業が学習データを囲い込む恐れがあるため、学習者の学習データを学校に提供するよう、契約時に盛り込むこと、②国全体で学習データを収集する際には、「データを個人情報保護に配慮して適切に処理して収集する」、「データの利用者は、個人を特定するような分析や地域や学校間の無用な比較をしない」、「教育に係る選択は本人が実施するものであり、学習データを利用した推薦や提案が本人にとって決めつけや押し付けにならないようにする」などである。一方、学習データの共有や利活用を促進するためには、データの書式や意味（言葉や数字の解釈）の標準化、学習要素の ID 化や、教育評価などで使う用語の統一などが必要である。この制度は、パブリックコメントの収集等により国民全体で議論しながら、学習者の人権と個人としての尊厳を脅かすようなことのないように、制度設計や実施がなされているかを倫理審査委員会のような第三者機関を設けて、学習データの悪用を防ぎつつ適切に利活用が進むように定期的に見直していく必要がある（参考資料 9）。

## (3) 学習データを収集・利活用するための情報環境の整備

学習データを効率的に収集するためには、情報端末を一人につき一台整備し、デジタル教科書や LMS などを各学習者が、授業内外で常時使える環境を整える必要がある。また、現在の初等中等教育の学校ではネットワーク接続に制限が多いため、各地方自治体での関連法令等を見なおし、クラウド環境等を適切に用いて学習データを収集・利活用するための情報基盤システムを構築する必要がある。加えて、このような環境の整備と同時に、情報モラルやセキュリティ教育の提供や、健康やメンタル面への配慮も必要である。

## (4) 学習データを収集・利活用するための人材の育成

学習データを教育現場で利活用するためには、それをうまく活用できる教員及び、教員をサポートする LA 専門員の育成が必須である。これには現職の教員への研修や大学の教員養成課程において、学習データを利活用したカリキュラム・授業・評価の設計・実施の方法をしっかりと教えていく必要がある。また、蓄積された学習データを十分利活用するためには、効果的な教育・学習方法、AI による新たな支援方法等を見いだすことができる、学習データサイエンティストを養成する必要がある。そのためには、新たな大学院の創設や、既に実施されているデータ科学教育との連携が必要であろう。この研究者の主な役割として、どのような学習データをどのように利活用することによって教育のどの要素が良くなったか等を継続して検証していくことも重要である。

# 目 次

1	はじめに	1
2	現状と問題点	2
(1)	学習データの利活用の現状と問題点	2
(2)	学習データを収集・利活用するための制度の現状と問題点	7
(3)	教育現場における情報環境整備の状況	12
(4)	教育現場における人材の状況	14
3	提言について	16
(1)	学習データの種類と教育改善のための利用	16
(2)	学習データを収集・利活用するための制度設計	17
(3)	学習データを収集・利活用するための情報環境の整備	18
(4)	学習データを収集・利活用するための人材の育成	19
4	おわりに	20
	<参考文献>	21
	<参考資料1> 審議経過	24
	<参考資料2> LAの研究事例	24
	<参考資料3> 質問紙調査の質問項目の例	26
	<参考資料4> 大規模調査によって収集されている学習データの事例	26
	<参考資料5> 教育データ利活用ガイドラインの宣言と原則[6]	28
	<参考資料6> 次世代医療基盤法について	28
	<参考資料7> 学習データの連結と個人情報保護	29
	<参考資料8> 安全な学習データの管理体制	29
	<参考資料9> 学習データの利活用に関する倫理審査委員会の例	31
	<参考資料10> 初等中等教育機関における学習者用コンピュータの整備状況[16]	32
	<参考資料11> 初等中等教育機関における学校ネットワークと SINET の接続形態の例	32
	<参考資料12> 高等教育機関における BYOD の導入形態の状況[19]	33
	<参考資料13> 教員の ICT に関する能力の推移[16]	33
	<参考資料14> 提言の概要（実施時期は 2020～2030 年度を想定）	34

## 1 はじめに

### (1) 従来を超えたきめ細かな教育の可能性

現在、学校などの教育機関においては GIGA スクール構想の推進や新型コロナウイルス対策でのオンライン授業の実施など、「教育のデジタル化」が急速に行われている。この「教育のデジタル化」によって、各教育機関や民間企業には、学習支援システムや校務支援システム等を用いて様々な学習データが蓄積されているが、この利活用に関しては明確な指針もなく、これを教育改善のために活用することはあまりされていない。この学習データを用いて、各個人の学習プロセスをきめ細やかに推測し、どの問題で分からなくなったか、どの概念が理解できていないか、などの理解につまずいた箇所を個人ごとに発見し、個別に対応することにより学習者の学力向上を図れる可能性がある。また、教員が、あるいは学習者自身が、学習データの分析結果を基にリアルタイムに教育方法や学習方法を検討できる点は非常に重要である。このような新しい教育に対し、教員-学習者間、学習者同士、学習支援システムとの間のインタラクションを分析して知見を統合することで、新しい教育学的な理論を構築できる可能性がある。

米国や英国などの諸外国では、高等教育機関を中心に、LA の研究・実践が盛んに行われ、学力の向上や退学者の減少など、学習データの利活用の効果が出始めている。日本では、2016 年に九州大学が LA センターを設置し、大学全体を対象に学習データの収集と分析を行っているが、ほとんどの教育機関では実現されていない。GIGA スクール構想の実施や新型コロナウイルス対策でのオンライン授業への移行等による教育のデジタル化の推進、並びに AI・ビッグデータの利活用が急速に進みつつある現状では、教育改善のために国全体で学習データの収集と利活用をどのように行うべきかを議論することは急務である。また、このような取組は、学校等の教育機関だけでなく、学習塾などの産業界や自治体・行政などの産学官の垣根を超えて、社会全体で取り組む必要がある。なお、このような教育ビッグデータを活用したデータに基づく科学的教育取り組みの重要性は、日本学術会議「国立大学の教育研究改革と国の支援—学術振興の基盤形成の観点から—」（2017 年 6 月）の提言にも示唆されており、文部科学省「新時代の学びを支える先端技術活用推進方策（最終まとめ）」にも教育ビッグデータを効果的に活用していくための様々な取り組みが提案されている[1]。なお、文献[1]は、主に初等中等教育での学習データの収集と利活用を対象としているのに対し、本提言は、初等中等教育から高等教育までを含めた学習データを対象としている点で異なる。なお、本提言は文科省・経産省・総務省などの国の政策に向けたものである。また、本提言は一人一台の情報端末や学習支援システム等を用いて蓄積される学習データに限定したものであり、教育活動全般で利用される広範な教育データの取り扱いについては、慎重に議論を深め、継続して検討する必要がある。

## (2) エビデンスに基づく教育の実現のために

これまで、多くの教育政策に対して効果・課題がまとめられてきたが、それらは多様な状況にある学校・自治体の全てに適用できるものではなかった。その原因の1つが、政策立案時におけるエビデンスの利活用の欠如と考えられる。このエビデンスの蓄積のためには、国全体での学習データの収集が鍵となる。また、学習データの利活用によるエビデンスの利用は、国レベルにとどまらず、学習者や教員の個人レベル、学校等の教育機関レベルでも利用可能である。例えば、日本の教育現場が継承すべきベテラン教員の優れた授業ノウハウや、情報技術を駆使して可能となる新しい授業の方法に対する効果・課題をエビデンスとして集約・蓄積し、柔軟に利活用して、各教員や各学校・自治体がそれぞれの現状に適したものと模索できる環境を整えることは重要である。このようなエビデンスに基づく成果を、学校・自治体レベルから国の学習指導要領レベルまでのカリキュラム・マネジメントに役立てることは、次の社会を担うべき人材を多様な地域から育てることに寄与しうる。このような考え方に基づいた事業は既に文部科学省が模索を始めているが[2][3][4]、本提言は、このような国の政策が進むべき方向性を示すことを意図している。

## 2 現状と問題点

学習データの利活用のあり方について、以下の4つの視点からその現状と問題点を示す。

- ・学習データの種類と教育改善ための利用
- ・学習データを収集・利活用するための制度設計
- ・学習データを収集・利活用するための情報環境の整備
- ・学習データを収集・利活用するための人材の育成

### (1) 学習データの種類と教育改善のための利用

#### ① 学習データの分類

本提言では、学習データとは、教育のデジタル化によって、一人一台の情報端末を使って学習支援システムや校務支援システム等を用いて蓄積されるデジタル情報と定義する。これは、いくつかの学校を対象とした、これまでの実証実験によって、表1に示すように、授業支援系データと校務系データに大別される[42]。表1の1～3は、学習支援システム等を用いて、学習者の学習プロセスを細かく記録したデータである。4～9は、1～3のデータを分析するために必要となるため、国全体で学習データを共有するためには必要となる。10～12は、主に一部の研究で利用されてきたデータであり、共有する必要性は低いと考えられる。なお、この学習データは、教育・学習活動の一部を切り取ったものであり、教育の効果や達成度などを全ては

かることはできず、ある意味限定されたものであることに注意が必要である。また、実際の教育現場では、病歴や障がいへの配慮策やいじめ・虐待など、表1の内容に含まれる場合があるが、本提言では、国全体でのデータの共有は、このような情報を除き、1～9を基本項目として収集することを提起する。現在、学習者アンケートなど、これらの一部は、現在は紙も含めた様々な方法で集められているものもあるため、教育現場にはデジタル化のための作業負担を強くないよう、既に電子化されているところからデータの利活用を始めるべきであろう。

表1 学習データの種類（灰色は国全体で匿名化して共有する基本項目）[42]

区分	番号	データの種類	説明
授業・学習系データ	1	学習支援システム学習履歴	デジタル教材閲覧履歴、LMS等の利用履歴、デジタルノートの内容
	2	デジタルドリル学習履歴	デジタルドリルの回答や正答率等
	3	学習者アンケート結果	学習者に対するアンケート結果
校務系データ	4	学籍情報	学習者の学年等の基本情報
	5	出欠席情報	学習者の日々の出欠情報
	6	指導計画情報	授業ごとの指導計画やシラバス
	7	テスト結果	小テストや定期テスト等の結果
	8	成績評定情報	通知表や単位取得等の評定結果
	9	教員アンケート結果	教員に対するアンケート結果
	10	健康観察記録	学級担任等が朝に行う児童生徒の健康状態を確認した記録
	11	日常所見情報	児童生徒の日々の様子や気付いた点などを記録した情報
	12	保健室利用記録	児童生徒が保健室に来室した記録

## ② 学習データの利活用の現状

学習データの利活用を扱う研究分野として、LAがある。LAとは、情報技術を用いて、教員や学習者からどのような情報を獲得して、どのように分析・フィードバックすればどのように学習・教育が促進されるかを研究する分野である。LAの研究は、2011年以降、LAK(Learning Analytics and Knowledge)などの国際会議の始まりと共に、活発になった。その当初は、LMS(Learning Management System)やMOOCs(Massive Open Online Courses)等を用いたオンラインコースのデータを扱った研究が数多くみられ、対面型の授業を扱った研究は少なかった。近年は、対面型の授業のデータを扱った研究も増えており、欧米の大学では成績の向上や退学率の低下などの効果がではじめている（参

