

記 録

文書番号	S C J 第 25 期 050908-25530800-089
委員会等名	日本学術会議情報学委員会・心理学・教育学委員会合同 教育データ利活用分科会
標題	教育データの利活用のさらなる促進に向けた考察 ～データ駆動型教育への対応に向けた論点整理～
作成日	令和5年（2023年）9月8日

※ 本資料は、日本学術会議会則第二条に定める意思の表出ではない。掲載されたデータ等には、確認を要するものが含まれる可能性がある。

この記録は、日本学術会議情報学委員会・心理学・教育学委員会合同教育データ利活用分科会の議論の結果を取りまとめ公表するものである。

日本学術会議情報学委員会・心理学・教育学委員会合同
教育データ利活用分科会

委員長	美濃 導彦	(第三部会員)	国立研究開発法人理化学研究所情報統合本部本部長
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)	九州大学理事・副学長
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
	大場みち子	(第三部会員)	京都橘大学工学部情報工学科教授
	宮地 充子	(第三部会員)	大阪大学大学院工学研究科教授
	相原 玲二	(連携会員)	広島大学上席特任学術研究員特命教授
	稲垣 成哲	(連携会員)	立教大学文学部特任教授
	乾 健太郎	(連携会員)	東北大学大学院情報科学研究科教授
	岩瀬 峰代	(連携会員)	株式会社サイバコ代表取締役
	江村 克己	(連携会員)	福島国際研究教育機構 (F-REI) 理事
	柴山 悦哉	(連携会員)	東京大学情報基盤センター教授
	椿 美智子	(連携会員)	東京理科大学経営学部経営学科教授・経営学部長・大学院経営学研究科長
	中山 迅	(連携会員)	宮崎国際大学教育学部教授
	野澤 祥子	(連携会員)	東京大学大学院教育学研究科附属発達保育実践政策学センター准教授
	前田 香織	(連携会員)	広島市立大学最高デジタル責任者 (CDO)・特任教授
	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

本記録の執筆にあたり、以下の方にご協力頂いた。

大久保 文哉
久富 望

九州大学システム情報科学研究院准教授
京都大学大学院教育学研究科助教

要 旨

1 作成の背景

デジタル社会が進むにつれて教育のデジタル化も進んでいるが、教育で使われるデジタル端末上のデータを適切な形で蓄積し、教育に役立てることは、世界的にもまだ発展途上であり、大きな可能性を秘めている。本分科会では、小学校から大学までの教育機関における学習におけるデータ（健康観察・日常所見に関する情報は含まない）の利活用について、教育現場にも学術的にも配慮された議論を促し、政策にも反映させることを目指し、以下の4点にまとめ、2020年9月に提言「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—」を公開した[1]。

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

本文書ではこれら4点について、2020年の提言以後3年間の教育データ利活用の進展を踏まえ、現状に対する報告と、AIの透明性やデータの公平・公正な利用など、追加で考慮すべき事項をまとめ、日本における教育データ利活用による教育の質の向上に資することを目指したものである。なお、これは、文部科学省・経済産業省・総務省・デジタル庁・こども家庭庁などの国や自治体の教育政策に関する機関に向けたものである。以下、上記4点(A)～(D)に対する現状報告と追加すべきことを示す。

2 内容

(1) 2020年の提言の現状に対する報告

提言後の3年間における社会状況の激変により、教育DXの必要性が浸透し、文部科学省やデジタル庁からの各種委員会での文書や、経済産業省や総務省も含めた各省庁からの事業が進められている(A)。特に、GIGAスクール構想により全国的に小中学校における1人1台の端末環境やネットワークは整備され、教育のデジタル化が大幅に進んでいる。また、全国各地の大学・研究機関、各省庁などの取組もあり、教育データ利活用を担う環境が整いつつある(C)。一方で、不十分な点も多い。学習データの種類や教育改善の目的に基づいた利活用について、根本的な問題が先送りされている(A)。教育データへの当事者によるアクセスの確保や教育データの標準化が不十分であり、学校や自治体における教育データの一次利用はあまり進んでいない(B)。また、全国で統一したルールの不足もあって、匿名化による教育データの共有・二次利用はほとんど進んでいない。また、教育現場で教育データを利活用する教員の育成や研究者の相互協力・共同研究の事例もまだまだ少ない(D)。

(2) 2020年の提言に加えて追加して推進すべきこと

2020年の提言の対象であった公的教育機関の教育データに加え、乳幼児や社会人におけるデータ利活用や、公教育以外とのデータ連携についての議論を始めるべきである(A)。また、各教育機関がデータの名寄せの必要性を認識し、適切なIDの連携と管理体制を構築し、円滑かつ適切な二次利用のための制度設計も必要である(B)。特に、デジタル教科書を用

いて行われた学習活動については、全国で統一したログデータの流通基盤システムについて設計し、コンテンツを用いた教育的内容の改善に直結するようにする必要がある(C)。さらに、教育データ利活用のベストプラクティスを共有し、教育分野の魅力を高め、優秀な人材を惹きつけ、デジタル化に適した社会を実現するための努力を続ける必要がある(D)。

