

(案)

提 言

我が国における教育データの
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）〇月〇日

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

(案)

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

教育は次世代を育てる国の基盤であり、その高度化と充実は国全体の課題である。現在、教育機関では、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への移行など、情報通信技術(ICT)を利用した教育の新たな取り組みが行われている。この「教育の情報化」によって、デジタル教科書の閲覧履歴や試験の答案など様々な教育データが蓄積されているが、これを教育改善のために利活用する方法はあまり検討されていない。例えば、成績や質問紙調査などの従来のデータだけでなく、ICT利用によって記録される教育・学習のプロセスデータと統合して解析することで、いつ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等がきめ細かく把握できるようになる。さらに、国全体で、教育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、教育データの種類やその利活用の必要性について論じ、教育データを利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、国の政策として提言するものである。なお、ここでの教育データの利活用は、分析結果を学習者個人にフィードバックするという狭義のラーニングアナリティクス(以下、LA)に限定せず、これまで行われてきた質問紙調査に基づくデータ分析や大学運営のための教学IR(Institutional Research)を含むものである。

2 提言

本提言は、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。

(1) 教育改善を目的とした教育データの収集と分析

教育データの利活用の目的は、教員や学習者（学習者の保護者）を含めた個々の教育・学習活動を支援する個人レベル、教育機関全体で分析してカリキュラムの最適化等をはかる機関レベル、政策立案者・研究者・市民などがエビデンスに基づく教育政策の立案や検証などのために利用する国レベルに分類できる。教育の情報化が進展している中で、これらの目的を達成するためには、時間割や成績、教材、質問紙調査などの従来のデータだけでなく、情報機器やLMS(Learning Management System)等を利用した細かな学習プロセスデータも基本項目として収集すべきである。なお、これらの基本項目は人種や病歴などの「要配慮個人情報」を除外しつつ、実践を通して精査していく必要がある。

(2) 教育データを利活用するための制度設計

教育データの収集は、それが比較的行いやすい小学校から大学・大学院までの公教育をまずは対象として集め、各教育機関等内で、教育の改善に利活用すべきである。そして小中学校等が属する地方自治体や国立大学法人、私立学校等の法

(案)

人内部で教育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が現実的であろう。このように、教育データの収集・利活用方法などの制度設計を国が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、契約時に盛り込むこと、②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、教育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、教育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。なお、この制度は、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、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必要がある（参考資料 9）。

(3) 教育データを収集するための情報環境の整備

教育データの中でも学習プロセス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、授業内外で制限なく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教育データを収集・利活用することも考える必要がある。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) 教員・LA 専門員・研究者の育成

教育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、教育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、蓄積された教育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、教育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、どのような教育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。例えば、過去の事例では、成績の向上などの効果が示されているが、自主性や創造性などの他の要素への効果については、今後、長期的な実践を通して確かめていく必要がある。

目 次

1	はじめに	1
2	現状と問題点	2
	(1) 教育データの利活用の現状と問題点	2
	(2) 教育データを収集・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	15
3	提言について	17
	(1) 教育改善を目的とした教育データの収集	17
	(2) 教育データを収集・利活用するための制度設計	18
	(3) 教育データを収集するための情報環境の整備	19
	(4) 教育データを利活用するための人材の育成	20
4	おわりに	20
	<参考文献>	21
	<参考資料1> 審議経過	23
	<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	25
	<参考資料4> 大規模調査によって収集されている教育データの事例	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]	27
	<参考資料6> 次世代医療基盤法について	28
	<参考資料7> 教育データの連結と個人情報保護	28
	<参考資料8> 安全な教育データの管理体制	29
	<参考資料9> 教育データの利活用に関する倫理審査委員会の例	320
	<参考資料11> 教員のICTに関する能力の推移[16]	321
	<参考資料10> 提言の概要（実施時期は2020～2030年度を想定）	32

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、教育機関では、タブレット端末の導入など、情報通信技術 (ICT) を利用した教育の新たな取り組みが行われている。この「教育の情報化」によって様々な教育データが蓄積されているが、これを教育改善のために利活用する方法はあまり検討されていない。多種多様な教育データを用いて、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、など成績が上がらない理由を把握し、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、教育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、教育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、大学全体を対象に教育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情報化の推進、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教育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教育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教育データを対象としている点で異なる。

