

連合研究科共同研究プロジェクト研究成果報告書

プロジェクト の名称	子どもの意欲と学力を向上させる教育ビッグデータの利活用ネットワークの形成		
研究期間	平成30年4月1日～令和3年3月31日	プロジェクト記 号	V
チーム構成員の氏名・職名等・所属（配属）大学（◎：チームリーダー）			
◎寺澤 孝文（教授・岡山大学）	大場 浩正（教授・上越教育大学）		
青木多寿子（教授・岡山大学）	Hiroyo Nishimura（Senior Lecturer・Yale University）		
山田 剛史（教授・岡山大学／現横浜市立大学）	Noriko Mori-Kolbe（Lecturer・Georgia Southern University）		
Scott Willard GARDNER（教授・岡山大学）	濱口 和弥（課長・徳島県立総合教育センター教育情報課）		
皆川 直凡（教授・鳴門教育大学）	清水 衆（局長補佐・長野県高森町教育委員会事務局）		
山崎 勝之（教授・鳴門教育大学）	益岡 都萌（特任助教・岡山大学）		
川上 綾子（教授・鳴門教育大学）	川崎 由花（教授・兵庫教育大学）		
越 良子（教授・上越教育大学）	澤山 郁夫（助教・兵庫教育大学）		
大場 浩正（教授・上越教育大学）			
プロジェクト全体の研究経過及び研究成果			
<p>・3年間における研究活動（研究会開催、調査活動、研究成果の公表等）の概況や個々の分担研究の特筆すべき事項等について、簡潔に箇条書きでお書きください。</p> <p>【背景】</p> <p>ICTの進展と情報端末の普及により、個人にヒモづけできる縦断的履歴データが大量に収集できる状況が生まれている。現在ほとんどの調査研究で収集されている横断データに比して、因果関係を議論できる情報を含む縦断データの価値は社会科学の領域において非常に大きい。その縦断データが大量に集約される状況は、社会科学全体の研究方法を根底から変える可能性を持っている。この状況は、文科省のギガスクール構想による一人一台端末の配備により加速される可能性もある。</p> <p>しかし、実のところ、縦断データから有益な情報を抽出することは容易なことではない。なぜなら、縦断データには、これまで社会科学で扱われてこなかった、「いつ」という時間次元の要因の影響が新たに含まれることになり、かつ、人間の行動に影響する様々な要因の中で、時間次元の要因の影響は想像以上に大きいためである。例えば、ある子どもが、明日の試験で、ある問題に解答できる確率を推定する場合、今日その問題をこの子が勉強していれば、正答できる確率は高くなるが、1か月前に勉強していた場合は低くなる。さらに1か月間勉強していない場合でも、1年前から何度も学習をしていた場合は、正答率は高くなる。そして「いつ」という条件は無数想定できるため、単純に収集された学習ログを解析すれば、「いつ」に起因する大量の誤差が生み出されることになり、精度の高い推定は原理的に難しい。</p> <p>それに対して、教育の領域では学習者ごと、学習内容ごと、学習やテスト等のイベントの生起タイミングを操作できる状況がある。それを年単位で緩やかに制御し、学習とテストのタイミングやインターバルを連続的にコントロールし、縦断データを一元的に集約する技術（マイクロステップ・スケジューリング法）が確立され、時間軸条件がそろった大量の縦断データ（以下、高精度教育ビッグデータと呼ぶ）が収集できるようになった。</p> <p>さらに、高精度教育ビッグデータの解析により、英単語や漢字など、日々のわずかな学習効果の積み重ねを学習者ごとに可視化できるようになり、漢字や英単語のドリルでは、ほぼ全ての学習者の成績が直線的に上昇することが明らかになった。さらに、その結果を個別にフィードバックすることも実現され、それにより、特に意欲を失っていた児童生徒の意欲を劇的に向上させられる状況が生まれた。</p> <p>本プロジェクトは、個別フィードバックを基本とした新たなeラーニング（マイクロステップ・スタディ）の社会実装の拡大と、そこで収集される高精度な縦断的学習データ（以下、高精度教育ビッグデータと呼ぶ）の学術利用を促進することを目的とした。</p> <p>具体的な目的として、①社会実装（費用を学校等が負担する形での導入）を拡大し、対象となる学習者を増やしていくためのネットワークを作ること、②収集される学習データと意欲のデータから新たな知見を見出すこと、③eラーニングの最後に回答を促す心理尺度項目をわずかずスケジュールリングすることで、研究者から提供を受ける心理尺度について年単位で縦断データを収集し、④そのデータを完全無作為化の処理を行った上で研究者に提供する、データの利活用ネットワークを構築することを目指した。</p>			