(3) 2020年の提言に加えて、注意すべきこと

社会におけるデータ利活用の拡がりに伴い、既存の個人情報保護法等の法令、ポリシーやガイドラインの枠組みを超えた広い観点から議論が必要である。新たな法令、ポリシーやガイドライン等を考慮する上で、研究開発の社会実装時に生じる倫理的・法的・社会的課題(ELSI)の観点や、プライバシー保護に関する法規制の世界的情勢との関わりに加え、教育を受ける権利や教育の自由に関わる問題としても注意されなければならない。より具体的には、まず、教育データ利活用におけるデータの収集・保存・処理について透明性の確保が必要である。さらに、それらのデータに基づいた分析・フィードバックや、そもそもの利用目的を理解するために、教育データ利活用の利用者へのリテラシー教育を進めなければならない。一方で、分析の一部を担うであろうAIに対して、人間による判断を介在させるような一定の規制が必要であることが議論されており、その原則の教育分野への応用も議論が必要である。また、教育データ利活用が生涯に亘ったとき、その利点だけでなく、負の側面についても十分な議論が必要である。このように、教育データ利活用は社会的な困難も多い課題ではあるが、全体としてサステナブルな人間社会のシステムの一部として機能するよう、十分に配慮されるべきである。

目 次

1. はじめに	1
2. 2020年の提言後の3年間について	1
3. 2020年の提言に加えて推進すべきこと	4
4. 2020年の提言に加えて、注意すべきこと	8
5. まとめ	10

1 はじめに

デジタル社会が進むにつれて教育のデジタル化も進んでいるが、教育で使われるデジタル端末上のデータを適切な形で蓄積し、教育に役立てることは、世界的にもまだ発展途上であり、大きな可能性を秘めている。本分科会では、小学校から大学までの教育機関における学習におけるデータ（健康観察・日常所見に関する情報は含まない）の利活用について、教育現場にも学術的にも配慮された議論を促し、政策にも反映させることを目指し、以下の4点にまとめ、2020年9月に、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—」[1]を公開した。

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

本記録は、2020年の提言以降3年間の教育データ利活用の進展を踏まえ、現状に対する報告と、AIの透明性やデータの公平・公正な利用など、追加で考慮すべき事項をまとめ、日本における教育データ利活用による教育の質の向上に資することを目指したものである。また、本記録では、教員が教育を行う上で生じる教育に関するデータも教育データとして扱い、小中高大の前後である、乳幼児や社会人における教育データ利活用について、メリット等の考察を行った。本分科会での議論の結果を参考資料1にまとめる。なお、本記録は、文部科学省・経済産業省・総務省・デジタル庁・こども家庭庁などの国や自治体の教育政策に関する機関に向けたものである

2. 2020年の提言後の3年間について

2.1. 提言後の3年間で進んだこと

2.1.1. 教育データ利活用の機運の高まり

この3年あまりの間に、COVID-19感染拡大に伴ってデジタル技術の必要性が広く認識され、利用が促進され、ビッグデータ・AI技術の発展は著しく進んだ。また、デジタル・トランスフォーメーション（DX）についても広く浸透し、国においてもSociety 5.0を実現すべくデジタル庁が設置された。OECD、ASSISTments¹、九州大学のような大規模な例[2]に加え、文科省教育DX・実証事業などの予算化[3]に伴い小規模な学校でも教育データを用いた授業改善の例が積み重なり始め、個々の変遷を踏まえた学校内での縦断的な分析が可能になり始めている[4]。

また、GIGAスクール構想によって実現した1人1台端末の環境のもとでは、従来では不可能だった量と質の教育データが収集できる。このことは、従来の教育的知見や理論を吟味して発展させることや、光を当てるのが困難であった教育の対象・活動・課題についての新たな抽出・検討に繋げうるが、この可能性は、研究者・教育現場の教員など様々な立場に認識されつつある。

さらに、匿名化による教育データの共有と、その意義の浸透も進んでいる。教育データの分析技術の普及を目的に、2020年にはSIP教育データ解析チャレンジコンテスト、2021

¹ <https://new.assistments.org/>

年には一般社団法人エビデンス駆動型教育研究協議会²が設立され第1回教育データ分析コンテストが、教育・学習支援システムを利用したログデータを共有したコンペティションとして実施されている。

2.1.2. 2020年提言の要点(A)-(D)における進展

前提言(A) 学習データの種類と教育改善のための利用においては、教育現場において新型コロナウイルスによるオンライン教育の検討や実施が幅広く起き、教育DXの必要性が浸透した。内閣府の教育再生会議が教育データ利活用を政策として含め、文部科学省、経済産業省、総務省は各々の立場から事業を進め、デジタル庁からはロードマップ[3]において教育データの種類や利用方法が整理されている。それらは、「教育・学習は、技術に優先すること」「最新・汎用的な技術を活用すること」「簡便かつ効果的な仕組みを目指すこと」「安全・安心を確保すること」「スモールスタート・逐次改善していくこと」の大原則を基盤とし、様々な観点から教育データ利活用について具体的かつ多様な議論を展開している[5]。

前提言(B) 学習データを収集・利活用するための制度設計に関しては、文科省の教育データ利活用に関する有識者会議において、教育データの収集方法や管理方法に関する議論が進められている。令和5年3月には、初等中等教育の公立学校における教育データの取得や管理に関する留意事項がまとめられた「教育データの利活用に係る留意事項」³が文科省から公表された。また、近年の著作権法の改訂や SARTRAS による運用など、著作権に関する制度の社会への浸透も進んでいる。