(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

(案)

れる。このエビデンスの蓄積のためには、国全体での教育データの収集が鍵となる。また、教育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

教育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。

- ・教育データの利活用の目的と集めるべき教育データの整理
- ・教育データを収集・利活用するための制度設計
- ・一人一台の情報端末、インターネット接続等の情報環境の整備
- ・教育データを活用して授業や教育改善を行う人材の養成

(1) 教育データの利活用の現状と問題点

① 教育データの利活用の現状

教育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、対面式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教育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を行うには教育データの収集・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

(案)

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。

表 1 教育データの利活用の目的の分類

	対象	誰のため	目的の例
1	個人	学習者	・個人に適した教材や問題の推薦による学習効果の向上 ・過去や現在の教育データを用いた、理解度の予測などによる、個人の学習状況の把握
		教員	・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 ・自動採点など、教育データの利用による教員の負荷の軽減
		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	・教育データに基づくカリキュラムの最適化 ・教員や学習者の最適な配置
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

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。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教育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。また、表1に示すように、教育データは、1の学習者

(案)

や教員による利用だけでなく、2 教育機関や研究者、社会全体での利用も考えられる。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、教育データを1～3の全ての目的で利用するものとする。

表2 教育データの種類（灰色は収集すべき基本項目）

	データ項目	具体例	利用目的
1	学校データ	学校情報、科目名、時間割など	全ての分析に必要な
2	授業設計データ	教育目標、カリキュラム、シラバス、指導内容など	
3	教材データ	教科書、補助教材、問題集など	
4	人的データ	教員や学習者の性別、学年など	
5	学習評価データ	最終成績、小テスト・レポート点数、入試情報、出欠など	
6	質問紙データ	学習時間や学習方法、学習意欲、家庭環境などの質問紙調査（参考資料3）、授業評価アンケートなど	個人の成績とその他の要因との関係等の分析研究
7	学習プロセスデータ	デジタル教材閲覧履歴、LMS 等のログイン・ログアウト時間、e ポートフォリオ、オンラインテスト回答時間など	学習者の学習プロセスや教員の教育プロセスを個別に細かく分析するために必要
8	環境データ	講義映像、気温、湿度、天気、部屋の明るさ、騒音、二酸化炭素濃度など	教室内の環境や活動量等のデータと学習効果との関連性の調査などの研究目的で主に利用 (要配慮個人情報には含まないように配慮が必要)
9	健康データ	身長・体重などの健康診断データ、50m 走・反復横跳び等の体力測定データ等の学校がもつデータ	
10	個別取得データ	日々の歩数等の運動量・食事等のデータ、脳波・視線等のセンサー情報など	

② 教育データの分類

表1の1～3の全ての対象のために必要となる教育データの概要を、表2に示す。表2の1～5は、多くの分析において必要となるデータである。6は、家庭の学習環境などのアンケートでのみ取得できるデータもあるため、多面的な分析をする際に必要となる。7は、学習者や教員の教育・学習活動

(案)

のきめ細かな学習プロセスのデータを用いて、教育・学習支援を行うために必要となる。従って、ここでは、表2の1～7の教育データを基本的に集めるべきであろう。これらのデータは、今後のギガスクール構想や教育のトランスフォーメーションの進展によって、情報端末や教務支援システム、LMS等を用いて、電子的に蓄積され、LAの研究で効果が示されてきたものが中心である。しかし、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いられないよう、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る。8, 9, 10は、主に研究目的で利用されるデータであり、共有する必要性は低いと考えられる。なお、実際の教育現場では、人種や病歴などの要配慮個人情報や、障がいや性自認にかかわる教育活動上の配慮策やいじめや虐待など、表2の内容より多くの情報が扱われているが、本提言では、それらは扱わず、表1の目的を達成するために必要であり、かつ、共有が比較的しやすい1～7の項目に焦点をあてている。また、表2では例をあげるため、「など」と記述しているが、実施にあたっては収集するデータの内容を限定する必要がある。今後、実践を進めながらこれらの項目を精査していく必要がある。

③ 収集されている教育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・民間機関による教育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教育データの概要）。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教育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、個票の教育データを収集し、利活用できるような制度設計を検討する必要がある。