以下、それぞれの目的に対応させ、それぞれの成果を説明する。

(1)社会実装の拡大

- ・当初数百人の小学生と高校生を対象としていた支援は、2021 年度当初で約 2,200 人の小中高生に広がり、それに加えて岡山大学では 2019 年度に課外学習支援システムとして正式採用され、1 年生全員に提供されることになり、一部の授業で学習量が成績に反映されることになった。その他の大学でも一部導入されるなど、プロジェクト期間で導入が広がり、2021 年度には、年間で約 1 万人対象にeラーニングを提供し始めることができた(プロジェクトの研究期間の中では、年度ごとに区切ると、のべ 1 万人の学習者に支援を提供)。
- ・特記事項としては、岡山大学での導入事例に関して、民間企業で評価の高い日本 e-learning アワードに応募した結果、「文部科学大臣賞」の受賞(2019 年 11 月)を得た。
- ・英語教育のメンバーの協力を得て、小学校英語で習得が推奨される英単語と中学校の教科書で出てくる英単語について、音声データを含むコンテンツを作成し、スケジュールリングしたコンテンツを開発し、小学校におけるeラーニング教材として実際に提供し、学習データを収集した。また、米国のメンバーの協力を得て、TOEIC 対策用コンテンツの出版社とコンタクトをとり、教材の提供を受け、スケジュールリングを開始した(実際のデータは期間内には収集できなかった)。

(2)収集される学習データと意欲のデータから見いだされた新たな知見

- ・2019 年度の岡山大学の 1 年生(約 2300 人)のデータ解析を行ったところ、語彙力の低い学生にとって、有意な語彙力の向上が認められた他、その結果のフィードバックを受けることにより、内発的動機付けが有意に向上することも明らかになった(岡山大学,2020)。
- ・マイクロステップ・スタディを提供している高校と小学校で収集されている意識データ(進路選択自己効力感尺度、動機づけ尺度等)の分析により、フィードバックに対応して、自己効力感が上昇する傾向や動機づけが有意に向上する等望ましい結果が得られた(平松,2019; 羊,2019; 今井,2020; 寺澤,2021)。
- ・2020 年度の岡山大学の 1 年生の英単語学習の学習量と総合的英語力の指標となる GTEC の得点(リスニングとリーディングの総得点)に有意な関係が検出され、マイクロステップ・スタディによる語彙習得学習が総合的英語力の成績に効果を持つこと等、非常に興味深い結果が得られた(山本・益岡・宮崎・寺澤,2021[発表予定])。英語力の向上には語彙力が重要であるとしばしばいわれているが、語彙習得に特化したeラーニングが GTEC の得点に明確な効果を持つことを検証した研究は初めてであり、特筆できる成果といえる(2021 年 8 月プレスリリース予定)。

・ネットトラブル、通信費等の問題を解決し自宅学習を高度化する教育専用端末の導入

GIGA スクール端末が配備されたことで、eラーニング環境が急激に整うことになったが、端末を自宅に持ち帰ることにより、子どもがネットトラブルに遭遇する問題や通信費の負担といった新たな問題が生じるため、自宅に端末を持ち帰らない自治体がほとんどとなっている。一方で、これまでの研究の知見から、自宅に情報端末を持ち帰る等の手立てを講じ、自宅学習を高度化せず、学校にいる時間だけで漢字や英単語などの知識の完全習得を実現することは、実質的に不可能であり、逆に自宅学習を高度化することで全ての子どもの知識習得を完全習得のレベルまで支援することが可能であることが明らかになっている。