前提言(C) 学習データを収集・利活用するための情報環境の整備に関しては、教育データ利活用の大前提である1人1台端末の情報環境や、その通信環境の整備は、小中学校のみであるがGIGAスクール構想によって劇的に進んだ。また、文部科学省ではMEXCBT・アンケート調査ツール(edusurvey)が開発されるなど、全国レベルでの教育のデジタル化が大幅に進んでいる。さらに、深層学習やChatGPTなどのビッグデータを用いたAI技術の教育利用も進展している。

前提言(D) 学習データを収集・利活用するための人材の育成に関しては、教育データ利活用を担う組織や人材も、国立教育政策研究所に教育データサイエンスセンターが設置され、現場の教員への研修が計画されている。教員養成の段階においても教師のICT活用指導力充実に向けた取組の中で教育データ利活用が紹介され始め、岡山大学においては教育データサイエンス学位プログラムが教育学研究科に新設される予定である。また、データサイエンス教育自体の重要性が広く浸透し、文科省の「数理・データサイエンス・AI教育認定制度」が着実に進められ、全国で学部や大学院の新設が行われ、教育未来創造会議による人材育成のための提言にも取り上げられている[6]。

2.2. 2020年提言において、まだ不十分なこと

2.2.1. 前提言(A) 学習データの種類と教育改善のための利用について

教育データ利活用は第一義的に学習者が最大の受益者であり、教師一人一人が主体的に

² <https://www.ederc.jp/>

³ https://www.mext.go.jp/a_menu/other/data_00007.htm

自らの力を教育に活かさなければならない。そのため、個人情報保護や、データによる推薦や提案による決め付けや押し付けの回避のような原則の元、自らに関わる教育データに安全・安心な形で自由にアクセスでき、最新・汎用的な技術による簡便かつ効果的な仕組みが図られなければならない。文科省やデジタル庁のロードマップにおいて、これらのような教育データの種類や利活用の目的の整理は進んだが、実行に移す点では不十分であると考えられるため、本分科会にて、引き続き考察を進める必要がある。教育データ利活用は本質的に困難の多い、社会的合意の取りにくい課題であることに変わりはなく、とりわけ根本的な問題については先延ばしにされているのが現状であり、2. 2. 2 (1)～(4)に示すような取組が必要であろう。

2. 2. 2. 前提言(B) 学習データを収集・利活用するための制度設計について

(1) 教育データへの当事者によるアクセスの確保

一般的に、公教育から得られるデータが特定の企業だけを利するべきでなく、データを提供した学校や自治体にデータを集約すべきである。残念ながら教育データを特定の企業で囲い込もうという動きはまだ多く、データを提供した学校にはデータが提供されない例も見られる。企業が参画しやすい環境が担保されながら、当事者である学習者（保護者）・教員がデータへアクセスし、解釈などについて吟味できる環境を維持するべきである。すなわち、まず、学校や自治体に教育データを集約して、一次利用・二次利用できるようにするべきである。

(2) 教育データ標準化の推進

文部科学省は教育データ標準として、教育データを主体情報、内容情報、活動情報に区分して順次策定が進められている[7]。このうち、どのような学習活動を行ったのかに関する活動情報については、多くの活動について標準化はまだなされていない。学習データの共有や適切な分析のためには、フォーマットや意味を統一するだけではなく、具体的にどのような項目をどのように記録するかを規定する必要がある。活動情報の標準的な技術規格として xAPI (Experience API)⁴や Caliper⁵等の利用が進んでいるが、これを元に標準化を行い、学習者や教員・学校にサービスを提供する企業がこの標準化に準拠して教育データを学校に提供できるような枠組みを構築することが重要である。特に、デジタル教科書の閲覧データにおける活動情報の標準化は重要な課題である。

(3) 教育データ収集のための国全体で統一したルール作り

大学 ICT 推進協議会 (AXIES) は、2020 年に提言[1]が公開されたことを踏まえて、大学等の高等教育機関による学習データの利活用を図るため、「教育・学習データ利活用ポリシー」のひな型を策定し公開している[8]。初等中等教育においても、個人情報の利活用を促進するために統一的なポリシーの策定が早急に必要とされる。初等中等教育における学習データは原則として学校などの各教育機関で管理することが望ましいが、現実的には地方自治体や教育委員会等の支援が必要となることも考えられる。学習データの管理者において、個人情報保護法や条例などに基づく適切な取り扱いができるよう、ポリシーを策定す

⁴ <https://adlnet.gov/projects/xapi/>

⁵ <https://www.msglobal.org/activity/caliper>

る必要がある。

(4) 教育データの二次利用

国や自治体における教育政策へのフィードバック、政策立案への活用、また研究利用のために教育データを共有し二次利用することのできる枠組みを構築することは重要である。現状は、国立教育政策研究所教育データサイエンスセンターが提供する公教育データプラットフォームや教育データ分析コンテストなどで教育データを提供するなど、進展は限定的であるが、これらはスモールステップで着実に進める事が求められる。