④ 教育データの収集対象となる機関

現在、教育データは、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、教育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・大学までの教育データを対象に、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または公的機関が収集し、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

(案)

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、データの収集の対象は、最初は、生涯にわたって教育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・評価のために教育データを収集・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 教育データの利活用の段階の分類

レベル	説明	教育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教育データを共有・活用ができていない。	・法人内で教育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	法人内では教育データを共有し、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。	・同一法人内で教育データの活用は可能 ・教育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	法人内で教育データを共有・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。	・法人内外での教育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教育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	法人内で教育データを共有・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。	・法人内外での教育データの利活用は可能 ・個人情報保護に配慮した教育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教育データを計算機で処理することが容易

⑤ 教育データの保存期間

教育データは、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、学習プロセスデータなどは、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1の目的のためには、過去から現在までの大量の教育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、大規模な教育データは、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、大規模な教育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たり、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教育データを長期間保存したり、国レベルでは、個人情報保護に配慮して教育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、教育データは個人情報であるため、国全体で教育データを収集・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) 教育データを収集・利活用するための制度の現状と問題点

教育データの多くは、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もある。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 教育データ利活用のレベル

表3に示すように、教育データの共有と利活用の範囲に応じて、教育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。教育データを最大限に利活用するには、国全体で教育データの利活用を行うレベル3を実現すべきであろう。しかし、現状ではレベル0、つまり、教育機関内でも教育データの共有・活用ができていない場合がほとんどである。先行して教育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、教育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、法人間をまたいで教育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、法人間をまたいで教育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を実現した方がよいと考えられる。そうしなけ

(案)

れば、仮に法人間をまたいで教育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、その内部で教育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。むしろ自らの教育データを内部に囲い込む場合が多い。しかし、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り、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教育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 教育データ利活用ポリシーの現状

教育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、個人情報の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教育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教育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。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、教育データの利活用における注意点も盛り込む必要がある。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、教育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、教育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教育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教育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

(案)

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。

- ・教育に係る選択は本人が実施するものであり、教育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、教務において教育データの利用によって知りえたことがら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、過去の教育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教育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）。

③ 教育データの管理体制

現状の個人情報保護法のもとで、国または公的機関が教育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ろう。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、教育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であろう。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であろう。一方、教育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

(案)

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必要がある。

教育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：教育データの連結と個人情報保護）。

表4 教育データの共有に伴って必要となる匿名化の段階の例

	教育データの保存	教育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教育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、教育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、教育データを比較的、容易に利用できること。

- ・教育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、教育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：安全な教育データの管理体制）。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 教育データのフォーマットと意味の統一

将来的に教育データは、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教育データを共有し、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、異なる機関間での教育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、教育データの利活用には非常に重要となる。

表5 学校の情報環境整備のレベル

レベル	説明	教育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教育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

⑤ 教科書や教材等の教育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、このような教育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、教育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。

(3) 教育現場における情報環境整備の状況

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2の教育データの中でも学習プロセスデータは、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

(案)

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（図 1）[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。教育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教育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]、これを用いて教育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図 2 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（図 2 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、教育データの収集・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、教育データ利活用の PDCA サイクルの効率化に繋がる。また教育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。

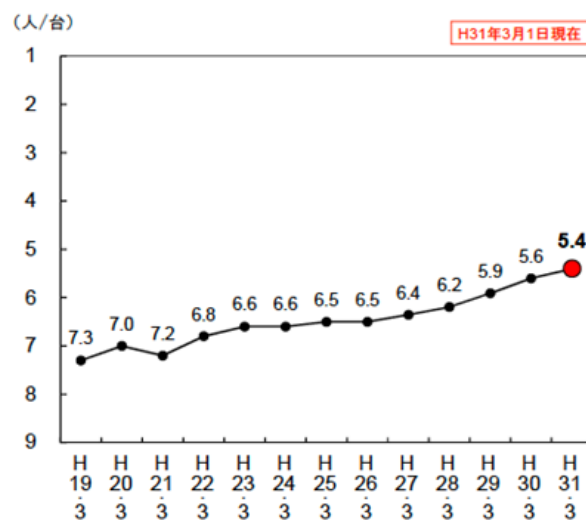


図 1 学習者コンピュータの整備状況[16]