具体的には、小学校 5 年の頃になると漢字の学習をあきらめ始める児童が 1 クラスに 1 人弱程度出てくる可能性が高い。その子にマイクロステップ・スタディを紙媒体で提供すると、他の児童とさほど変わらないスピードで漢字ドリルの成績が上昇する結果が出てくる。その結果をフィードバックすることでその児童は驚くほど学習を主体的に継続し、一定数の漢字を完全に習得できる事実が得られる。その結果からすると意欲を失った原因は、単に学習を継続できなかったためであり、フィードバックにより意欲を高めることで、完全習得までサポートできることは明らかである。しかし、このような結果は、自宅に持ち帰られる、紙媒体を使ったマイクロステップ・スタディで得られた結果であり、学校にいる時間のみでは完全習得は不可能であることも逆に明らかである。このような背景から、自宅学習を高度化する上で、自宅に端末を持ち帰ることで生まれてくる種々の問題を解決するシステムを考案し、それを別予算により端末に実装した。

その端末は、小学校低学年の子どもでも自宅と学校を容易に持ち運べ、体験学習や総合の授業などで映像や音声も容易に記録できるサイズで、将来的には家庭が文房具と同様に購入できる程度に安価(1 万円台)なスマホ端末とし、学校では校内の Wi-Fi により教師の管理下で様々なアプリが利用できるが、学校を離れると、ネットトラブルに遭わず、ネットの映像なども自由に見ることができないように画面が自動で切り替わる、エリア依存処理技術(寺澤,2020)を実装した端末である。さらに、クラウド SIM 機能という、SIM カードと同様の機能をソフトウェア的に実現する機能を持たせることで、自宅に Wi-Fi 環境がない家庭の子どももeラーニングが提供でき、さらに、知識習得に特化した自宅学習に特化させることができるため、わずかな通信量で学習が可能になるため、格安の通信料金(95 円/1 月・1 台)で自宅学習を高度化することが可能になった。

このエリア依存処理システムの開発と端末予算を別予算で確保し、岡山大学附属中学校の他、2 つの自治体のマイクロステップ・スタディに導入した。その結果、附属中学校では、マイクロステップ・スタディを導入している学校等でなされた学習データの 2 倍を超える学習データが収集された。そのデータの増加に解析システムとフィードバックシステムを対応させることが難しく、膨大な時間を費やしたが、安定してサービスを提供できる状況に近づいている。また、後述する意識データも想定を超える量のデータが手に入ったが、その解析も従来の方法では有意義な知見を抽出することが難しい問題も明らかになってきたなど、今後の課題が明らかになった。

・教育専用端末とゲーミフィケーションの導入による経済格差と教育格差の負の連鎖を断ち切る実践の成果

上記自宅学習を高度化するための教育専用端末を、児童養護施設の子どもの他、民間の塾に通えない子ども向けに自治体が開く公設塾(鳥取県倉吉市未来塾)に導入し、さらに、岡山大学教育学研究科の修士課程の PBL(project based learning)グループが主体となり開発した、ゲーミフィケーションの要素を加えたフィードバックシステム(学習の森シ

システム)を導入した結果、フィードバックにより成績の上昇が可視化されるまでの学習期間において、意欲が有意に向上する結果が得られた(山本,2021)。学習の森システムは、学習量の増加に伴い、一人ひとりの子どものフィードバック画面に木の芽が出て、葉が出て、実がなるような変化をフィードバックできるシステムであり、クラスごとに学習量の増加に対応して森が大きくなっていく特徴も持つ。マイクロステップ・スタディでは成績の上昇が出てくるまでに時間がかかることが課題であった。このシステムによりeラーニング開始当初の意欲の向上が可能になった。この成果は、家庭の経済状況などに依存せず、全ての子どもの知識習得を完全サポートするしくみにつながっていきと考えている。すなわち、経済格差と教育格差の負の連鎖を断ち切るしくみづくりにつながる成果といえる。