2.2.3. 前提言(C) 学習データを収集・利活用するための情報環境の整備について

GIGA スクール構想を通じて情報環境は大きく改善したものの、学校の無線 LAN が多数の同時接続に頑健でない場合が見受けられる。また、宿題などで端末を持ち帰る事例もみられるが、家庭環境によっては無線 LAN が整備されていない場合もあり、公平な教育環境を維持するための十分な支援が必要である。加えて、健康への配慮や端末を適切に使うことの指導についてのサポートは十分とは言えず、合理的配慮への対応も進んでいない。学校外部からの支援体制などをうまく作り、学校本来の教育活動に集中できるようなインクルーシブな情報環境を目指して改善するべきである。

2.2.4. 前提言(D) 学習データを収集・利活用するための人材の育成について

(1) 教員のデータ利活用能力の向上

学校現場の ICT 機器をめぐる環境はこの3年間で大きく改善しているが、教員のデータ利活用能力の向上にはまだほど遠いのが現状である。教員の ICT を用いた指導力向上自体が不十分であり、4人に1人程度は十分に自信を持っていない一方で、ICT 活用指導力に関する研修の受講率は75.8%に留まり、地域差も依然として大きい[9]。すべての高校生が学ぶようになった「情報 I」の一部の内容は、官民が提供するビデオ・教材などを積極的に活用して現役教員も身に付け、自らの指導に活かすべきである。また、10年後に教員になる人材のためにも、教員養成課程だけでなく高校・大学やそれ以前のカリキュラムを一層強化して、データ活用能力を高めるべきである。

(2) 教育現場と研究者の相互協力

教育データ利活用においては、教育学・心理学の素養とデータ科学の技術が必要であるため、これまで以上に教育現場と研究者の相互協力が重要となる。協力の形として、研究者が教育現場の一人としてデータの一次利用を行う場合と、複数の学校設置者から得られた匿名加工データを用いて二次利用の形が主に考えられる。また、初等中等教育の教員が大学にきて教育データサイエンスを学ぶことを通じ、共同研究に発展するような形も考えられる。

3. 2020年の提言に加えて推進すべきこと

3.1. 2020年提言(A) 学習データの種類と教育改善のための利用

本分科会の2020年の提言では、学習者が教育を受けたり、学習活動する中で生じる学習データを主な対象とし、その収集範囲を小学校・中学校・高等学校・大学・大学院などの公的教育機関を対象として始め、徐々に広げていくことを提案していた。社会のデジタル

化が加速度的に進む中、これらの問題について議論を始めるべきであると考え、本記録では、教員が教育を行う上で生じる教育に関するデータも教育データとして扱い、小中高大の前後である、乳幼児や社会人における教育データ利活用について、メリット等の考察を行った(参考資料1)。この他、学校教育以外の教育機関との連携や名寄せの問題など、3.2～3.3節で論じるようなより広い観点からも検討する必要がある。

表 学習データ・教育データの活用に関する2020年の提言と本記録の範囲

	2020年の提言	本記録
対象	公教育(初等中等教育と高等教育)	公教育に加えて幼児教育、社会人、私教育も対象
データの種類	主に学習者に関する学習データ	学習データに加えて教育データも対象
名寄せ	名寄せを行わない	名寄せを行う

3.1.1. 未就学児・小学校低学年

未就学児の就学前教育や保育の状況、たとえば既に広く流通している保育・育児関係のアプリや、保育所・保育園・幼稚園等における子供に関する活動記録を適切な範囲において、デジタル化との関係について議論を始めるべきである。発達・発育の状況における個人差が大きいという、フィードバックの内容自体を本人が理解出来ない場合がほとんどであり、基本的に保護者へフィードバックすることになる。その場合、どのような利点や問題点があるのか、実証実験や専門的な知見による検討を具体的に始めるべきである。これは、小学校低学年についての実証実験にも重要な知見を与えるだろう。

3.1.2. 社会人

リカレント教育の重要性に加え、近年、新しい職業に就くため、あるいは今の職業で必要とされるスキルの大幅な変化に適応するためのリスキリングの重要性が指摘されている。これら社会人や企業における学習についても、学習の履歴や学習によって得られたスキルについてのデータを用いることで、企業におけるニーズとのマッチングを行うような就労支援にも学習データ利活用を活用することができると考えられる。一方で、社会人の学習データは企業の利潤と結びついているケースが多いと考えられるため、データの収集や共有などにおいて、学校教育における学習データとは異なる制約が生じるケースが考えられる。また、社員教育については企業の規模が大きいほど取組やすいが、中小企業も恩恵を得ることができるよう、社会人教育における知見を得るための研究データの収集、フィードバックのあり方について検討の必要がある。

3.1.3. 公教育以外の機関との連携

公教育以外との連携についても、具体的な議論を始めるべきである。たとえば、学習者が同意し、学校における成績や学習状況を学習塾などへ提供することはあるが、教育データ利活用はその連携を容易にする。逆に、学習塾やオンライン教材等における学びのデータを学校に提供することで、学習者にとっての教育効果を高め、全体として学校の教育的