考資料 2 : LA の研究事例)。しかしながら、日本を含め海外でも、まだまだ研究者個人の授業でデータを収集した研究が数多くみられ、大学等の教育機関全体で学習データを収集して利活用している研究はそれほど多くない。理由としては、教育機関全体で LA を行うには学習データの収集・管理・利活用を行うためのポリシーとして標準的なものがないことがあげられる。今後、LA を国全体に広げるためには、教育機関や国全体のポリシーを提案していく必要がある。日本においては、2016 年 2 月に九州大学が LA センターを、日本を含めアジアで初めて設置した。九州大学では 2012 年度から、全学一年生(約 2,700 名)を対象に PC 必携化(BYOD: Bring Your Own Devices)を行い、2014 年 4 月からは、LA のためのシステムを一年生全員(約 2,700 名)に導入した。同年 10 月からは、全学の学生・教職員約 27,000 人がこのシステムを利用している。ここでは、主にデジタル教材の閲覧ログ、出席、レポートの提出、小テストの成績、最終成績等を分析・可視化して、教員や学習者にフィードバックし、教育・学習改善を行っている。具体的には、予習・復習の学習行動の分析、最終成績の予測、リアルタイムな学習活動分析などが行われている。

表 2 教育改善のための学習データの利活用の分類

	対象	誰のため	目的の例
1	個人	学習者	<ul style="list-style-type: none"> <li>・個人に適した教材や問題の推薦による学習効果の向上</li> <li>・過去や現在の学習データを用いた、理解度の予測などによる、個人の学習状況の把握</li> </ul>
		教員	<ul style="list-style-type: none"> <li>・クラス全体や個々の学習者のつまづき箇所の発見などによる教材や授業設計の改善</li> <li>・自動採点など、学習データの利用による教員の負荷の軽減</li> </ul>
		保護者	自分の子供の学習状況、学習意欲などの把握
2	教育機関	機関の管理者	<ul style="list-style-type: none"> <li>・学習データに基づくカリキュラムの最適化</li> <li>・教員や学習者の最適な配置</li> </ul>
3	国や地域	政策立案者	エビデンスに基づく教育政策の立案と評価
		研究者	大規模な縦断的・横断的データを用いた学習者の成長過程の研究
		市民	教育に関する諸問題を、データを用いて社会全体で共有・議論

さらに、2019 年からは、京都市内の中学校・高等学校において、学校全体を対象とした LA の実証実験が行われている[5]。さらに、教育機関や国全体で

学習データを大量に収集することによって、人工知能(AI)技術を用いた様々な教育・学習活動の支援が考えられる。例えば、レポートや答案の自動採点や授業の自動設計などの教員の支援、最適なカリキュラムの提案やクラスの自動編成などがある。このように、学習データは、1の学習者や教員による利用だけでなく、2教育機関や研究者、3社会全体での利用も考えられる(表2)。

1は狭義のLA、2は教学IR(Institutional Research)、3は根拠に基づく政策立案(Evidence-Based Policy Making: EBPM)、オープンデータ、研究データ管理などが関連する。本提言では、LAを広義で捉え、学習データを1～3の全ての目的で利用するものとする。

### ③ 収集されている学習データの現状と問題点

既に、紙のテストや質問紙などを通じて、様々な公的・民間機関による学習データの収集が行われている(参考資料4:収集が行われている学習データの概要)。しかし、統計的に集計された集計データは公開されているが、研究目的であっても個票データを利用することができないことが多い。この問題は、デジタル教科書やLMS等のシステムを用いて電子的に収集された学習データについても同様である。現状では、システムに収集されたデータを研究利用するための手続きや制度などの整備が不十分である。従って、将来的には個人情報保護に配慮して適切に加工した上で、個票の学習データを収集し、利活用できるような制度設計を検討する必要がある。

### ④ 学習データの収集対象となる機関

現在、学習データは、小学校・中学校・高等学校・大学などの公的教育機関、及び、学習塾などの私教育機関あるいは教育サービスを提供する民間企業などに分散して管理されているが、まずは、学習データを収集する対象として、小学校・中学校・高等学校・大学・大学院などの公的教育機関から始め、自らの教育改善に利活用できるようにするべきであろう。そして小学校・中学校・高等学校・大学までの学習データを対象に、教育機関ごと、次いで当該教育機関の所属する法人格(地方自治体、大学法人、学校法人など)ごとに個人情報保護に配慮して適切に処理して国が認定する公的機関が収集し、利活用すべきであろう。そうしなければ、国レベルで、データに基づく教育政策の検証や立案が行なえず、各学校や学習者個人レベルでは、全国の平均との比較や分析が行えなくなる。また、研究者は、より大規模なデータを利用して、教育や学習効果の向上のための研究を実施できなくなる。なお、このデータの共有は、問題がないことを確認しつつ、慎重に議論を深め徐々に行われるべきである。また、データの収集の対象は、生涯にわたって学習データを収集することはせず、比較的データの収集がしやすい、小学校から大学までの学校教育内のデータのみを対象として始め、教育政策の立案や改善・評価のために学習データを収集・利活用して、その効果を検証することから始めた方が良い

と思われる。また、最初は何らかの公的な機関がデータをしっかりと集めて管理し、利活用において問題がないことを確認した後、民間企業がデータを集めたり利用したりすることは国全体で慎重に議論しながら検討すべきであろう。従って、最初は、民間企業が国全体のデータを集めることをせず、公的機関が収集すべきであろう。

表3 学習データの利活用の段階の分類

レベル	説明	学習データの活用方法と情報保護
0	学校や地方自治体、大学などの法人内で学習データを共有・活用ができていない。	・法人内で学習データが電子化されていない、もしくは、分散管理されている。
1	法人内では学習データを共有し、データの分析や学習者への介入等の利活用できているが、教育機関の所属する法人外にはデータを共有していない。	・同一法人内で学習データの活用は可能 ・学習データは個人情報であり、適切に管理されることが必要
2	法人内で学習データを共有・活用し、個人情報保護に配慮して適切に処理した後、学校や地方公共団体、大学などの外部の法人ともデータを共有しているが、意味やフォーマットの統一できていない。	・法人内外での学習データの利活用は可能 ・個人情報保護に配慮して適切に処理して学習データの利活用に対するルールが必要 ・法人内外でデータのフォーマットが統一されていないため計算機処理が困難
3	法人内で学習データを共有・利活用できている。また、個人情報保護に配慮して適切に処理した後、意味やフォーマットを統一した形で法人外にデータを提供し、法人内外で統一したデータを利用できている。	・法人内外での学習データの利活用は可能 ・個人情報保護に配慮した学習データの利活用に対するルールが必要 ・法人内外でデータのフォーマットや意味が統一されているため学習データを計算機で処理することが容易

#### ⑤ 学習データの保存期間

学習データは、成績などの一部のデータは、教育機関に長期間保存されるが、それ以外の小テストの結果や質問紙データ、授業・学習系データなどは、現在、学校を卒業後、一定期間で破棄されてしまう。しかしながら、表2の

目的のためには、過去から現在までの大量の学習データを蓄積しておき、利活用することは重要である。例えば、大規模な学習データは、個人レベルでは、問題や教材の推薦、学習者の理解度の予測などの精度を向上させたり、教育機関レベルでは、カリキュラムの最適化などに役立つ。また、国全体では、大規模な学習データから様々なエビデンスを抽出してエビデンスに基づく教育を実現し、心理学的・教育学的・社会学的等の多様な観点からデータを分析して学術的な知見を蓄積する上で、極めて貴重なデータとなるであろう。従って、教育機関レベルでは学習データを長期間保存し、国レベルでは、個人情報保護に配慮して学習データを加工して収集して、長期間、利活用できる状態にしておくことは重要である。なお、学習データは個人情報であるため、国全体で学習データを収集・利活用するためには、そのための制度設計や実装方法をしっかりと検討して実施する必要がある。

## (2) 学習データを収集・利活用するための制度設計

学習データの多くは、教育を受けた学習者個人のものであるが、教育を提供した教員や学校のものである。また国全体で教育を良くするためには、そのデータを国全体で共有する、ということも必要である。ここでは、そのために必要となる制度について論じる。

### ① 学習データ利活用のレベル

表3に示すように、学習データの共有と利活用の範囲に応じて、学習データの利活用のレベルを4段階に分けることができる。現状ではレベル0、つまり、教育機関内でも学習データの共有・活用ができていない場合がほとんどである。先行して学習データの利活用をしている九州大学でもレベル1の段階である。レベル1の状態では、学校の規模に応じて収集されるデータ量が大きく異なり、結果的には、学習データの利活用による支援にも差が出てしまいかねない。また、レベル2の段階では、法人間をまたいで学習データの共有はできていても、フォーマットや意味が統一されていないため、計算機上のプログラムを用いた分析が困難となる。従って、計算機で処理することを考慮し、かつ、法人間をまたいで学習データを収集すること考えた場合、フォーマットや意味が揃った形でデータを収集する、レベル3を実現すべきである。そうしなければ、仮に法人間をまたいで学習データを収集できたとしても、そのデータを計算機で処理することが困難となる。一方で、学校等の公教育機関に対して、LMS等の教育サービスを提供する民間企業は、その内部で学習データ利活用を行っているレベル1の状態であるが、多くの場合レベル2には至らず、外部へデータは提供していない。むしろ自らの学習データを内部に囲い込む場合が多い。従って、社会全体で教育改善を行うためには、レベル3を実現する必要がある、公的教育機関は、それが利用する教

育サービスを提供している民間企業と、指定したデータフォーマットや意味で学習データを提供してもらうよう、最初に契約を結んでおく必要がある。

## ② 学習データ利活用ポリシーの現状

学習データの利活用については、同一機関内の利用であっても、個人情報保護などの観点から未だに躊躇しているのが現状である。そこで、大学 ICT 推進協議会 (AXIES) では、大学等の高等教育において学習データの利活用を促進するための基本原則やガイドラインなどについて議論している [6] (参考資料 5 : 教育データ利活用ガイドラインの宣言と原則)。将来的には、初等・中等・高等教育機関において学習データの利活用を促進するためには、このような「教育データの利活用ポリシー」を国レベルで定めていく必要があるだろう。また、ほとんどの大学では情報セキュリティポリシーが策定されている一方、学校を対象とした情報セキュリティポリシーを策定している教育委員会は 64.1% (策定中含む、平成 29 年 2 月) に留まっている [7]。加えて、セキュリティポリシーは自治体ごとに極めて多様であり [8]、これも整理していく必要がある。さらに、教員への情報モラル教育など、このようなポリシーを実践するための教育を同時に行っていく必要がある。