(案)

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。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]。

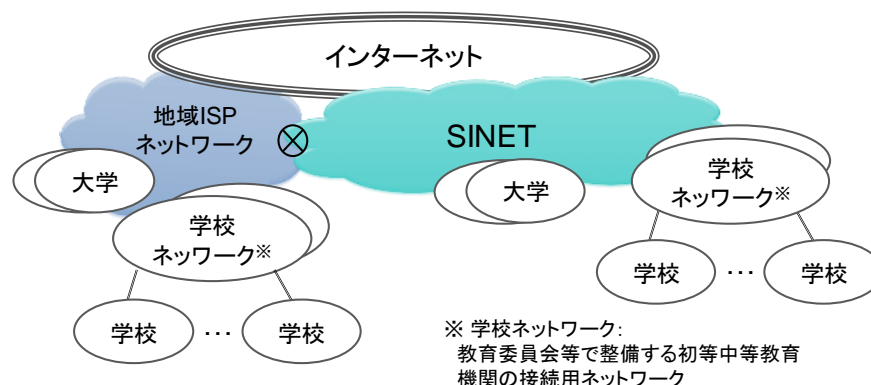


図2 初等中等教育機関における学校ネットワークと SINET の接続形態の例

教育データの蓄積には、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

(案)

[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。

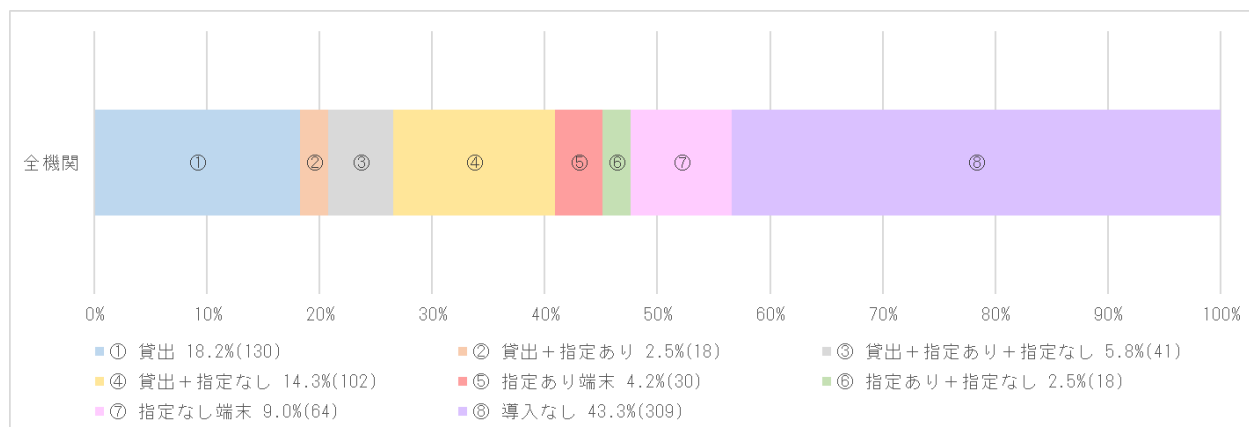


図3 高等教育機関におけるBYODの導入形態の状況(2016年度)[19]

(4) 教育現場における人材の状況

① 教員養成課程における教育の情報化への対応の遅れ

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、教育の情報化を直接扱う学科・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教育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、教育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、教育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「教育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、教育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

(案)

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 10)。[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、教育の情報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、教育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、さらに教育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、教育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 教育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、教育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、教育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 教育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、教育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教育データサイエンティストの育成が急務である。このため、教育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、教育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、教育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、教育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 1 参照)

(1) 教育改善を目的とした教育データの収集

① 教育データの収集目的

教育データ利活用の目的としては、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、国全体で教育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1)。

② 集めるべき教育データの種類

教育データは、(1)学校データ、(2)授業設計データ、(3)教材データ、(4)人的データ、(5)学習評価データ、(6)質問紙データ、(7)学習プロセスデータ、(8)環境データ、(9)健康データ、(10)個別収集データがあり、この提言では収集すべき基本項目として(1)～(7)を提起する(表2)。また、教育データの取得はできる限り電子的な形で一人一台の情報端末の環境で効率的に、かつ教員や学習者に負担の少ない方法でデータを収集すべきであろう。