・高精度教育ビッグデータに関しては多数のメディアで取材がなされた他、IEEE 等での招待講演、書籍の出版もなされた。

・AI をはじめとするデータサイエンスは、扱えるデータがなければ大きな展開は期待できない。それに対して、本プロジェクトで集約され始めた高精度教育ビッグデータは、桁外れにデータ量が多く、時間軸の条件がそろった非常に質の高い特徴を持っている。本プロジェクトは、新たな技術を用いたeラーニングの支援拡大を目指してきたが、その取り組みとデータサイエンスの拠点づくりは表裏一体の特徴を持つ。すなわち、2021 年度の時点で、約 1 万人の学習者の質の高い縦断データが、年間を通じて集約できる状況が構築された。「この規模の高精度な学習データが集約できている拠点は世界的にもない(京大の研究者の言葉)」のは事実である。

・個別最適化処理の実装：形成的評価の自動化

学習者の学習状況に応じて指導法や学習内容を変えていくことが個別最適化といわれているが、それを実施する上で困難となるのは、学習者の学習状況を詳細に、また正確に把握することにある。例えば、ある学習者の英語の語彙力が低いことが分かったとしても、何千もある英単語のうちどれが弱いのがわからなければ、適切な指導は提供できない。さらに、ある英単語の習得度をテストで推定する場合、その英単語を、たまたま直前に学習していた場合成績は一時的に高くなる。その結果を受けてその英単語を学習しなくてもよいと判定してしまえば、取り返しのつかないことになる。以前から重要とされてきた形成的評価も、学習内容の量と実力の推定精度の問題を解決できなければ、実質的な評価にはならず、個別最適化も高い精度で実施できない。特に、形成的評価の自動化には、個々の学習内容に関するテストの精度向上が必須となる。

それに対して、マイクロステップ・スタディでは、テスト項目の一つひとつについて、テスト前のようなタイミングで学習し、それからどのくらいのインターバルをあけてテストを実施するのかをコントロールできる。それにより、何百という学習内容の一つひとつについて成績を定点観測できるようスケジュールを組むことで、学習から比較的長いインターバルをあけ、実力を測定することが可能となり、その複数の得点の時系列変化データから、その学習内容(例えば1つの英単語)の習得度を正確に推定することを可能にした。本プロジェクトでは、一つひとつの学習内容の成績の解析結果を基に、実力レベルで最高点を超えたと判定された学習内容を特定し、学習から排除していくことを可能とした(個別最適化機能の実装)。さらに、その結果から実力レベルで習得が必要な内容が何個残っているのかを学習者にフィードバックする仕組みも導入した。この個別最適化処理は 2020 年に全支援対象者に提供を開始した他、検証実験も開始したが、新型コロナ禍により学校等でのeラーニングの導入がストップするなど当初の計画と大きくずれたため、十分なデータが蓄積できなかった。

ただし、新型コロナ等の影響や、対象群の設定ができていないため明言することはできないが、個別最適化そりを実装した 2020 年度の岡山大学 1 年生対象のマイクロステップ・スタディでは、収集された学習データの量が 2019 年度に比べて 2 倍程度に増加した。

(3)心理尺度項目をわずかずスケジュールリングすることで、研究者から提供を受けた心理尺度について年単位で縦断データを収集することが可能に

・日々のeラーニングの最後に 4 問程度の心理尺度項目をスケジュールリングし、回答を求めることで、様々な心理尺度について、縦断データを個別に収集できるしくみを精緻化し、年単位で縦断データを集約することを可能にした。初年度はほとんどの支援対象校で既にeラーニングが開始されていたため、一部の対象校に限定し質問項目の提示順序をランダムにする方法と研究者が指定した順序で提示する方法をシステムに実装した。

・当初想定していなかった問題や事実も明らかになった。すなわち、現在心理学の分野において、横断的調査研究で収集されている心理尺度構成では、信頼性の高い尺度、すなわち複数回繰り返し調査データを収集した場合に因子構造などに変動が少ないことが求められる。しかし、一般の学校現場で子どもごとの意識変動に注意を払う場合には、逆に、微妙な意識変動をとらえる尺度が求められる。つまり、変動が少ない尺度は、現場の要望に応えられない可能性がある。具体的には、自尊感情尺度系の尺度ではフィードバックに対応した変動が検出できないことが明らかになってきた。一方、学芸大式学習意欲検査(簡易版)の尺度の自主的学習態度、達成志向の因子等ではフィードバックに対応し望ましい方向で有意な変化が検出されやすかったことが明らかになってきた。この事実は、心理学における尺度構成の方法に一石を投じることが予想される。