このポリシーには、学習データの利活用における注意点も盛り込む必要があるだろう。とりわけ、「エビデンスに基づく教育」に対する国内外の批判的議論 (例えば, [9][10]) を十分踏まえた上で、利活用することが重要である。特に、原理的な次元においては、教育が常に異なる人や環境を対象にして行われていることに伴う再現不可能性 (非再現性) などの問題もあり、学習データからエビデンスを抽出し、教育体系を生成するようなことは、困難であろう。しかしながら、近年は、個々の教育実践によって蓄積された教育ビッグデータを科学的に分析してエビデンスを導出する取り組みも行われはじめており [41]、学術的には、このような研究のアプローチの歩みを止めずにチャレンジし続けることは、極めて重要であると思われる。例えば、学習データの利用における注意点には、以下のようなものが考えられる。

- ・従来の教育活動の業務以外に学習データを利用する際には、表 2 のような教育改善や授業改善の目的外には学習データを利用しない。なお、教育改善を示す評価指標は成績や学習者の発言回数・積極性・創造性など、たくさんあるが、実践を通じてそれぞれの指標を検討しながら、範囲を広げていく必要があるだろう。
- ・教育に係る選択は本人が実施するものであり、学習データを利用した推薦や提案が、本人にとって決めつけや押し付けにならないようにする。
- ・教員や LA 専門員は、教務において学習データの利用によって知りえたことに関して守秘義務を負う。これ以外のデータの利用者、例えば研究者等は、個人を特定するような分析や地域や学校間の無用な比較を

しない。

- ・教員は、過去の学習データに基づくエビデンスを過度に過信してしまい、前例主義になってはいけない。これまでの自分の経験や学習者の状況と照らし合わせて、上手にエビデンスを利用する必要がある。
- ・エビデンスを、政策立案の根拠、政策や実践の評価などに利用するときは、データが収集された文脈や方法を吟味し、その信頼性や妥当性をチェックし、限界を踏まえることが重要である。例えば、限られた指標だけで政策評価を行うのではなく、多面的なデータ（例：大規模調査の量的データと学校や学級における実践に依拠した質的データ）を用いて行う必要がある。

なお、上記のようなデータ利活用の制限は、本提言を政策として実現する際に検討する必要がある。その際には、国が倫理審査委員会のような第三者機関を設けて、定期的に見直し、個人の人権と個人と尊厳を最大の基軸として学習データの利活用の促進や悪用を防ぐなど、制度を改良する体制を構築することも重要である（参考資料9：倫理審査委員会の例）。また、この制度は、Web ページでパブリックコメントを集めたり、各地で説明会やシンポジウムを開催して、マルチステークホルダープロセスにより、国全体で議論しながら、進めていく必要がある。

### ③ 学習データの管理体制

現状の個人情報保護法のもとで、国または公的機関が学習データを収集するには、学校ごと、または学校が所属する法人格（地方自治体、大学法人、学校法人など）の中で、個人情報保護に配慮して適切に処理された後のデータを収集すべきである。この場合、学習者本人の同意は不要となる。また、法人をまたいでデータを収集する仕組みは、医療データを共有する仕組み（参考資料6：次世代医療基盤法）などを参考に検討が必要であろう。ここで、もし、学習者あるいは保護者から同意をとることにした場合、同意しなかった学習者は、学習データを用いた問題の推薦や介入などの支援をうけることができず、教育現場では公平性に欠けることとなる。従って、教育現場では、これまで成績データなどの管理で行ってきたように、本人同意は不要としてデータを収集した方が良い。一方、学習データを研究利用する場合は、各研究機関の研究倫理委員会の規則に則って、これまでと同様に、収集するデータの項目や利用目的、管理方法などを事前に説明をしっかりと行い、本人あるいは法的代理人から自発的な同意をえること、また、同意をいつでも取り下げることができるようにすること、などが必要である。これら同意の取得方法や取り扱いについては、社会実装をする際に、上述の倫理審査委員会において、しっかりと議論して進めていくべきである。

学習データを異なる機関から収集し解析する上で必ず重要となるのは、デ

ータを正確かつ最新の状態で管理する、データ完全性である。例えば、第三者が解析するためだけであれば、個人情報を提供させず、必要なときに各教育機関がもつ分散データを連結し、完全性を保証する方法[9]やブロックチェーンを用いた方法などを検討する必要がある（参考資料7：学習データの連結と個人情報保護）。

表4 学習データの共有に伴って必要となる匿名化の段階の例

	学習データの保存	学習データの提示方法の例
同一法人内	個人にフィードバックするため非匿名のまま個人情報を保存	学習者本人・その保護者・授業の担当教員には個人識別可能な情報を表示
		学習者本人以外の学生、他の保護者には個人識別可能な情報は非表示
同一法人外で研究者・行政などが利用	個人情報保護に配慮して適切に処理された情報を保存	不可逆的に匿名化された情報等を表示

情報セキュリティ等の技術により保護可能な個人情報や、運用で守る必要がある内容を明確にする必要がある。例えば、表4のような個人情報の保存方法や提示方法の制度化も合わせて検討されるべきであろう。特に、研究者・行政機関などの法人外で学習データを利用する際は、不可逆的な匿名化の技術を適用してデータを個人情報保護に配慮して適切に処理した上で、「個人を特定するような分析をしてはいけない」のようなデータ利活用時の禁止事項（3.(2).②節参照）の遵守も合わせて実施すべきであろう。一方、学習データの管理体制としては、次の二点を両立させることが重要であろう。

- ・学習者（保護者を含む）・教員・研究者・政策立案者・市民らが、それぞれの立場で教育をよりよくしていくために、学習データを比較的、容易に利用できること。

- ・学習データの管理体制や制度が国民の理解と信頼を得ること。

そのための適切なシステム・アーキテクチャ、学習データの運用管理体制、ポリシー等を設計するにあたり、収集・フィードバックの段階においても、その後の長期保存の段階においても、安全性を十分に考慮する必要がある（参考資料8：安全な学習データの管理体制）。前者についてはクラウド環境を前提に、医療等の領域で既に導入実績のあるシステムの仕組みなどについて検討を行っていく必要がある。

#### ④ 学習データのフォーマットと意味の統一

将来的に学習データは、タブレットやパソコン等を用いて、LMS等の学習支援システム（仮想学習環境）の利用によって集められる場合が多くなると

考えられる。しかし、学習支援システムには数多くの種類があり、それらのログデータは独自のフォーマットで記録されている。これらの学習支援システムを用いて、異なる教育機関で集められた学習データを共有し、計算機で処理するためには、共通のフォーマットが必要である。この学習履歴をデータ表現するための世界標準の規格としては、xAPI と Caliper の 2 つがある。xAPI は 2013 年 4 月に米国 ADL(Advanced Distributed Learning Initiative)によって公開された。その後、IMS Global によって Caliper も提案された。xAPI や Caliper 形式の学習履歴データは、多くの場合 LRS (Learning Record Store)に蓄積される。これらの標準的なデータ形式を用いることにより、異なる機関間での学習データの交換が容易に行える。ただし、データの意味(言葉や数字の解釈)までは規定されていないため、海外では、エビデンスに基づく政策立案の実現を動機として、米国の Common Education Data Standards[11]や英国の Common Basic Data Set [12]などのデータの標準化が存在している。このようにフォーマットだけでなく意味も統一して蓄積することが、学習データの利活用には非常に重要となる。

#### ⑤ 教科書や教材等の学習データの著作権処理の現状

デジタル教科書や問題集等のデジタル教材の利用履歴は、教員や学習者がどのような教育・学習活動を行っているかを把握するために重要なデータとなる。しかし、最近の文学作品やある機関が所有する重要な文化財の画像など、著作権処理の必要なコンテンツを授業外あるいはオンラインで利用する場合、その利用が困難な場合がある。確かに、教育現場における教科書や教材の柔軟な利用は、平成 31 年の著作権法の改正施行により教育的に配慮が進んだが、著作権法を遵守するための手続きには大変時間がかかる場合がある。特に、紙の教科書や教材の著作権処理は済んでいるが、それらの電子版の著作権処理は手つかずの場合が多く、教材会社にとって大きな負担となっている。また、デジタル教科書や問題集等を利用する学習者や教員以外の人物が、それを利用して分析する場合、出版社に許諾を得る必要があるという問題もある。このため、デジタル教科書やデジタル教材を、授業内・外、オンライン・オフラインを問わず、教育改善のために十分利用でき、内容を分析できるように、教科書制度や著作権法を見直すことも検討が必要であろう。また、それが困難な場合においても、ワンストップの補償金や裁定制度による対応により、著作権への対応を簡便化する方法を検討するべきであろう。

さらに、教科書や教材だけでなく、学習者が提出したレポートや作文など文章も収集して分析する場合、その著作権の取り扱いも考えておく必要がある。例えば、このような学習データを収集する場合は、開始時にあらかじめ同意を得ておくなどの対応が必要であろう。このように、

学習データの利活用においては、著作権法などを遵守して、十分な対応策を考えておくことが重要である。

表5 学校の情報環境整備のレベル

レベル	説明	学習データの収集と分析
0	情報端末の環境が整備されていない。	主に紙面での試験結果やアンケートの集計とその分析
1	学校内において教員にのみ情報端末を提供しているが、学習者には情報端末を提供していない。 (電子黒板やプロジェクター等を利用するものの教員のみが情報端末を利用)	教員の情報端末を用いて授業を行い、学習者からは紙で問題への回答などの学習データを収集して分析可能(学習者からは細かな学習プロセスのデータは取得不可能)
2	学校内のコンピュータ室でのみ、学習者が一人一台の情報端末とネットワークを利用できる。	主に授業内の学習プロセスデータを用いた分析が可能
3	教室の内外で学習者が一人一台の情報端末をネットワークに接続し、学習に利用している。(BYODや情報端末貸与等の形式がある)	授業内外の学習プロセスデータを用いた分析が可能

### (3) 学習データを収集・利活用するための情報環境の整備

教室の内外の学習活動のデータをリアルタイムに収集するためには、情報端末やインターネット接続などの情報環境の整備が必要不可欠である。学校の情報環境の分類を表5に示す。無論、レベル3を目指すのが理想であるが、以下の節で述べるように、現状ではまだまだ理想とは程遠い環境であり、情報環境の早急な整備が望まれる。

#### ① 初等中等教育における情報環境整備

表1の学習データの中でも授業・学習系データは、学習者が一人一台の情報端末を用いることにより効率的に収集できると考えられるが、初等中等教育の教育現場では、一人一台の情報端末などの情報環境整備は遅々として進んでいない。例えば、5年ごとに定められる教育振興基本計画において、第二期[13]における教育用コンピュータ一台当たりの児童生徒数 3.6 人目標の達成のため「教育のIT化に向けた環境整備4か年計画(平成26~29年度)」[14]を進めたものの、様々な要因[15]により、平成30年度末において

十分に達成されていない（参考資料 10）[16]。また、学校のネットワーク接続に関しては、30Mbps 以上のインターネット接続率は 91.5%であるが、普通教室の無線 LAN 整備率は 41.0%[16]とまだまだ低いのが現状である。