な機能を高める可能性がある。公教育以外との適切な連携を実現する制度設計(2.2.3)や、ログデータの流通を実現する情報環境(2.3.2)において議論し、公教育と教育関連企業の相互作用による新たなステージを模索し始めるべきである。一方、教育活動に直接影響しない生体情報やセンサーなどのデータは、前提言に引き続き教育データに含めるべきではない。しかし、研究目的に限って教育以外の分野との連携も含めて模索し続け、多種多様なデータを用いるメリットとともに、4章で論じるような諸問題との関連を注意深く議論するべきである。

3.2. 2020年提言(B) 学習データを収集・利活用するための制度設計について

データの名寄せについて技術・運用面において十分に検討し、その概念・意義と必要性を教育関係者に広く浸透することが求められる。

3.2.1. データの名寄せの必要性の浸透

各教育機関が(不可逆に)匿名化を行った場合、データの名寄せは不可能になり、機関を超えてのデータの連結ができないため、長期的な視点による各個人の状況に合わせた教育改善の取組は難しい。たとえば、ある市の教育委員会が中学3年生に対して長期的な視点で何らかの施策を取ったとき、その施策が適切な効果を生んだかを評価して次へ生かすためには、子どもたちを追跡調査することが必要である。しかし、たとえば各学校の「ユーザー情報」の部分を破棄してしまうと、長期的な視点から、施策がどのような子どもには適切・不適切であり、施策を公正に最適化するための知見を得ることは極めて困難になる。また、教育を受ける子どもの視点で言えば、自分自身に関する情報は、どのような範囲まで、どのように利用されることなら許可するかについて、コントロールする権利は本人にある。(本人がデータを確認でき、修正を依頼できる)たとえば、学校に集積された教育データを卒業後に削除依頼をするようなオプトアウトの可能性も残すべきであるし、逆に、学習履歴の振り返りや今後の学習計画、自己調整能力や自己主導能力の育成、生涯学習やキャリア形成等に活かすために自らの教育データや学修状況をエクスポートするような、データポータビリティを高める可能性も必要である。そのいずれもが、不可逆な匿名化では実行できない。4章で論じるような問題に配慮しつつ、データの名寄せを拒否する選択肢を残しながら、本人の意思を尊重してデータの名寄せを可能にする手段が用意されるべきである。さらに、2020年の提言にもあったように、間違った利用をしないように倫理審査委員会を設置して、安全・安心に適切に教育データを利活用していく枠組みを作ることが必要である。

3.2.2. IDの連携と管理体制について

上記の議論から、学習者のIDや個人情報情報を国の機関が一元管理すべきでないと考えられるが、教育に関するサービスを一括して提供する特定の企業が、結果的にID管理を独占的に担う事態も避けるべきである。公的なインフラとして学習者個人ID、公教育データプラットフォーム、私教育データなどを連携させ、データの名寄せを行う制度と仕組みを用意することは技術の進化ででき、学術的にも社会的にも意義がある。また、名寄せによる長

期的な学習データの収集と利活用を考慮したポリシーについても検討する必要がある。

3.2.3. データの適切な二次利用のためのポリシーの作成

教育データを異なる機関で安心して共有できるよう、教育データの二次利用に関するポリシーを策定し、個人が特定されないための匿名化等の処理を行うこと、データの利用目的を特定した上で、学校、地域の序列化などに利用しないことを事前に規定するべきである。例えば、ルールに違反した場合には、罰則を科すなどの規程を設けることも考える必要があるだろう。一方で、適切な企業との連携へのハードルを下げるため、学習者のデータを適切に取り扱うための「学習者プライバシー宣言」を作成するなどして、ログ連携を実現するための、適切な制度設計が必要である。

3.3. 2020年提言(C) 学習データを収集・利活用するための情報環境の整備について

3.3.1. デジタル教科書・教材のあり方

日本においては、デジタル教科書に関しては、数十年前から学校への導入が試みられているが、いまだに紙の教科書が日常的に利用されている。現状では、多くのデジタル教科書のビューアが乱立している状況であり、ユーザー認証やユーザーインターフェースの統一などの問題がある。それゆえ、デジタル教科書上で行われた学習活動のログデータに対する扱いについても議論が不十分である。デジタル教科書・教材の整備においては、教員や学習者の利便性を中心に考えて、ビューアや配信システムについては、様々な規格を乱立させず、標準化を行って統一の方向性を目指すことは、利用時の負担の面でも、独自システムのメンテナンスに伴うコストの面でも、ログデータの書式統一性の面でも重要である。また、紙の教科書とデジタル教科書の間における著作権の扱いや、教科書検定のあり方、デジタル教科書・教材の流通など、様々な問題を統合的に早急に解決しつつ、日々の教室内で起こっている教育・学習活動の改善に素早く活かせるように教育データを活用した仕組みを模索するべきである。そのうえで、教科書の内容の改訂にログデータを利用したり、学習者が何をどのように学んだかをフィードバックしたり、集められた教育データに対する最適な分析ツールを開発する[10]など、より広い視野での標準化や制度設計が用意されるべきであろう。