・本プロジェクト以前の研究で、抑うつ傾向を測定する尺度で縦断的に抑うつ傾向を測定した結果、上記と同様、ほとんどの中学生においては抑うつ傾向の変動が見られないことが明らかになってきた。一方で、抑うつ傾向の変動が大きな生徒を特定することが可能になることも明らかになってきた(矢地・寺澤,2012)。つまり、変動量を指標にすることで危機的状況にある生徒を検知できる可能性があった。この知見の再現性を大規模に確認することを計画したが、eラーニング自体は自治体単位で導入を検討していただく必要があり、教育委員会等がその導入の必要性を認識していただく必要がある。比較的ナイーブな質問項目がある抑うつ尺度の導入に対しては自治体単位での導入は困難であることが明らかになった。そのため、本プロジェクトでは抑うつ傾向については、縦断データの収集はできなかった。大規模な縦断的意識調査を実施するための新たな

課題が明らかになったといえる。抑うつ傾向の変動と共変動しやすいタイプでない複数の尺度で危機的状況の子どもを検知する方法等を開発する必要がある。

- ・学校等では、学習意欲の変動に関心が高いことが明らかになってきた。学年が変わっても継続して意識変動を把握するため、本プロジェクトで導入する心理尺度は、年度ごとに新たに導入が始められる学校とクラスなどに限定されることになった。また、上述したような変動量の少ない尺度は別の尺度に入れ替える他、逆に、変動が出て学校等が必要とする尺度は継続してデータを収集する必要が出てきたため、当初の計画通りには心理尺度の導入はできなかった。さらに、2019年度より新型コロナの影響で、自治体の予算編成でマイクロステップ・スタディの導入が停止された自治体が出た他、学校での新規導入をストップせざるを得ない状況や、業者が入校禁止になったためサーバの修理ができない時期が3か月以上続き、フィードバックが2020年度の夏休み明けまでできない状況が生まれた。これにより、正常な支援の提供や新規にeラーニングが導入される学校でデータ収集が半年以上遅れることになった。
- ・上記のような新型コロナの影響や、計画的に尺度を導入することが予想以上に困難であったが、以下にあげる大部分の尺度については、半年以上にわたって縦断データを収集した。なお、学校等によってはeラーニングを導入したものの生徒に学習を促す指導をされない学校等も存在した。そのため、一度も学習をしなかった生徒が多い高校などで研究に供せないものもある。

学芸大式学習意欲検査(簡易版)(心理尺度集)

特性的自己効力感尺度(成田ら1995)

中学生用自律的動機尺度:西村・河村・櫻井(2011)+上田(岡山大学、大学生用に改編)

英語学習行動:島田・鈴木・田中(2018)+上田(岡山大学)

英語および英語学習に対する信念:島田・鈴木・田中(2018)+上田(努力信念因子)

学習意欲尺度(皆川)

学習方略尺度(福谷・皆川)

日本語版時間的焦点尺度:Chishima(2017)+越:中学生向けに修正

邦訳版 AGQ-R: Tanaka & Murayama(2019)+越:中学生向けに修正

DIDS-J: 中間・杉村・畑野・溝上・都築(2014)

児童生徒用品格尺度(青木)

主観的な感覚としての人格特性的自己効力感尺度(SMSGSE)

ホープ尺度

生活充実感尺度

日本語版 IPANAT(下田他, 2014)+山崎

本来感尺度(伊藤・小玉,2005)

メタ認知的方略(市原)

児童用コンピテンス尺度:自己価値(桜井,1999)

- ・心理尺度データに関しては、欠損値が出てきた場合の対処方法の検討も難しい課題になることが明らかになった他、学習をしないため、データが収集できない児童生徒も少なからず存在するため、収集されるデータには意欲の高い児童生徒のデータが多く含まれる可能性も高くなるなど、データの偏りが出てくるのが新たな問題として浮上してきた。従来の横断研究では個人差を無視した平均的なデータ解析が行われてきたが、長い期間にわたり縦断的に学習・意識データを収集する状況では、欠損データの扱いや、人間の行動特性の影響を考慮した解析を行う必要が出てくる。現時点で利用可能な解析法としては、マルチレベル分析の活用が妥当であると考えられるが、今後新たな縦断データ解析法の開発が求められるのは間違いない。