このような現状において、文部科学省は新たに「教育の ICT 化に向けた環境整備 5 か年計画（2018～2022 年度）」[17]を策定し、三人に一台の学習者用コンピュータの整備を目指している。学習データ利活用の前提となる一人一台環境の実現はまだ遠いと言わざるをえない。このため、OECD 生徒の学習到達度調査（PISA2018）「ICT 活用調査」において学校の授業におけるデジタル機器の利用時間は OECD 加盟国中最下位である。また、他の OECD 加盟国と同じように学校外で多様な用途にデジタル機器が利用されている一方、コンピュータを使って宿題をする頻度は OECD 加盟国中最下位である[18]。

現在の学校現場では外部とのネットワーク接続に制限があるため、家庭から学校内の学習支援システムにアクセスしたり、逆に学校の中から外部の学習支援システムにアクセスしたりして学習データを蓄積するためには大きな障害となる場合が多い。高等教育の現場で広く用いられている SINET（Science Information NETwork）は、将来的に初等中等教育に開放される予定であり[1]、これを用いて学習データを収集することを検討する必要がある。初等中等機関の SINET への接続形態は、例えば、既にその形態をもっている（参考資料 11）のような広島市の地域ネットワークの接続をモデルとすることができる。全国の教育委員会の学校ネットワークはその単位で集約して上位のネットワーク（参考資料 11 の場合は地域の ISP ネットワーク）に接続しているので、上位ネットワークを SINET に向けることで接続が可能である。このような接続形態により、初等中等教育と高等教育機関が SINET のみを介してセキュアに接続され、通信効率や通信品質が期待できる。また、SINET により、学習データの収集・利活用が行われる初等中等教育と、それらを学術的に支える高等教育が接続され、学習データ利活用の PDCA サイクルの効率化に繋がる。また学習データが利用できれば、授業の再開などの災害からの復旧時に役立つと考えられる。

## ② 高等教育における情報環境の整備

大学、短期大学、高等専門学校等の高等教育機関においては、教育機関が情報端末を貸し出したり、学習者が自分の端末を持参したりする BYOD の形態で一人一台の情報端末の環境を実現している場合が多いと考えられる。2016 年度の AXIES の調査（回答率 59.2%=713/1204）によると、56.7%が学内の一部で BYOD を導入しており、43.3%が全く導入していないとの回答であった[19]。高等教育機関は、初等中等教育に比べると進んでいるが、まだまだ環境の整備が進んでいないことが分かる（参考資料 13）。高

等教育機関のネットワーク環境については、99.7%の国公立大学が対外インターネット接続を整備しており、国立大学の98.8%、公立大学の83.7%、私立大学の54.1%が、国立情報学研究所が運用するSINETに加入してインターネット接続をしている[20]。学習データの蓄積には、LMSなどの学習支援システムを利用される場合が多い。日本においては、大学ICT推進協議会(AXIES)が主体となって高等教育機関でのICT利活用教育の状況を調査したところ、2015年度の時点で約65.3%の国内大学でLMSが普及しているとのことである[21]から、LAの基盤となる仕組みが多く大学の既に整いつつあるといえるだろう。一方で同じ調査によれば、LMSを利用する科目数は20%以下とのことであり、今後の進展が望まれる。なお、米国の大学ではLMSの導入率100%、利用率62%と報告されており[22]、日本とは大きく差がある。

### ③ 情報環境の整備に伴う情報モラル教育や健康への配慮

情報端末を各自でもつことにより、学習者への情報モラル教育や情報セキュリティ教育が必要となる。一部の自治体では、スマートフォンなどの情報端末を学校に持ち込んで教育活動に利用可能とする方針を示しており、同時に情報モラル教育の取り組みなどの重要性も強調しているが、学校で行うには十分なリソースがなく、また、教員自身がデジタル端末に慣れていない場合が多く、それほど容易ではないという問題点もある。情報環境の整備にあたっては、学習者の健康面への配慮をするべきであり[23][24]、安心して教育活動に使用でき、教育現場から支持されるような情報端末が満たすべき要件[25]について検討する必要がある。これらについては、(一社)教科書協会や(公財)教科書研究センターなどが紙の教科書に対して行ってきた教科書の体様に関する調査を拡大して継続し、情報端末や環境に対する認定制度なども考慮する必要がある。

## (4) 学習データを収集・利活用するための人材の育成

### ① 教員養成課程における教育のデジタル化への対応の遅れ

現在は、大学の教員養成課程において、情報セキュリティや個人情報の扱いに関する最低限の知識も含め、情報学に関する教育は完全に不足している。日本には、教育学や教員養成のための課程を提供する学部は、国公立・私立を合わせて300以上存在するが、教育のデジタル化を直接扱う学科・講座は少ない[26]。例えば、学生数が5,000人近い東京学芸大学においては「教育AI研究プログラム」の設置などの動きがみられるが、教員養成を担う学部には学習データ利活用に近い学科・講座は存在しないことが多い。このため、学習データの解析を踏まえて教育の改善方法などを提案できる、教員養成課程を修了した人材は不足している。また、教育現場に寄り添いながら、教員

や学習者個人の長所・短所の発見と適切な改善の提案をデータに基づいて行う研究や、学習データを幅広く分析して教育制度の問題点と改善策を提案するような研究は、情報学・教育学・心理学・認知科学等にまたがった学際研究となるが、そのような「学習データサイエンティスト」と言うべき人材を育成する準備が、日本の高等教育機関においてほとんど行われていない。その一方で海外には、米国コロンビア大学のように LA の基礎を学ぶ修士課程のプログラムをもつ大学や、16,000 人以上の学生が在籍する台湾国立師範大学のように、学習データ利活用に近い分野の学科を複数備え、それぞれが教育現場との連携を盛んに行っているような大学が存在する。今後はこのような人材を育成するための教育制度の改革も検討する必要がある。上記のような事情も影響して、情報端末機器を活用していくための教育現場の準備は十分でない。[16]によると、教員用の校務用コンピュータは日本全体で平均一人一台を平成 23 年度に達成したが、統合的な校務支援システムの導入は全国平均で 52.5%に留まっており、都道府県によって 100%から 9.6%まで格差が激しい(2019 年 3 月現在)。また、教員の ICT 活用指導力は年々向上してはいるものの、どの項目も 8 割程度で頭打ちになっている(参考資料 1 3)。[15]によると教育現場に ICT 機器の整備が進まない現状と問題点として「推進計画が立てられていない」「推進体制ができていない」「調達のための知識が不足」など、教育のデジタル化について十分な理解と経験をもつ人材の不足が含まれている。日本においては教員の勤務時間が世界的にみても非常に長く[27]、このような人材を現在の教育現場から育てる対策だけでは不十分であることは明らかであり、教員養成課程レベルの対策が早急に必要である。既に現職で働いている教員については、学習データを利活用した授業設計について、新たに学ぶためのシステムを構築する必要がある。[28]のように教員の ICT 活用指導力の向上を意図した解説書や教材を利用し、さらに学習データ利活用に対応していくことで、教育現場の教員が自らのクラスにおいて、学習データを利活用した指導を支えていく必要がある。さらに、情報技術の有無にかかわらず重要となりうる、長年培われてきたシニア教員のもつ経験を若手教員が継承するような観点からのシステムづくりも必要である。

## ② 学習データを専門的に扱う人材の不足と現状

LA を教育現場で担うべき人材は、現在であれば ICT 支援員と考えられるが、[29]に示すように、ICT 支援員の数は不足している。四校に一人の配置の目標[17]を達成するには現在の 34,000 校に対して 8,500 人が必要となるが、現在は 2,000 人しかいない。学校における他の専門スタッフであるスクールカウンセラー(7,344 人)、ALT(Assistant Language Teacher)(15,432 人)、学校司書(50%以上)の状況に対し、現状も目標も少ないと言わざるを得ない。その上で、ICT 支援員に求められる機能や業務は非常に多岐にわた

り、年々増加している。近年では、ICT 機器・ソフトウェアを活用した効果的な授業づくり・教材作成の提案・助言や、利用者の情報モラルに関する機能・スキルのような、その1つだけでも専門性を必要とされるような機能・業務が含まれる。教育現場において、学習データの収集や分析によって教育改善をサポートするための人材としては、現実的には現在の ICT 支援員が担当することが考えられるが、学習データを取り扱うため、ICT 支援員の特別な職種として資格制度をとり入れる必要があるだろう。例えば、ICT 支援員の資格に加えて、民間企業や大学等におけるデータ分析の経験や、LA 研究の実証実験のような経験を重視した、LA 専門員を導入することが考えられよう。

### ③ 学習データを用いた研究を行う研究者の育成の遅れ

教育改善のためには、学習データの利活用によって現場の問題点を発見し、改善策を提案し、効果的な学習方法や教育方法を発見する研究者の養成が必要である。これには教育や心理学・認知科学の素養をもち、データ科学や教育工学の技術をもつ学習データサイエンティストの育成が急務である。このため、学習データ科学を教えるための大学院の専攻の創生や既存のカリキュラムの見直しが望まれる。また、学習データ利活用に関する研究を行う研究者のキャリアアップの道筋を整えると共に、学習データ利活用が発展するための人材エコシステムの構築について、議論する必要がある。

## 3 提言について

以上の議論を踏まえ、学習データの利活用のあり方について、以下の提言を行う。(参考資料 1 4 参照)

### (1) 学習データの種類と教育改善を目的とした利用

#### ① 学習データの種類

本提言では、学習データとは、教育のデジタル化によって、一人一台の情報端末を使って学習支援システムや校務支援システム等を用いて蓄積されるデジタル情報と定義する。これらは、授業・学習系データと校務系データに分類できる(表 1)。なお学習データは、教育・学習活動の一部を切り取ったものであり、教育の達成度評価などに用いる場合には注意が必要である。

#### ② 学習データの利用目的

学習データ利活用の目的としては、教育改善のために、教員や学習者を含めた個々の教育・学習活動を支援するため、教育機関内で分析してカリキュラム等の最適化をはかるため、国全体で学習データを収集して政策立案者・研究者・市民などが利用するため、に分類できる(表 2)。

#### ③ 学習データ収集対象となる機関

本提言では小学校・中学校・高等学校・大学などの公的教育機関に関わる、現在入手可能な学習データの収集を提起する。そして、各公的教育機関では、

まず、同一機関内で学習データを収集し、自らの教育の改善に利活用できるようにする必要がある。そして、各公的教育機関の内部で学習データを個人情報保護に配慮した適切な加工をした後で、国や地方自治体と共有すべきであろう。この提言では、国全体で共有すべき学習データの基本項目としては、表1(1)～(9)を提起する。学習データの管理は、各公的教育機関が行い、学習データを収集する公的教育機関が匿名化のための処理プログラムを提供して、そのための作業負担を軽減する等の方策が必要であろう。

#### ④ 学習データの保存期間

学習データは学校教育の終了とともに破棄されるべきではなく、必要なデータを、個人情報保護に配慮して適切に処理された後で、国あるいは公的教育機関が収集して所定の期間保存し、研究者や行政機関が利用してエビデンスに基づく教育を実現すべきであろう。