3.3.2. 教育ソフトウェアによるログデータのエコシステムの整備

個人の特定の知識・技能を伸ばすためのデジタル・ドリルや、より高次な能力の向上に繋げるためのアンケートを伴うサービスなどが急速に普及しているが、適切な利用許可の下で安全・安心してデータが収集される仕組みが実装され、適切な手続きを踏んだ教育関係の研究者・企業などが利活用できるような社会基盤を早急に整備するべきである。例えば、教育関連企業が提供する教育ソフトウェアが保存するログについては、今後、教育データの標準化に準拠して保存されると考えられるが、ログデータを、このような標準ログと、そのソフトウェアが独自に保存する独自ログに分け、標準ログについては学校にデータを集約して学校教育の改善に繋げていくべきであろう。一方で、このようにして、複数の教育ソフトウェアから学校に集約されたデータの内、各教育ソフトウェアが利用可能なものはこの利用についても学校標準ログとして教育ソフトウェアが使えるようにすること

も検討すべきであろう。

3.4. 2020年提言(D) 学習データを収集・利活用するための人材の育成について

3.4.1. 魅力的な学校現場につながるベストプラクティスの発信

近年、小中高における教師の負担の大きさが広く知られるようになる一方で、教員の年齢層の偏りに伴う急速な世代交代が進み、学校現場における実践知の継承が不十分な場合が生まれている。このような状況を打破するため、教育データ利活用だからこそ可能なベストプラクティスを共有することはますます求められている。さらに、事例報告を超え、エビデンスとして効果のある事例を共有することも重要である。従来ではやりたくてもできなかった授業を実現し、教育分野の魅力を増やして優秀な人材を惹きつけることを社会全体で目指すべきであり、優れた取組は、学習者のデータやAIに対するリテラシーをも高めるであろう。これらの事例には、教員の負担軽減、働き方改革といった従来の課題に対する解決策も含まれる。

3.4.2. デジタル技術を活用した社会の実現に向けて

ChatGPTなどのAI・ビッグデータ技術を教育に適切に活用することを推進することは重要である。特に、教育に限らず、データ利活用において重要なことのひとつが、データやAIによる判断に対する「人間の主体性」であり、そのためのデータを多面的に吟味する力である。このような力の必要性は、学習指導要領においても、主に算数・数学の「データの活用」において育成が求められている。加えて、データを多角的に吟味して判断する力の育成が、現代科学の発展において不可欠であった事を考慮すると、科学的態度の育成に直結する可能性がある。学習者にとって、自らの学習データを用いて利活用の試行錯誤を行うことは、社会に出てからのいろいろなデータ利活用の環境を生き抜く糧になるであろう。データサイエンスの発展に繋がる人材・データ・技術・制度を一体化したエコシステムを、広い視点で教育的に配慮して構築する視点が必要である。このような視点は、自然科学に偏る事なく人文・社会科学においても拡がるべきであり、また、一般化や抽象化への志向だけでなく、個々の事象を扱うような臨床的な志向も包摂して議論されるべきである。

4. 2020年の提言に加えて、注意すべきこと

2020年の提言では、既存の個人情報保護法等の法令、ポリシーやガイドラインの範囲内で作成したが、社会におけるデータ利活用の拡がりに伴い、既存の枠組みを超えた広い観点から議論する必要性が高まっている。たとえば、様々な分野における研究開発の社会実装時に生じる課題であるELSI (Ethical, Legal and Social Issues: 倫理的・法的・社会的課題) は、学習データの利活用においても重要な観点である。JST-RISTEX『科学技術の倫理的・法制度的・社会的課題 (ELSI) への包括的実践研究開発プログラム』プロジェクト企画調査「学習データ利活用 Edtech (エドテック) のELSI論点の検討」では、日本型公教育における教育制度・仕組みと学習データ利活用技術におけるELSIの論点をまとめている[11]。

海外におけるプライバシー法とポリシーに関しては、現在、世界各国の個人情報やプライバシー保護に関する法規制の基本原則として、OECDプライバシーガイドラインにおける

8原則が取り入れられている。EU では、2018年にEU一般データ保護規則（General Data Protection Regulation: GDPR）が施行され、個人情報の取得における同意などについて厳しいルールが定められている。米国では、家族教育権とプライバシー法（Family Educational Rights and Privacy Act: FERPA）によって、学生の教育情報へのアクセスや開示に関する権利について公的な教育機関が守るべきルールについて規定している。2020年には、教育用アプリケーションにおけるプライバシーやその他の安全性の審査を統一的な評価で行うための仕様であるIMS Global 1EdTech Data Privacyが策定されている。

国内では、AXIESが高等教育機関における「教育・学習データ利活用ポリシーのひな型」を策定している。また、GIGAスクール構想推進委員会は、学習者プライバシー宣言書を策定している。これは学習サービスの提供企業が、学習者の個人情報保護や個人プロフィール等の取り扱いについて宣言するものである。

憲法学の観点からは、特に教育を受ける権利、プライバシー権、教育の自由についての論点が提示されている[12]。教育を受ける権利に関して、学習データを活用した教育の個別最適化は重要である。他方、学習者の学習権の側面から、その提案内容を本人の意思に優越させるべきではない。分析上は最適な選択でない場合であっても、そのような選択をする権利を保障していくことが必要である。プライバシー権については、学習者の同意の扱いや同意しなかった場合の不利益への対処などの論点がある。教育の自由の観点からは、国や企業が学習データ利活用技術の提供を通して教育内容の決定に深く関わっていく可能性がある。これらの技術を教師が利活用できるよう支援していくことが重要である。