(4)収集される縦断データに完全無作為化の処理を施した上で研究者に提供する、データの利活用ネットワークを構築。

- ・国内最大規模の学習・意識データを集約できるプラットフォームの拡大

近年、AIを始めとするデータサイエンスの研究が注目されているが、そこで共通の課題となり始めているのは、データがないことである。現在注目されているAIの技術は、人間と同様、様々な情報を学習(経験)することにより成長していく原理に従っている。つまり、学習させるデータがなければAIは作れないのである。GAFA(Google、Amazon、Facebook、Apple)が、世界のICTを牽引している理由は、それぞれの企業が、一般個人に大きなメリットを提供できるサービスを持っていることにある。それに比して、日本にはそのようなビッグデータを集約できるサービスが存在しない点が最大のネックとなっている。それに対して本プロジェクトは、一般の学習者に自覚できない学習効果の積み重ねをフィードバックすることでメリットを提供し、それにより個人ごとの膨大な学習ログと意識データを集約できるプラットフォームを、さらに拡大することを目指してきた。収集される学習と意識の縦断データは国内では最大規模といえる。また、「現時点で世界の中でもこのようなシステムはない(京都大学の研究者)」と言われていた他、東京大学エドテック連携研究機構の研究会と京都大学の同様の外部団体への参画が求められた。これからも現状でこれまでに収集できなかったデータを集約できるようになったことは意義があると考えられる。特に、上記のような、研究者から提供を受けた心理尺度をスケジューリングし、時系列条件がそろった縦断データを実際に一元的に集約した研究はこれまでにない。

ただ、本プロジェクトでは、新型コロナの影響の他、以下に説明する、当初想定していなかった問題により、プロジェクトの助成期間内には、大量の縦断データを収集することに留まらざるを得なかったことは非常に残念といえる。今後尺度

の提供を受けたプロジェクトメンバーの研究者には、十分なデータが収集できた尺度に限られるが、収集されたデータを完全匿名化した上で提供し、縦断データに基づく新たな研究を開拓していただくことを期待している。前掲した尺度の中には、岡山大学の学生約1800人の年単位の意識変動データ、700人の高校生の年間を通じた意識変動データなどが集められており、2021年度中には、高校生データで3年間を通じた縦断的意識変動データが集約できる所属も出てきている。本プロジェクトでは、コロナ禍によりデータの収集が困難になったこと以外に、データ構造の理解の困難さや簡易的な解析方法が十分開発されていないため、そのデータを活用した成果を論文として出すことはできなかったが、収集されるデータ自体に学術的価値が認められる状況が情報系の学問分野で出てきている。その点で、前述した岡山大学のGTECの得点との関連と同様、収集された高精度教育ビッグデータから新たな研究成果が出てくることが期待される。

・高精度教育ビッグデータの解析法の開発とその活用スキルの向上

収集される学習データに関しては、エクセルでハンドリングできない規模のデータが容易に集まるため、プログラムを作りデータを集約するスキルが必要になる。一方、本プロジェクトで収集した心理尺度データに関しては、一定の規模以下であればエクセルで操作できるため、エクセルのピボットテーブルを活用したデータの集約方法を開発し、プログラミングなく縦断データを心理尺度の因子ごとに集約し、学習者一人ひとりの意識変動はもとより、学校、学年、クラス単位で平均的な意識変動を可視化する方法を確立した。その集計方法については、プロジェクトメンバーで複数回研修会を開催した他、学習データと意識データのデータ構造と、意識データに関しては解析の方法をマニュアルとしてまとめ、心理系の研究者が高精度教育ビッグデータを直接操作できる方法を確立した。また、徳島県総合教育センターと連携し、高校生向けのビッグデータを解析するコンテストに過去に収集したビッグデータを提供し、本プロジェクトで確立した解析法を紹介し、コンテストに寄与することができた。本プロジェクトの成果とした「高精度教育ビッグデータで変わる記憶と教育の常識ーマイクロステップ・スケジューリングによる知識習得の効率化ー(風間書房)」の後半には、スケジューリングの具体的な方法やスケジュールの定義等がまとめられている。本プロジェクトで収集された縦断データの解析には、このようなスケジューリングの基本的原理の理解が必要であり、その基本的な解説書として本書は出版された。