### (2) 学習データを収集・利活用するための制度設計

#### ① 社会全体での学習データの利活用

学習データの利活用の実現は段階的に進められるとしても、基本的にはレベル3(表3)の、国全体でフォーマットと意味を統一した状態で学習データの共有と利活用を目指して制度を整備すべきである。そのための制度設計を国が行う場合に注意すべきこととしては、(i)教育機関が民間企業等のシステムや試験等を利用する場合、企業が学習データを囲い込む恐れがあるため、学習者の学習データを学校に提供するよう、契約時に盛り込むこと、(ii)国全体で学習データを収集する際には、「データを個人情報保護に配慮して適切に処理して収集する」、「教育に係る選択は本人が実施するものであり、学習データを利用した推薦や提案が本人にとって決めつけや押し付けにならないようにする」などである。

#### ② 学習データの利活用ポリシーの策定

大学ICT推進協議会(AXIES)での議論[6]を参考にしながら、初等中等教育においても、学習者、教職員、保護者が安心して活用できるように、基本原則・ガイドラインをまとめる必要がある(参考資料5:教育データ利活用ガイドラインの宣言と原則)。特に、「従来の教育活動の業務以外に学習データを利用する際には、教育改善や授業改善の目的外には学習データを利用しない」などの学習データの利活用における注意点もポリシーに盛り込む必要がある。さらに、国が倫理審査委員会を設けて、データの悪用や行き過ぎた利用をしないように、定期的に制度を見直すような体制を作る必要がある。

#### ③ 学習データの管理体制

現状の個人情報保護法のもとで地方自治体等の法人を越えて学習データを収集するには、各法人格が個人情報保護に配慮して適切に処理した後のデータを国あるいは公的教育機関が収集すべきであろう。この場合、本人同意は

不要となる。また、学習データの管理体制としては、次の二点を両立させることが重要である。

- ・学習者（保護者を含む）・教員・研究者・政策立案者・市民らが、それぞれの立場で教育をよりよくしていくために、学習データを比較的、容易に利用できること。

- ・学習データの管理体制や制度が国民の理解と信頼を得られること。

そのための適切なシステム・アーキテクチャ、学習データの運用管理体制、ポリシー等を設計するにあたり、収集・フィードバックの段階においても、その後の長期保存の段階においても、安全性を十分に考慮する必要がある。

#### ④ 学習データのフォーマットと意味の統一

学習データの書式は、xAPI などの世界標準を参考にするべきであろう。また、教材や問題の体系化のためには学習要素の ID 化や、授業計画・設計や教育評価などで使う用語の統一が必要である。具体的には、単元 ID や学習要素 ID のようなコードセットだけでなく、評価の観点、欠席の理由の一覧のようなデータについても整理し、フォーマットと意味を統一する必要がある。また、公教育機関はもとより、学習塾等の私教育でも可能な範囲で共通化すべきであろう。

#### ⑤ 学習データの著作権

デジタル教科書や問題集等のデジタル教材の利用履歴は、教員や学習者がどのような教育・学習活動を行っているかを把握するために重要なデータとなる。しかし、紙の教科書や教材の著作権処理は済んでいるが、それらの電子版の著作権処理は手つかずの場合が多く、教材会社にとって大きな負担となっている。このため、デジタル教科書やデジタル教材を、授業内・外、オンライン・オフラインを問わず、教育改善のために十分利用でき、分析できるように、教科書制度や著作権法を見直す必要があるであろう。

### (3) 学習データを収集・利活用するための情報環境の整備

#### ① 学習データを収集するための情報端末の整備

学習データを、授業内だけでなく授業外でも継続的に収集する必要性から、情報端末は一人につき一台整備する必要がある。また、教育を受ける人それぞれの個人の特性に合った新しい教育を行うにも、一人一台であることが必須である。このため、タブレットやスマートフォン等の既存の情報端末を学校でも利用する BYOD を検討し、学校でも家庭でも常時利用可能にする必要があるであろう。

#### ② インターネットの常時接続と情報基盤システムの整備

学習データをリアルタイムに蓄積するためには、一人一人がもつ情報端末を教室の内外でインターネットに常時接続できる環境を整備する必要がある。また、学習データはクラウド上に情報収集と利活用のための情報基盤シ

システムを構築することが想定されるが、それにはそれを利用する教育機関等からアクセスしやすい SINET の利用を検討すべきであろう。

### ③ 情報モラル教育の整備や健康への配慮

交通環境の整備に伴って交通安全教育が必要となったように、情報環境の整備に伴って、情報を安全に扱うための情報モラル教育が学習者や教員に対して必要である。例えば、国立情報学研究所が高等教育機関を中心に提供している情報セキュリティのコンテンツや文科省の「情報モラル教育推進事業」で作成されたコンテンツや活動[30]、京都府警察と連携しているネット安心アドバイザーの活動[31]のような、学校外の公的な機関との連携による改善を模索する必要がある。この情報環境は健康やメンタル面への影響も考慮して利用されることが望まれる。例えば、情報端末の長時間の利用による視力や体への影響、また睡眠不足や運動不足による精神面への影響などに配慮が必要であり、今後も、さらに継続して調査していく必要がある。

## (4) 学習データを収集・利活用するための人材の育成

### ① 一般教員の育成と養成

情報を収集し蓄積するのはあくまでも手段であって、そのデータを生かして少なくなりつつある子供をしっかりと教育するためには、教員が最低限、データを活用した教育について理解しなければならない。このため、現職の教員が新たに学ぶためのシステムを構築する必要がある。また、大学の教員養成課程においても、学習データを利活用したカリキュラム・授業・評価の設計・実施の方法をしっかりと教えていく必要がある。

### ② LA 専門員の育成

教育現場において、学習データの収集や分析によって教育改善をサポートするための人材としては、現実的には現在の ICT 支援員が担当することが考えられるが、学習データを取り扱うため、ICT 支援員の特別な職種として資格制度をとり入れる必要があるだろう。例えば、ICT 支援員の資格に加えて、民間企業や大学等におけるデータ分析の経験や、LA 研究の実証実験のような経験を重視した、LA 専門員を導入することが考えられよう。

### ③ 学習データサイエンティストの育成と人材エコシステムの構築

教育改善のためには、学習データの利活用によって現場の問題点を発見し、改善策を提案し、効果的な学習方法や教育方法を発見する研究者の養成が必要である。これには教育学や心理学・認知科学などの素養をもち、データ科学や教育工学の技術をもつ学習データサイエンティストの育成が急務である。これには、学習データ科学を教えるための大学院の専攻の設置や既存のカリキュラムの見直しが望まれよう。

#### 4 おわりに

教育は国の礎であり、社会全体で教育を良くするためには、個々の学習者が何をどのように教えられ学びをしてきたかという、学習データを個人情報保護に配慮して適切に処理して、統一された意味やフォーマットで、社会で共有・利活用していくことは重要である。特に近年のように、学習者も多様化する時代（外国人の学習者や教室での学習困難な学習者、遠隔地あるいは海外から参加する学習者などにも対応が必要な状況）において、それぞれの学習者の特徴に合わせた学習支援を提供するために学習データの利活用はますます重要となるであろう。また、現在 GIGA スクール構想の実施や新型コロナウイルス対応でのオンライン授業への移行などで、学習データ収集・分析の必要性は極めて高く、そのための情報基盤システムを早急に整備し、エビデンスに基づく教育を実現する環境を構築しておくことは非常に重要である。

## <参考文献>

- [1] 文部科学省. 新時代の学びを支える先端技術活用推進方策 (最終まとめ). 2019年6月25日 [http://www.mext.go.jp/a\\_menu/other/1411332.htm](http://www.mext.go.jp/a_menu/other/1411332.htm)
- [2] 文部科学省. 人口減少社会における ICT の活用による教育の質の維持向上に係る実証事業. [http://www.mext.go.jp/a\\_menu/shotou/zyouhou/1364592.htm](http://www.mext.go.jp/a_menu/shotou/zyouhou/1364592.htm).
- [3] 文部科学省. 次世代学校支援モデル構築事業. [http://www.mext.go.jp/a\\_menu/shotou/zyouhou/detail/1387543.htm](http://www.mext.go.jp/a_menu/shotou/zyouhou/detail/1387543.htm).
- [4] 文部科学省. 情報活用能力を育成するためのカリキュラム・マネジメントの在り方と授業デザイン 情報活用能力を育成するためのカリキュラム・マネジメントの在り方と授業デザイン. 2018.
- [5] 緒方広明、ラーニングアナリティクスを用いた「エビデンスに基づく教育」の実現に向けて、教育システム情報学会誌, Vol. 36, No. 3, pp. 215-217, 2019.
- [6] 上田 浩, 緒方 広明, 山田 恒夫, 高等教育機関における教育・学習データの利活用に関する方針の検討, 情報処理学会研究報告電子化知的財産・社会基盤研究会 (EIP), Vol. 2018-EIP-81, No. 21, pp. 1-6, 2018.
- [7] 文部科学省. 教育情報セキュリティポリシーに関するガイドライン. 2017. [http://www.mext.go.jp/a\\_menu/shotou/zyouhou/detail/1397369.htm](http://www.mext.go.jp/a_menu/shotou/zyouhou/detail/1397369.htm).
- [8] 湯浅壘道: 個人情報保護法改正の課題—地方公共団体の個人情報保護の問題点を中心に—. 情報セキュリティ総合科学 第6号 2014.
- [9] ガート・ビースタ (藤井啓之・玉木博章訳) よい教育とはなにか—倫理・政治・民主主義—白澤社 2016
- [10] 杉田浩崇・熊井将太 (編)、「エビデンスに基づく教育」の関を探る、春風社 2019.
- [11] Common Education Data Standards, <https://ceds.ed.gov/>
- [12] Common Basic Data Set. <https://www.gov.uk/government/publications/common-basic-data-set-cbds-database>.
- [13] 文部科学省. 第2期教育振興基本計画. 2015, H25.6.14 閣議決定. [http://www.mext.go.jp/a\\_menu/keikaku/detail/1336379.htm](http://www.mext.go.jp/a_menu/keikaku/detail/1336379.htm).
- [14] 文部科学省. 教育の IT 化に向けた環境整備4か年計画. [http://www.mext.go.jp/component/a\\_menu/education/micro\\_detail/\\_\\_icsFiles/afieldfile/2018/08/03/1369638\\_1\\_1.pdf](http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2018/08/03/1369638_1_1.pdf).
- [15] 文部科学省. 地方自治体の教育の情報化推進事例—ICT活用教育アドバイザー派遣—. 2016. [http://www.mext.go.jp/a\\_menu/shotou/zyouhou/1370125.htm](http://www.mext.go.jp/a_menu/shotou/zyouhou/1370125.htm).
- [16] 文部科学省. 平成30年度学校における教育の情報化の実態等に関する調