④ 教育データ収集対象となる機関

表1の目的を達成するために、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、現在入手可能な教育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、同一機関内で教育データを収集し、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教育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。教育データの管理は、各公的教育機関が標準的なシステムを用いて行い、教育データを収集する国や公的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。

⑤ 教育データの保存期間

教育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) 教育データを収集・利活用するための制度設計

① 社会全体での教育データの利活用

教育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教育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ろう。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、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、契約時に盛り込むこと、②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、教育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 教育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教育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教育データを利用しない」などの教育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ろう。

③ 教育データの管理体制

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教育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、教育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、教育データを比較的、容易に利用できること。

- ・教育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、教育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 教育データのフォーマットと意味の統一

教育データの書式は、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 教育データの著作権

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) 教育データを収集するための情報環境の整備

① 教育データを収集するための情報端末の整備

教育データを、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、新しく買い与えるか、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利用する BYOD を検討する必要があるであろう。

② インターネットへの接続

教育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接続できる環境を整備する必要がある。また、教育データはクラウド上に情報収集用の情報基盤が構築されることが想定されるが、その情報基盤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 を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

(案)

連携による改善を模索する必要がある。また、健康やメンタル面にも配慮した適切な情報環境の構築が望まれる。

(4) 教育データを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、教育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、教育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、教育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。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 教育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、教育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教育データサイエンティストの育成が急務である。これには、教育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、教育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教育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、教育データ収集・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
- [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.
- [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.
- [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.
- [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.
- [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.
- [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.

(案)

- [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.
- [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.
- [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
- [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.
- [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.
- [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
- [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
- [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

(案)

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., Lauri´a, 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
- [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
- [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.

<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）

(案)

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

<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教育・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、

(案)

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

(案)

(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教育データの事例

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

(案)

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、教育学習データは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、多岐にわたって教育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叢智を公開し、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。

(案)

- (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に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＞教育データの連結と個人情報保護

教育データの利活用には2つの観点がある。第1は表1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教育データを個人が利用する場合である。第2は表1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教育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るとよいだろう。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のである。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、教育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教育データ解析には必須といえる。

<参考資料 8> 安全な教育データの管理体制

安全な教育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、教育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教育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

(案)

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。教育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。教育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。仮に各機関が教育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教育データを保存・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＞教育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、教育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 制度の関する国民からの意見の収集
- 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

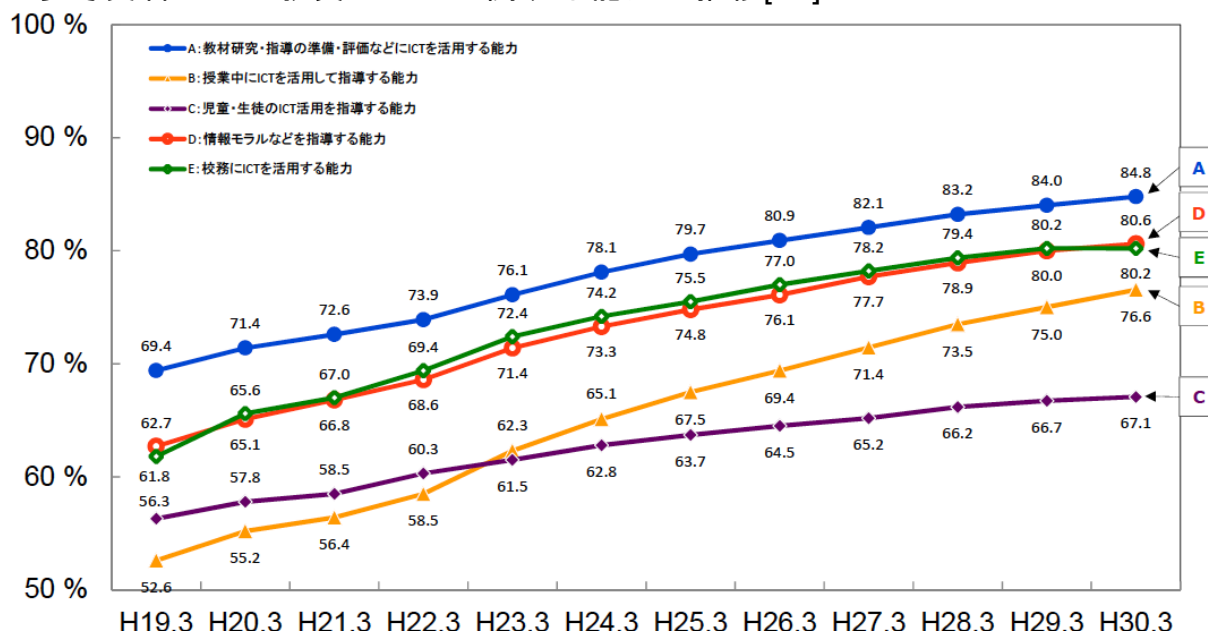
(案)