・個人情報の保護と高精度教育ビッグデータの共用に関して新たに明らかとなった課題

当初、プロジェクトメンバーの研究者から心理尺度を提供してもらい、それをスケジューリングし、大量の縦断データを集約し、それを共用し、成果を広く公表していくことを計画していた。ところが、現在ビッグデータの収集に当たっては、データの提供元となる個人に、データの利用目的を明示し、同意をとることが一般的になされている。当初から研究利用を予定し、その旨同意を得ていたが、支援を継続していくうちに、全ての学習者から同意を得ることが困難な状況が出てくること明らかになった。大学におけるeラーニングでは、強制的に学習を求めることができない。そのため、やる気のある学習者からは同意が得られても、やる気のない(学習を行わない)学習者からは同意を得ることができず、収集されるデータには意欲に関して学習者の偏りが生じる事態が生まれることが明確になった。この状況で収集されるデータでは、厳密な研究はできない。このような状況は当初想定できなかったため、同意書の再検討が必要となった。

一方、現在eラーニングを教育支援として提供している岡山大学や学校等では、教育実践の一部として収集されるデータを分析し、指導等に活用することは認められる。その範囲では、実践研究としてデータを活用して成果をまとめることはできる。ただし、教育実践のためのデータから意味のある知見が得られても、その学校等の固有の情報や、公開に問題のある知見は公表することは難しい。そこで、実践研究としてデータは活用するが、得られた知見については全て公開せず、所属から承諾を得られたものに限定して公開していくという趣旨の同意書にする方向で動くことになった。岡山大学では、全学教育・支援機構(教学関係の最高の意志決定機関)で、大学の固有情報(例えば、GTECの大学や学部ごとの平均点)等は公開しない条件で成果の報告が認められることになった。

さらに、eラーニングの学習量がGTECの成績に効果を持つという冒頭の成果が得られる前から、私立高校などでは、独自に全校の生徒対象に実施している英検等の外部試験データも提供し、指導に生かしたいという要望も出されていたため、2021年度の支援のための同意書を根本的に検討し直した。

その他、個人ごとにまとまりをもつ縦断的なビッグデータの場合、完全匿名化した場合でも、後日、その他の端末のネットワークのアクセスデータ等と照合することで個人が特定される危険性もあるため、かなり慎重な判断が必要となることも明らかになってきた。特に、教育に関わるデータに関しては、個人が特定されることで大きな問題が出てくる可能性が高く、現在、一般的にデータを共有する方向で動いている学術の領域でも、慎重な判断が求められる領域であることは間違いない。現時点で最も多くの所属と学習者からデータを収集している本プロジェクトでは、その点を同意書の中で明示していく必要があった。そのため、学校等で実施される外部試験データを研究利用できる研究者の枠、第3者に提供できるデータの範囲など、これまで我々が想定していたデータの利活用のルールより詳細なルールを厳密にし、それに基づき同意書を作成した。その同意書のうち、データの扱いについての最新版を以下につけるので参考にさせていただきたい。

(同意書の該当部分)

◆e-Learning 実施に必要な個人を特定し得る情報の取得と取り扱いについて

e-Learning の提供と指導、データ解析および結果のフィードバックに必要な、最低限の個人を特定し得る情報(管理番号、学年・組・出席番号のみ)を所属機関(学校・自治体等)からセンターへ提供を受け登録させていただきます(氏名はセンターへ提供されません)。その情報の管理は、センターおよび学習者の所属機関が責任を持って行います。なお、学習結果のフィードバックに電子メールを利用する場合には、メールアドレスも個人を特定し得る情報として扱います。

提供を受けた個人を特定し得る情報は、インターネットから独立した専用の記録装置に保存し、岡山大学内の鍵のかかる部屋にて学習者の卒業までの期間厳重に保管の上、その後消去いたします。また、e-Learning 開始後に個人を特定し得る情報を取り扱えるのは、岡山大学の研究実施責任者*および研究実施責任者が監督する岡山大学の関係者のみです。