以上の点などを踏まえ、今後、学習データ利活用技術において注意すべき点を以下に挙げる。

4.1. データ利活用における透明性の確保

学習データの利用の際には、データの収集・保存・処理の方法等について、学習者への透明性を確保する必要がある[13]。個人情報保護法では、本人が個人情報の取り扱いについて合理的に予測・想定できるように利用目的を特定する必要があるとあり、個人情報保護委員会が事例を挙げている[14]。学習データを利用した推薦や提案においては、学習者が自ら学習行動を選択するために必要な情報を提供することが必要であると考えられる。この際、学習データの分析方法やアルゴリズム、得られた推薦情報の解釈方法などを、どのように学習者に提示すべきか今後さらに検討を進めていく必要がある。また、教育データ分析に利用するAIの説明性や解釈性に関する研究を進めることも重要である。

データバイアスの観点からは、学習データが適切な方法で偏りなく収集されているか、またデータや分析結果の解釈の方法に偏りが存在しているかについて、教員や学習者が確認することのできるよう、透明性を確保する必要がある。

4.2. 利用者のリテラシー教育

教員や学習者が、上記のような学習データの分析における利用目的やフィードバックを理解し、それを活かした学習につなげるためには、利用者に対してデータ活用に関するリテラシー教育を行うような体制を構築する必要がある。リテラシー教育による学習データの分析結果への理解の下で、データに基づいて提案された教育が学習者本人の望むもので

あるかを判断し、学習者の意思によって選択をできることが望ましいと考えられる。

4.3. AIの規制

学習データの分析によって得られた結果をどのように学習者にフィードバックするべきかについても、検討の必要がある。EU 一般データ保護規則 (General Data Protection Regulation: GDPR)⁶ では、自動化された取扱いのみに基づく判断を受けない権利を定めており、データ主体に大きな影響を与えるような場合には人間による判断を介在させる必要性を求めている。日本の法令にはこのような規定はないが、進路の選択や決定、学習状況の診断などの重大な学習データの分析結果のフィードバックにおいて、完全に自動化された場合の問題点について検討する必要がある。特に、アルゴリズムに監査や認証が必要とするようなAI規制⁷の原則は整備されつつあるが、教育分野特有の問題に対処できる原則も検討を進めるべきである。

4.4. 生涯にわたる教育データ利活用について

リカレント教育やリスキリングの必要性の高まりなどを考慮して、公的教育機関より外へ教育データ利活用の議論を広げた時、生涯に亘る教育データ利活用がどのような危険や可能性に繋がるのか、倫理的な観点も含めて検討を始めるべきである。データを生涯に亘って連結することには学習履歴の振り返りや証明、学習計画の立案など様々な利点がある一方で、学習者がそれまでの履歴をリセットして再スタートすることを望むようなケースや、教師に不要な先入観を与えるような履歴が提供される可能性も考えられる。このようなケースでは、学習者はデータ項目によっては学習履歴を提供しないという判断ができる機会が与えられるべきである。また、働く方の自発的なキャリア開発や、再就職のための社会人の学び直しのためのリカレント教育は、働き方改革を推進していくためにも重要であるが、そのような個別の問題を一つなぎにした時、全体としてサステナブルな人間社会のシステムになっているべきである。

5. まとめ

教育データ利活用は、様々な分野が絡む社会的に合意を取りにくい課題である。COVID-19による社会の激変を受けてもなお、手付かずな根本的な問題は多いうえに、近年の劇的なAIの進化は、教育の分野を超えてデータ利活用そのもののあり方を揺るがしている。しかし、すべての人がなんらかの教育を経て大人になる以上、数十年後のデジタル社会に向け、教育の分野に適切なDXを実現していくことは社会全体の責任であろう。そのための人材を育成しながら、分野や立場を超え、制度設計を行い、情報環境を整備し、技術的にも社会的にも合意を形作っていくことが求められる。

⁶ <https://gdpr-info.eu/>

⁷ <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

参考文献

- [1] 日本学術会議「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—」(2020)
- [2] 九州大学ラーニングアナリティクスセンター<https://la.kyushu-u.ac.jp/>
- [3] デジタル庁 (2022)「教育データ利活用ロードマップ」
https://www.digital.go.jp/posts/a5F_DVWd
- [4] 文部科学省：教育データの利活用に係る論点整理 (中間まとめ) (2021)
https://www.mext.go.jp/b_menu/shingi/chousa/shotou/158/mext_00001.html
- [5] 緒方、江口：学びを変えるラーニングアナリティクス(2023)日経 BP 出版
- [6] 教育未来創造会議「我が国の未来をけん引する大学等と社会の在り方について (第一次提言)」(2022)<https://www.cas.go.jp/jp/seisaku/kyouikumirai/teigen.html>
- [7] 文部科学省「教育データ標準」https://www.mext.go.jp/a_menu/other/data_00001.htm
- [8] 大学 ICT 推進協議会 (AXIES) 学術・教育コンテンツ共有流通部会 (2020)『「教育・学習データ利活用ポリシー」のひな型の策定について』<2023年1月に改訂>
- [9] 文部科学省 (2021)「令和3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」
- [10] OpenLA, <https://limu.ait.kyushu-u.ac.jp/~openLA/>
- [11] 「**学習データ利活用 EdTech の ELSI 論点の検討**」プロジェクト、EdTech (エドテック) ELSI 論点 101, https://elsi.osaka-u.ac.jp/system/wp-content/uploads/2022/01/20220201_EdTech_ELSI101.pdf
- [12] 堀口悟郎、連載『デジタル社会と憲法』第2回 EdTech と子どもの権利、法学館 憲法研究所、https://www.jicl.jp/articles/topics_digital_202206020.html
- [13] 個人情報保護委員会、個人情報の保護に関する法律についてのガイドライン に関する Q & A、https://www.ppc.go.jp/personalinfo/faq/APPI_QA/#q2-1
- [14] 個人情報保護委員会、GDPR (General Data Protection Regulation : 一般データ保護規則)、<https://www.ppc.go.jp/enforcement/infoprovision/laws/GDPR/>