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 教育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教育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> 教員の ICT に関する能力の推移[16]



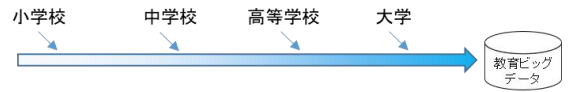
<参考資料 1 1> 提言の概要（実施時期は 2020～2030 年度を想定）

①教育データの種類とその必要性



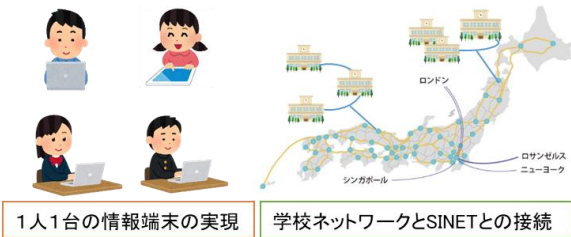
- ・教育データの利活用は、個人、組織、国レベルに分類
- ・時間割や成績、教材、質問紙調査、LMS(Learning Management System)等の利用による学習プロセスのデータを基本項目として収集
- ・人工知能(AI)技術を用いて、教師の負担軽減やテラーメイド教育を実現

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

提言等の提出チェックシート

このチェックシートは、日本学術会議において意思の表出（提言・報告・回答、以下「提言等」という）の査読を円滑に行い、提言等（案）の作成者、査読者、事務局等の労力を最終的に軽減するためのものです¹。

提言等（案）の作成者は提出の際に以下の項目を1～11をチェックし、さらに英文タイトル（必須）、英文アブストラクト（任意）、SDGs との関連の有無（任意）を記載し、提言等（案）に添えて査読時に提出してください。

記入者（委員会等名・氏名）：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

和文タイトル 我が国における教育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

英文タイトル（ネイティブ・チェックを受けてください）

The Effective Use of Educational Data in Japan --- Towards Evidence-Based Education ---

	項目	チェック
1. 表題	表題と内容は一致している。	1. はい
2. 論理展開 1	どのような現状があり、何が問題であるかが十分に記述されている。	1. はい
3. 論理展開 2	特に提言については、政策等への実現に向けて、具体的な行政等の担当部局を想定していますか（例：文部科学省研究振興局等）。	1. 部局名：(1)文部科学省初等中等教育局・高等教育局・教育再生会議、(2)経済産業省商務・サービスグループ教育産業室、(3)総務省情報流通行政局
4. 読みやすさ 1	本文は 20 ページ（A4、フォント 12P、40 字×38 行）以内である。※図表を含む	1. はい
5. 読みやすさ 2	専門家でなくとも、十分理解できる内容であり、文章としてよく練られている。	1. はい

¹ 参考：日本学術会議会長メッセージ、「提言等の円滑な審議のために」（2014 年 5 月 30 日）。
<http://www.scj.go.jp/ja/head/pdf/1>

6. 要旨	要旨は、要旨のみでも独立した文章として読めるものであり 2 ページ (A4、フォント 12P、40 字×38 行) 以内である。	1. はい
7. エビデンス	記述・主張を裏付けるデータ、出典、参考文献をすべて掲載 した。	1. はい
8. 適切な引用	いわゆる「コピペ」(出典を示さないで引用を行うこと) や、 内容をゆがめた引用等を行わず、適切な引用を行った。	1. はい
9. 既出の提言等と の関係	日本学術会議の既出の関連提言等を踏まえ、議論を展開して いる。	1. はい
10. 利益誘導	利益誘導と誤解されることのない内容である。	1. はい
11. 委員会等の趣旨 整合	委員会・分科会の設置趣旨と整合している。	1. はい