- 査結果. 2019,  
[https://www.mext.go.jp/a\\_menu/shotou/zyouhou/detail/1420641.htm](https://www.mext.go.jp/a_menu/shotou/zyouhou/detail/1420641.htm).
- [17] 文部科学省. 教育の ICT 化に向けた環境整備 5 か年計画. 2018-2022 年  
度. [http://www.mext.go.jp/a\\_menu/shotou/zyouhou/detail/1402835.htm](http://www.mext.go.jp/a_menu/shotou/zyouhou/detail/1402835.htm).
- [18] 文部科学省. 国立教育政策研究所. OECD 生徒の学習到達度調査 2018 年調  
査 (PISA2018) のポイント. 2019.  
[https://www.nier.go.jp/kokusai/pisa/pdf/2018/01\\_point.pdf](https://www.nier.go.jp/kokusai/pisa/pdf/2018/01_point.pdf)
- [19] AXIES. BYOD を活用した教育改善に関する調査研究結果報告書, 2018.
- [20] 文部科学省. 平成 30 年度学術情報基盤実態調査, 2019.
- [21] 緒方広明、大学教育におけるラーニングアナリティクスの導入と研究、日  
本教育工学会論文誌、Vol.41, No.3, pp.221-231, 2018.
- [22] Campus Computing Survey, 2013.  
<https://www.campuscomputing.net/content/2013/10/17/the-2013-campus-computing-survey>
- [23] 文部科学省. 学習者用デジタル教科書の効果的な活用の在り方等に関する  
ガイドライン. 2018.  
[http://www.mext.go.jp/b\\_menu/shingi/chousa/shotou/139/houkoku/1412207.htm](http://www.mext.go.jp/b_menu/shingi/chousa/shotou/139/houkoku/1412207.htm).
- [24] 文部科学省. 児童生徒の健康に留意して ICT を活用するためのガイドブ  
ック. 2014, [http://www.mext.go.jp/component/a\\_menu/education/micro\\_detail/\\_\\_icsFiles/afieldfile/2018/08/14/1408183\\_5.pdf](http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2018/08/14/1408183_5.pdf).
- [25] 久富望. 紙の教科書から推測される教育現場に支持されるデジタル教科  
書の特徴. デジタル教科書研究. vol.1, pp.37-49, 2014.
- [26] 旺文社, 螢雪時代 4 月臨時増刊 全国大学 学部・学科案内号 大学選び  
の基本 学部・学科がわかる!!, 2019.
- [27] TALIS2013,  
[http://www.mext.go.jp/b\\_menu/toukei/data/0thers/1349189.htm](http://www.mext.go.jp/b_menu/toukei/data/0thers/1349189.htm)
- [28] 文部科学省. 新学習指導要領を見据えた小中高等学校教員の「ICT活用  
指導力向上」のための ICT 活向上研修実施モデル 解説書. 2018.
- [29] (一社) 日本教育情報化振興会. ICT 支援員の育成・確保のための調査研  
究 成果報告書. 平成 29 年度文部科学省委託事業. 2018.
- [30] 文部科学省. 情報モラル教育の充実.  
[http://www.mext.go.jp/a\\_menu/shotou/zyouhou/detail/1369617.htm](http://www.mext.go.jp/a_menu/shotou/zyouhou/detail/1369617.htm)
- [31] 京都府警察ネット安心アドバイザー.  
<http://www.pref.kyoto.jp/fukei/anzen/cyber/cyber13.html>
- [32] Arnold, K. E., & Pistilli, M. D., Course signals at Purdue: Using  
learning analytics to increase student success. In Proceedings of

- the 2nd international conference on learning analytics and knowledge, pp.267-270. 2012.
- [33] Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D., Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, 1(1), pp.6-47, 2014.
- [34] 独立行政法人統計センター. 国勢調査における匿名化マイクロデータの有用性と秘匿性の定量的な評価 製表技術参考資料 32 独立行政法人統計センター. 2016.  
<https://www.nstac.go.jp/services/pdf/sankousiryoku2805.pdf>
- [35] 国立教育政策研究所：教育情報共有ポータルサイト  
<https://www.contet.nier.go.jp/>
- [36] NHK 放送文化研究所  
[http://www.nhk.or.jp/bunken/about/research\\_2017.html](http://www.nhk.or.jp/bunken/about/research_2017.html)
- [37] ベネッセ総合教育研究所 <https://berd.benesse.jp/research/>
- [38] 東京大学社会科学研究所社会調査・データアーカイブ研究センター,  
<https://csrda.iss.u-tokyo.ac.jp/access/ssjda-guide/>
- [39] 次世代医療基盤法の施行について、[https://www.kantei.go.jp/jp/singi/kenkouiryoku/jisedai\\_kiban/houritsu.html](https://www.kantei.go.jp/jp/singi/kenkouiryoku/jisedai_kiban/houritsu.html)
- [40] Miyaji, A., Nakasho, K., Nishida, S., Privacy-Preserving Integration of Medical Data A Practical Multiparty Private Set Intersection, *Journal of Medical Systems*, Vol.41, No.3, pp.1-10, 2017. DOI: 10.1007/s10916-016-0657-4.
- [41] Kuromiya, H., Majumdar, R., Warriem, J., and Ogata. H., Data-driven Validation of Pedagogical Model -A Case of Blended LCM Model. Proc. of IEEE International Conference on Technology for Education (T4E) 2019, pp.38-45, doi: 10.1109/T4E.2019.00016.
- [42] 文部科学省. 教育の質の向上に向けた効果的なデータ連携・活用のポイントと学校改善事例集. 2019年. [https://www.mext.go.jp/content/1387543\\_02.pdf](https://www.mext.go.jp/content/1387543_02.pdf)

## ＜参考資料 1＞ 審議経過

2018 年

- 10 月 4 日 教育データ利活用分科会（第 1 回）  
役員の選出、現状の説明、今後の進め方について
- 11 月 13 日 教育データ利活用分科会（第 2 回）  
シンポジウムの開催について

2019 年

- 1 月 22 日 教育データ利活用分科会（第 3 回）  
提言の構成案について
- 3 月 22 日 教育データ利活用分科会（第 4 回）  
提言骨子案について
- 6 月 14 日 教育データ利活用分科会（第 5 回）  
提言の案について
- 7 月 23 日 教育データ利活用分科会（第 6 回）  
提言の分科会での承認について
- 10 月 4 日 教育データ利活用分科会（第 7 回）  
提言の修正について

2020 年

- 3 月 15 日 教育データ利活用分科会（第 8 回）  
提言の修正について
- 9 月 24 日 日本学術会議幹事会（第 299 回）  
提言「教育のデジタル化を踏まえた学習データの利活用に関する提言—エビデンスに基づく教育に向けて—」について承認

## ＜参考資料 2＞ LA（Learning Analytics）の研究事例

海外においては、より良い意思決定のためにはより良い情報が必要である、との原理のもと、米国では、教育省が 2005 年に州全域での時系列データシステム（SLDS: Statewide longitudinal data systems）の導入プログラムを開始し、初等中等教育を中心に履修科目や成績等の学生データの収集が始まった。2017 年には 50 州が参加しており、このデータを用いてドロップアウトの分析予測など様々な研究がされるようになった。さらに、大学教育においても、2012 年以降、ミシガン大学やスタンフォード大学、フェニックス大学、コロラド州立大学をはじめ、様々な大学に Learning Analytics (LA) に関するセンターが設置され、組織的に大学内の教育・学習活動のデータを収集して、LA の研究・実践が活発に行われるようになった。米国では各大学が独立して LA を行っているため、その

ツールやデータの共有のために、Unizin や LearnSphere 等のコミュニティサイトを立ち上げている。

英国では、オープン・ユニバーシティやエジンバラ大学が先駆的に LA を行っている。また、JISC(Joint Information Systems Committee)が、2014年から2017年にかけて、50以上の大学に対して LA の情報基盤システムの整備を行い、それぞれの大学において LA が実施されるようになった。

ノルウェーでは、国レベルで LA を推進するために 2016 年にベルゲン大学に SLATE(Center for the Science of Learning & Technology)を組織した。Uni-NETT が全大学に対して LMS の提供を行い、LA の基盤システムの開発を検討している。Uni-NETT は大学入試もオンラインテスト(CBT)で提供しているので、LA 基盤が整備されれば、大学入学時から卒業までの全ての学習データを集中して蓄積可能となる。

2011 年以降、教育ビッグデータの利活用を目的とした LA への関心が高まり、これに関する国際会議として LAK (Learning Analytics and Knowledge) や EDM(Educational Data Mining) が開催された。LAK は、2011 年以降毎年開催され、それを運営する国際学会として、2011 年に SoLAR 学会を設立した。EDM は主に機械学習や知識発見のアルゴリズムの応用に重点をおいており、2008 年以降毎年開催され、2011 年に国際学会 IEDMS (International Educational Data Mining Society) を設立した。

米国パデュー大学では、2007 年から、過去のデータから退学や留年の可能性の学習者にメッセージを送るコースシグナルシステム(CSC)を運用し、2010年には約 24,000 名の学習者を対象に学習者へのリスクのある学習者への自動フィードバックを行っている[32]。その結果、最終成績が A と B の学生は 10.37%増加し、D と F の学生は 6.41%減少した。また、大学一年生の継続率は、CSC を利用しないグループ(4,221 名)は 81.69%、CSC を利用したグループ(2,690 名)は 96.25%となった。学生の意見は 89%が CSC に好意的であり、74%が意欲を高める効果があると回答した。

米国マリスタ大学では、学部生 1,739 名を対象に、システムが何も介入しなかった場合 23.3%が退学したのに対し、リスクのある学生にシステムが介入した場合 13.5%のみが退学した[33]。

この他にも、米国ミシガン大学、ニューヨーク大学、アリゾナ州立大学などでは、LA を実践するセンターを設立し、大学全体で組織的に LA を実践している。リスクのある学習者への介入以外にも、GPA 等の成績の予測、e-ポートフォリオに記述された文章の分析、コース選択の提案、カリキュラムの提案、映像や音声、脳波・心拍などのセンサー情報などの多種多様なデータの分析を対象としたマルチモーダル学習分析、個人情報保護やポリシーの策定、教育機関へのスムーズな導入など、様々な実践研究が行われている。

### ＜参考資料 3＞質問紙調査の質問項目の例

文部科学省が実施している全国学力・学習状況調査＜参考資料 4 参照＞における児童生徒質問紙における質問項目を例示する。

#### 学習時間

学校の授業時間以外に、普段（月曜日から金曜日）、1日当たりどれくらいの時間、勉強をしますか（学習塾で勉強している時間や家庭教師に教わっている時間も含む）

（選択肢）① 3 時間以上／② 2 時間以上，3 時間より少ない／③ 1 時間以上，2 時間より少ない／④ 30 分以上，1 時間より少ない／⑤ 30 分より少ない／⑥ 全くしない