参考資料1：乳幼児や社会人における教育データ利活用について

		乳幼児への利活用	社会人への利活用
一 次 利 用	メリット	フィードバックを得て子どもへの理解を深め、今後の家庭での教育に活かしたり、保育・幼児教育施設での個人面談での活用などが考えられるが、原理的には個人の発育状況の確認であり、他者との比較や細かい数値のフィードバックは解釈が難しく、メリットがない。また、収集自体に対する異論も考えられる他、厚生労働省関係の保育・健康関係のデータと、文部科学省関係の教育関係のデータに分散しており、取り扱いが困難ではないか。	ある程度の大きさの企業であれば、社員教育におけるポリシーを定め、何らかの学習に対してフィードバックし、以後の最適な学習計画や、学修成果の証明に使えるかもしれないが、中小零細企業を含めると実施は難しいと考える。メリットは組織でのキャリアではなく、個人の職能を中心にキャリアを考えることが可能になる。
	集めるデータ	保育所児童保育要録や幼稚園幼児指導要録の元になるデータなど、言葉、表現等に関する指導の展開と子どもの育ちなど。また、保育者の評価による子どもの心身の発達に関するデータ。乳幼児期は発達の個人差が大きいと考えられるので、個人へフィードバックする場合は十分な配慮と説明が必要。	個人の持つ資格や研修の受講歴、民間テスト(TOEFL など)のデータ、リカレント教育の受講時の成績、職能を活用できる部署の在籍年限など
	同意の必要性	保護者に対して、opt-in も必要。opt-out も認める。	opt-in も必要。opt-out も認める。
	個人情報管理	原則教育機関で管理すべきと思われるが、現実的には地方自治体等の支援は不可欠か	当該機関で管理する
	家庭状況の有無	子どもに関する理解に活用できるなど、理想的には含めるべきだが、どの程度の粒度で調べるかは議論が必要。家庭状況は、ある程度は各施設で既に把握されている。	家庭状況は収集不要と思われるが、信頼性と説明責任を担える企業は含めてもよい。

	フィードバックの必要性	フィードバック内容の意味や意義や対応方法などの丁寧な説明が必要であり、保護者のメンタルへの配慮が必要。	フィードバック内容の意味や意義や対応方法などの丁寧な説明が必要であり、本人への伝え方に配慮して、動機づけを高め、成長につながるように伝える。本人には伝えない情報（選択や選抜、研究にかかわるデータ）もありえる。また、心理面への配慮が必要であり、インタラクションのデザインは重要。
		乳幼児への利活用	社会人への利活用
二次利用	メリット	教育研究のための基礎データとしてや、国や自治体における教育政策の効果のフィードバック等に活かし、真の乳幼児教育（保育）向上につながる可能性がある。ただし、教育組織間、地域間の差異を確認して、地域環境、教育プロセスの改善に使う場合も、序列化に使われないように注意する必要。	教育研究のための基礎データとして利用し、生涯学習に関する実態の把握と政策への活用や、職能を生かす働き方や人材育成に関するに利用できる可能性がある。
	同意の必要性	匿名化なので同意不要/匿名化利用にも同意必要の両意見有り。一方、opt-out について、行政実施の場合は認めない。研究機関実施の場合は認める、という可能性もありうる。	匿名化なので同意不要/匿名化利用にも同意必要の両意見有り。
	匿名化について	名寄せし、最新の匿名化技術で行える。	名寄せし、最新の匿名化技術で行える。
	複数のデータを連結する方法	全員のデータは必要ないが、個人が自らの学習履歴を個人利用するための許諾、研究者がおこなう長期縦断研究のための許諾を経る場合が考えられる	全員のデータは必要ないが、個人が自らの学習履歴を個人利用するための許諾、研究者がおこなう長期縦断研究のための許諾を経る場合が考えられる

<p>フィードバックの必要性</p>	<p>乳幼児教育（保育）施設に入って教育データを分析して、個人へフィードバックが行えるのはミクロな研究に限られ、データを保護者にフィードバックすることは容易ではない。全体的な分析としてや、教諭に伝えるところがスタートになるだろう。</p>	<p>社会人の学習の機関に入って、教育データを分析して、機関、教員、学習者にフィードバックするし、企業との共同研究はありえるが、企業が要望しない限り、個人にフィードバックするのは難しいのではないかと考える。全体的な分析結果をフィードバックすることは可能な場合もあるかもしれない。</p>
--------------------	---	---