※9で「はい」を記入した場合、その提言等のタイトルと発出委員会・年月日、既出の提言等との関係、相違点等について概要をお書きください

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－、学術振興の観点から国立大学の教育研究と国による支援のあり方を考える検討委員会、2017年6月、この提言は、国立大学の教育研究と国による支援のあり方をまとめたものである。一方、本提言は、国全体の教育データの利活用のあり方をまとめたものであり、目的を異にする。

※チェック欄で「いいえ」を選択した場合、その理由があればお書きください

◎ SDGs (持続可能な開発目標) との関連 (任意)

以下の 17 の目標のうち、提出する提言等 (案) が関連するものに○をつけてください (複数可)。提言等公表後、学術会議 HP 上「SDGs と学術会議」コーナーで紹介します。

1. () 貧困をなくそう
2. () 飢餓をゼロに
3. () すべての人に保健と福祉を
4. (○) 質の高い教育をみんなに
5. () ジェンダー平等を実現しよう
6. () 安全な水とトイレを世界中に
7. () エネルギーをみんなに、そしてクリーンに
8. () 働きがいも経済成長も
9. (○) 産業と技術革新の基盤をつくろう
10. (○) 人や国の不平等をなくそう
11. () 住み続けられるまちづくりを
12. () つくる責任つかう責任
13. () 気候変動に具体的な対策を
14. () 海の豊かさを守ろう
15. () 陸の豊かさを守ろう

16. () 平和と公正をすべての人に
17. () パートナースHIPで目標を達成しよう

※「持続可能な開発目標 (SDGs)」とは

2015年9月に国連総会が決議した「我々の世界を変革する：持続可能な開発のための2030アジェンダ」が掲げた目標。

詳細は国連広報センターHPをご覧ください。

http://www.unic.or.jp/activities/economic_social_development/sustainable_development/2030agenda/

提言等公表時のSDGs説明

この説明は、日本学術会議の意思の表出（提言・報告・回答、以下「提言等」という）を日本学術会議ホームページのSDGsコーナーで紹介し、多くの関係者の閲読を促進するためのものです。

提言提出時のチェックシートにおいてSDGsとの関連に記述した場合は、日本語紹介文と英文アブストラクトを記載し、提出してください。

記入者（委員会等名・氏名）：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

和文タイトル 我が国における教育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

◎ SDGs（持続可能な開発目標）との関連

チェックシートで選択した項目に○をつけてください。

1. () 貧困 2. () 飢餓 3. () 健康 4. (○) 教育
5. () ジェンダー平等 6. () 安全な水 7. () エネルギー
8. () 経済成長 9. (○) 産業と技術革新 10. (○) 不平等
11. () まちづくり 12. () つくるつかう責任 13. () 気候変動
14. () 海の豊かさ 15. () 陸の豊かさ 16. () 平和と公正
17. () パートナースHIP

◎ 和文紹介文 200字以内

教育の情報化により、デジタル教科書の閲覧履歴や試験の答案など様々な教育データが蓄積されているが、これを社会全体で共有・活用することはあまり検討されていない。これにより、児童・生徒・学生の多面的なデータの分析による研究が促進され、エビデンスに基づく教育が実現できる。本提言は、教育データの種類やその利活用の必要性について論じ、教育データを利活用するための制度設計や支援体制等のあり方について提言する。

◎ 英文アブストラクト 150 words 以内

Various educational data such as digital textbooks reading logs and responses of examinations have

been accumulated due to the digital transformation in education. But sharing and utilizing these data throughout our society is not considered much. It facilitates research based on data analysis and enables evidence-based education. This proposal discusses the types of educational data and the need to use them, and proposes the design of policy, systems and support to share and use educational data.

◎ キャッチフレーズ 20 字以内

教育データの利活用で教育改革を推進しよう

◎ キーワード 5つ程度

教育データ、スタディログ、学習分析、教学 IR、EBPM(根拠に基づく政策立案)