#### 学習方法

家で自分で計画を立てて勉強をしていますか

（選択肢）① している／② どちらかといえば，している／③ あまりしていない／④ 全くしていない

#### 学習意欲

国語の勉強は大切だと思いますか

（選択肢）① 当てはまる／② どちらかといえば，当てはまる／③ どちらかといえば，当てはまらない／④ 当てはまらない

#### 家庭環境(基本的生活習慣)

朝食を毎日食べていますか

（選択肢）① している／② どちらかといえば，している／③ あまりしていない／④ 全くしていない 10 92

毎日、同じくらいの時刻に寝ていますか

（選択肢）① している／② どちらかといえば，している／③ あまりしていない／④ 全くしていない 11 93

#### 家庭環境(家庭でのコミュニケーション)

家の人（兄弟姉妹を除く）と学校での出来事について話をしますか

（選択肢）① している／② どちらかといえば，している／③ あまりしていない／④ 全くしていない

### ＜参考資料 4＞大規模調査によって収集されている学習データの事例

文部科学省が実施している全国学力・学習状況調査は、2007 年度より日本全国の小中学校の最高学年（小学六年生、中学三年生）全員を対象とした悉皆調査（一時期、抽出調査。私立は任意）である。国語と算数・数学は毎年、理科は三年ごと（2012 年度～）に実施されている。また、2019 年度からは、中学校のみ英語も始まった（三年ごとに実施）。あわせて、生活習慣や学校環境に関する質

問紙調査（学習者および学校対象）が実施されている。これらのデータの分析結果の概要は web サイトで、また、詳細な分析結果は教育委員会や学校関係者に公開されている。しかし、ローデータの公開はほとんど行われてこなかった（現在では、個票データ等が大学等の研究者や公的機関の職員等に一定期間貸与されるようになっている）。OECD 生徒の学習到達度調査（PISA）のように、研究者が分析可能な形でローデータを広く公開することによって、学問的な分析と政策の効果分析と提言が可能になると考える。さらに、小学六年時のデータと中学三年時のデータを個人で紐付けることによって、中学三年間の学習者の成長と教育の効果を明らかにできる。

また、地方自治体教育委員会ごとに、地域の課題に即した学力調査、生徒質問紙、保護者質問紙、学校質問紙などが実施されている。これらのデータも、全国学力調査のデータと紐付ける形で、公開することによって、全国学力・学習状況調査の対象となっていない学年についても、同一学習者の追跡調査・縦断的分析などが可能になり、全国の研究者が多角的な分析を行い、優れた地域の教育実践の効果を明らかにできる。ただし、ローデータの公開にあたっては、地域のランク付けにならないように、個人名、学校名だけでなく地域情報も含めた匿名化データの作成や、学術目的に限定して公開するなどの措置が必要である。なお、学術研究のための匿名化マイクロデータの公開は国勢調査データで行われている[34]。なお、国立教育政策研究所では、国が実施してきた様々な調査（就学前児の調査、情報活用能力調査、いじめ調査など）の主な結果が web 上の研究成果アーカイブで公開されている。しかし、ローデータの公開はない。また、同研究所の教育情報共有ポータルサイト[35]は、同研究所や都道府県・市町村教育委員会等が作成した教育や学習に関するコンテンツ（学習指導案、教材・素材など）約三万件を集めていたが、現在更新は行われていない。

団体と民間教育企業のもつデータとして、特徴的なものを2つあげる。

NHK 放送文化研究所[36]では、次世代を担う子どもとメディアについての研究、幼児とデジタルメディアに関する研究などが継続的に行われている。その調査の概要は、web ページで報告されているが、ローデータの公開はない。

ベネッセ総合教育研究所[37]では、乳幼児・学習者調査、保護者調査などが継続的に実施されていて、調査の概要が報告されている。そのローデータは、次に述べる東京大学社会科学研究所社会調査・データアーカイブ研究センターで公開されている。

大学のもつ重要なデータアーカイブとしては、教育に特化したものではないが、東京大学社会科学研究所社会調査・データアーカイブ研究センター[38]には、学習データは239件（ベネッセ教育総合研究所、内閣府、全国生協実施の大学生データ、JGSS、労働政策研究・研修機構など）登録されている。ここで登録されているローデータ（個票データ）に基づいて、二次分析が可能である。

この他にも、各学校には毎年、指導要録・通知表・調査書のように学習者に関するものや、授業計画・シラバス・テストのように学校・教員に関するものなど、多岐にわたって学習データが収集されている。

## ＜参考資料 5＞教育データ利活用ガイドラインの宣言と原則[6]

### ＜教育データ利活用宣言＞

（地域を代表・貢献する）教育機関として、日々の教育や学習に関するデータを安全な方法で収集・保持・分析し、客観的データに基づく教育改善や学習者等の学習支援を図るととともに、データ解析から得られた叡智を公開し、国民と人類の福利に貢献します。

### ＜教育データ取扱い原則＞

- (1) 使用目的を明示し、目的外には使用しません。
- (2) 使用方法と使用結果を明示します。
- (3) いつでも本人または保護者同意を取り下げることができます。
- (4) 個人情報保護法などの関連する法令を遵守します。
- (5) いつでも自分のデータにアクセスできるようにします。このためのデータ分析ツール（ダッシュボード等）を提供します。
- (6) データの分析結果の公表については個人が決して特定されないようにします。
- (7) 教職員は自らの職務で知りうる範囲の個人情報を用い、学習者の学習支援の発展に寄与します。職務の範囲を超えた個人情報の利用を禁止します。
- (8) 研究利用を可能にするためのデータは適切に匿名化を行って提供し、学術の発展に寄与します。データに対して個人を特定するような処理を禁止します。
- (9) 研究成果やデータの共有によって、人類の福利に貢献します。

## ＜参考資料 6＞次世代医療基盤法について

医療の分野では、2018年5月11日から、カルテや検査データなど医療機関などがもつ患者の医療情報を匿名加工し、大学や企業での研究開発などでの活用ができるようにするための法律（次世代医療基盤法）が施行されている[39]。個人情報保護法に従えば、本来は、患者本人の同意が必要となるが、次世代医療基盤法では、患者本人が反対しなければ同意したとみなして、医療情報を第三者である認定事業者提供できるようにしている。また、本人同意があれば、医療等IDを用いて個人の健康・医療・介護のデータを統合し、医療・介護職等に提供可能である。このようにして収集した医療情報を分析して、政策立案や創薬、AIによる診断支援ソフトの開発、高度医療機器開発などに活用することが期待さ

れている。

### ＜参考資料 7＞学習データの連結と個人情報保護

学習データの利活用には2つの観点がある。第1は表2の項目1にみられるように、異なる教育機関で収集された個人の学習データを個人が利用する場合である。第2は表2の項目2, 3にみられるように、異なる教育機関で収集されたデータを学習データの解析に利用する場合である。第2は第三者への利用と考えてよいだろう。将来的には、異なる機関で収集されるデータにおいては、本人の利用、あるいは第三者利用の両方において、データの完全性、つまり、同一個人のデータを名寄せする方法を検討する必要があるかもしれないが、これについては、本提言では扱わず、今後の議論に委ねるが、現在の研究動向として以下を示す。

ここで、本人へ解析結果をフィードバックするためには、そのデータが確かに本人のデータであることを保証することは必須であるが、解析の段階では、1つ1つのデータが誰のデータであるかということは重要ではなく、データの完全性さえ保証できればよい。このような異なる教育機関で収集されたデータからデータを名寄せし、完全なデータベースを構築する方法として、データの集中管理手法が考えられる。しかしながら、集中管理はサイバー攻撃などによる個人情報漏洩の回避対策なども必要となり、データ管理コストが多くなる。それに対して、分散管理のまま、個人情報そのものは他機関に送付されずにデータ連結を行い、データの完全性のみを保証するセキュリティ手法がある[40]。このような手法を用いると、必要なときに各教育機関がもつ分散データからの名寄せが可能でかつ個人情報は漏洩しない。

一方、別の考えとして、個人情報の制御という方法もある。これは、第三者利用においても、個人にその情報を還元する必要がある、個人がそれを望む場合である。この場合、削除した個人情報を制御し、個人にトレースできるようにする必要があるだろう。このようなセキュリティ技術についても、今後、開発していく必要があるだろう。さらにセキュリティ技術には暗号化したまま平均値や分散などの統計処理を実現する技術（準同形暗号）や暗号化データへの検索クエリを実現する技術（検索暗号）などがある。今後、学習データ解析技術へのセキュリティ技術の適用、さらに両研究者での問題共有が個人情報を保護した学習データ解析には必須といえる。

### ＜参考資料 8＞安全な学習データの管理体制

安全な学習データの管理体制のためには、次の二点を十分に考慮する必要がある。

- (1) 数万に及ぶ学校や学習塾等の多数の機関で行われる学習、さらには自宅

等での自習に関するデータ収集と、横断解析の結果等を機関や個人にフィードバックにおける安全性

(2) 学習者の数年にわたる縦断解析のための長期保存における安全性

(1)に関して、伝統的には、境界防御(perimeter defense)の考え方にに基づき、機密性の高いデータを閉域網や物理的に隔離された室内に閉じ込める方式が主流であった。しかし、学習データを有効に活用するためには、全国に散らばる多くの学習者、教員、研究者等が、少なくとも間接的には学習データにアクセスする必要があるため、インターネットから完全に隔離するのは困難である(完全に隔離すると利便性が非常に大きく低下する)。また、学校や教育委員会などが個別にデータを保存する広域分散管理方式を採用するためには、各機関にセキュリティ・エンジニアを配置するなどの人的体制の整備が必要となるが、それだけの人材を養成・配置するのは容易ではない。そのため、伝統的な方式を徹底させるのは困難と思われる。一方、クラウド環境を前提に、より柔軟なアクセス制御方式を大規模に導入した事例としては、Google社のBeyondCorpやMITで進められているブロックチェーン技術を用いたMOOCs等の受講証明書の管理などが知られており、このようなシステム・アーキテクチャや運用方法の妥当性について検討を行う必要がある。

(2)に関しては、暗号の危殆化、証明書の失効、データ標準の変化、匿名化手法、機関の存続可能性などを合わせて検討する必要がある。学習データをセキュアに保存するために暗号技術は必須である。しかし、RSA、AESなどの今日の標準暗号方式は、計算論的安全性に基づくものであり、数十年のスパンで考えると十分な強度があるとは言えず、量子計算機の到来により解読される問題がある。この問題については、情報理論的安全性に基づく方式(たとえば、Secret Sharing)や耐量子暗号の研究が進められている。機密性ではなく真正性が問題になる場合には、電子署名を付加することがよくあるが、電子署名を検証するために必要な公開鍵証明書の寿命は短く、高々数年で失効する。この問題に対しては、RFC 5126などの方式が知られている。長期保存・長期利用を目指す場合に、収集・保存・解析の対象となるデータの属性が変化(おそらく増加)する可能性を意識する必要がある。学習データの有効活用のために、モジュラな拡張を許すようなデータの標準化が必須である。学校や学習塾は多数存在し、その全てが永続的に存続するわけではない。仮に各機関が学習データの保存に責任を有する場合、データの永続性を保障するためには、統廃合時のデータの継承方法を定める必要がある。また、個々の機関が適切なデータの保存と保護を行っていることを監査する仕組みも必要となる。一方、1つまたは少数の機関が集約的に学習データを保存・管理する場合、それらの機関の永続性を保証する仕組み、あるいは認定・監査の仕組みと認定を取り消された場合のデータの継承方法を定める必要がある。

### ＜参考資料 9＞学習データの利活用に関する倫理審査委員会の例

この倫理審査委員会では、学習者の人権と個人としての尊厳を脅かすようなことのないような制度設計や実施がなされているか、という観点から、以下の点を議論すべきであろう。

- ① 収集するデータの基本項目と取得方法
- ② 収集したデータの匿名化方法
- ③ 匿名化後のデータの管理・利用方法など

また、学習データの利活用に関する制度に関して以下の調査をする。

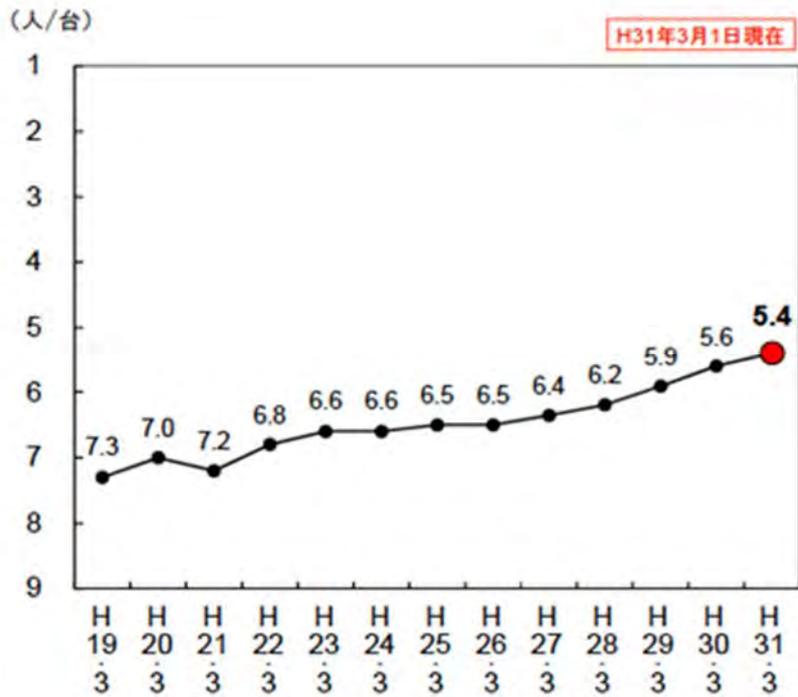
- ① 制度の内容に関して、各地で説明会やシンポジウムを開催したり、パブリックコメントを収集して国民全体から意見を収集
- ② 制度の改善のための調査研究・企画・報告書作成
- ③ 制度の実施状況の調査など

そのため委員は、以下の分野の専門家から構成されるべきであろう。

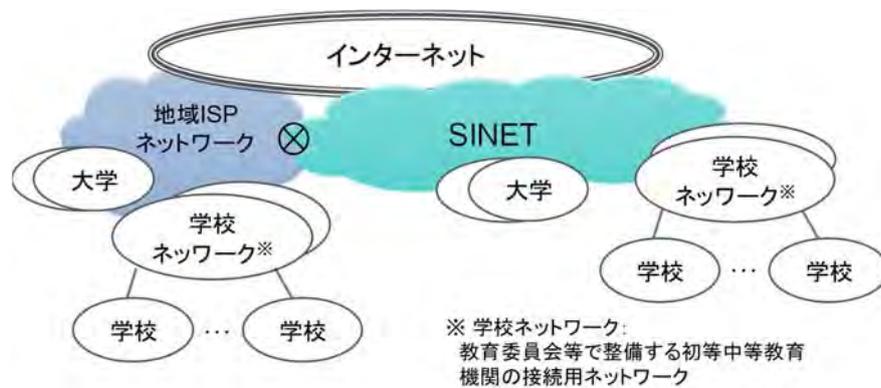
- ① 学習データの利活用に関連する分野の専門家
- ② データ科学・解析の専門家
- ③ 情報セキュリティの専門家
- ④ 個人情報保護法や人権などの法律の専門家
- ⑤ データ倫理の専門家
- ⑥ 教育委員会や学校長などの教育に携わる人など

なお、倫理審査委員会が関与するのは、あくまで国全体での学習データの収集方法と利活用の方法のみであり、学校や自治体内のデータの活用には関しては関与しない。倫理審査委員会の設置に関しては、情報銀行における、データ倫理委員会を参考に、国または第三者機関がその役割を担うことも考えられる。

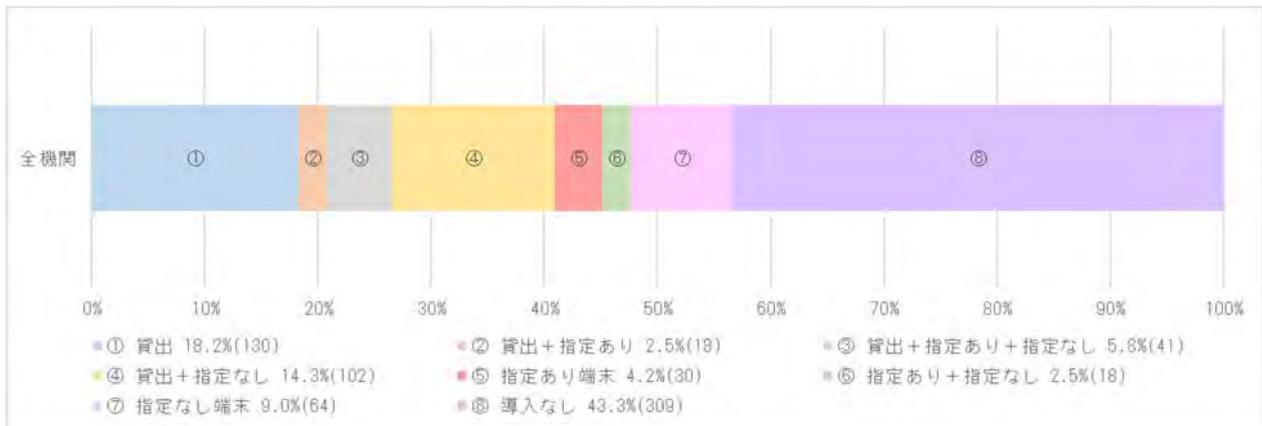
<参考資料 10> 初等中等教育機関における学習者用コンピュータの整備状況  
[16]



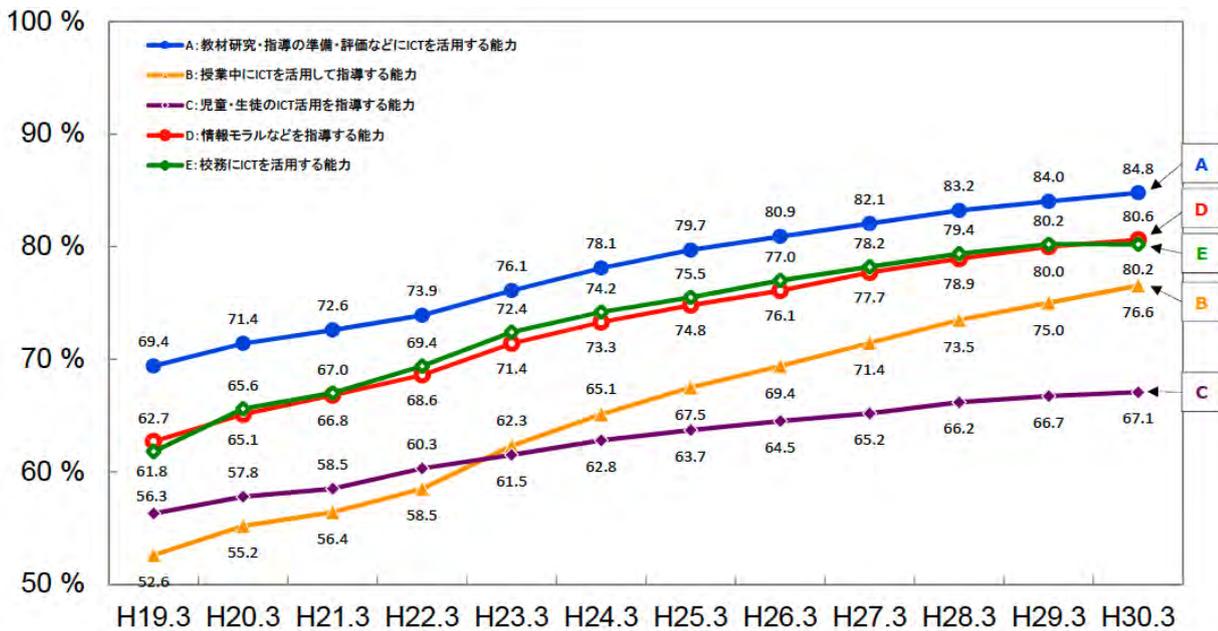
<参考資料 11> 初等中等教育機関における学校ネットワークと SINET の接続形態の例



<参考資料 1 2> 高等教育機関における BYOD の導入形態の状況[19]



<参考資料 1 3> 教員の ICT に関する能力の推移[16]



<参考資料 1 4> 提言の概要（実施時期は 2020～2030 年度を想定）

①教育データの種類とその必要性



- ・教育データの利活用は、個人、組織、国レベルに分類
- ・時間割や成績、教材、質問紙調査、LMS(Learning Management System)等の利用による学習プロセスのデータを基本項目として収集
- ・人工知能(AI)技術を用いて、教師の負担軽減やテラーメイド教育を実現

②教育データ利活用のための制度設計



- ・小学校、中学・高等学校、大学・大学院までを対象に、まずは、各教育機関で教育データを教育改善・授業改善のために利活用する。
- ・そのためには、教育機関が民間企業等のシステムや試験等を利用する場合、企業が教育データを囲い込む恐れがあるため、学習者の教育データを学校に提供するよう、契約時に盛り込む必要がある。
- ・国全体で教育データを収集する際には、「データを個人情報にあたらないように適切に処理して収集する」、「データの利用者は、個人を特定するような分析や地域や学校間の無用な比較をしない」などの制限を考慮した制度設計が必要。
- ・この制度は、国全体で継続して議論しながら、学習者の人権と個人としての尊厳を脅かすようなことのないよう、倫理審査委員会のような第三者機関を設けて、教育データの悪用を防ぎつつ適切に利活用が進むように定期的に見直していく。
- ・教育データの共有のためには、書式や意味(言葉の解釈)、データ標準化などを考慮すべき
- ・教育データを国レベルで管理するための体制づくりと長期的な予算措置が必要

③教育データ蓄積のための情報環境の整備



- ・一人一台の情報端末により、本格的なデジタル教科書や個人適応型学習が実現できる
- ・SINETに接続することにより、初等中等高等教育の機関が接続でき、教育データの収集と利活用がスムーズに行える。

④教員・LA専門員・研究者の養成



- ・教育データをうまく活用できる教員の養成とそれをサポートするLA専門員の配置
- ・効果的な教育方法や学習方法、AIによる新たな支援方法を見いだすことができる。教育サイエンティストの研究者を養成
- ・全国レベルで連結した教育ビッグデータを用いて、教育工学、教育学、認知科学などの様々な分野における学術研究を促進