◆e-Learning で収集される学習データ等の取り扱いについて

本 e-Learning により収集される学習データ等は、個人を特定し得る情報と切り離し、それ単体では個人を特定できない形で記録・保存します。本 e-Learning により収集されるデータは、学習者の支援や所属機関の指導に用いますが、最適な学習法を解明していく研究等にとっても、非常に利用価値の高いデータになります。そのため、収集されたデータは、学習者の支援と教育に資することを目的に、個人が特定できないよう匿名化の処理を行った後、センター及び連携する機関で研究等に利用します。その研究成果等の公表にあたっては、全てのデータを統計的に処理した上で個人が特定されない形に加工し、さらに所属機関に承諾を受けます。

また、そのデータはさらに匿名化の処理を施し、可能な限り社会全体で広く共用できるようにします。ただし、最終的に匿名化された後のデータについても、安易に第三者へ提供せず、将来に渡り個人が特定されないようにする他、社会に与える影響などを考慮して提供する予定です。なお、最終的な匿名化の処理を行った後のデータは、センターでも個人を特定することが困難になるため、データの削除や利用停止のご要望にはお応えできなくなります。これらを踏まえ、第三者へ提供可能なデータの管理と運用に関しては、センターおよび連携機関にご一任いただく予定です。

※匿名化されたデータの利用については別途説明を加え、データ利用に同意いただく手続きを取らせていただきます。今回の参加申込を行っただけで、匿名化したデータをセンターと連携機関以外の第三者へ提供することは一切ありませんのでご安心ください。

◆外部試験成績等の付加的情報の提供・利用について

より効率的な学習法や新たな教育支援技術を研究開発していくため、e-Learning で収集される学習データ等とは別に、学校等で実施される定期テストや外部検定試験等の成績、普段の学校生活に関わるアンケートデータ等、付加的な情報の提供を所属機関に依頼し、所属機関が許可した情報を分析に利用する場合があります。具体的には、e-Learning の学習状況(学習ベースや成績等)と外部試験等の成績との関係を分析し、より効率的な学習法を明らかにし、学習者の支援に活用することや、学習意欲等の意識変化と学校行事や生活指導のタイミング等との関連を分析し、学習者に寄り添った支援技術を研究開発する等の活用を想定しています。また、これらの付加的情報を用いた分析結果を、論文や学会発表、プレスリリース等で公表する場合は、全てのデータを統計的に処理した上で個人が特定されない形に加工し、さらに所属機関に承諾を得たうえで公表します。

これらの付加的な情報を取り扱えるのは、センターの研究実施責任者、研究実施責任者が監督する岡山大内関係者および連携研究者のみとし、その他の第三者は利用できないようにします。なお、連携研究者へ当該データを提供する場合は、連携研究者が上述したセンターのデータ管理の方針に従うこと、およびデータに関して機密保持を遵守することを、センターと連携研究者との間で契約を結んだ上で、データを匿名化し、個人が特定できない形に加工して提供します。

◆保護者の皆様へのお願いと免責事項

保護者の皆様にも簡単なアンケートに答えていただく場合があり、可能な範囲でご協力をお願いしております。その際に取得されるデータの扱いは、上記学習データの取り扱いと同じ方法によります。

なお、今回の e-Learning への参加申し込みをいただけない場合、他の学習者と同じような個人ごとの成績のフィードバックを受けることができなくなり、学習者の不利益となる恐れがあります。e-Learning の趣旨と内容および各種データの取り扱い等についてご同意の上お申し込みくださるようお願い申し上げます。今回の e-Learning への参加や、データの扱い等に関するご質問・ご不明点等がありましたら、申込書の提出までに下記の研究実施責任者およびセンターまでご連絡いただき(担任の先生を通じてお問い合わせいただいても結構です)、必ずご納得の上でお申し込みください。

また、マイクロステップ・スタディの学習中に発生した事故や障害等に起因するいかなる損害も、岡山大では一切の責任を負いかねますので、あらかじめご了承ください。特にスマートフォンを使用して学習を行う場合は歩きスマホは絶対に行わず、目の悪くなるような暗い場所や至近距離での使用はさせないよう保護者の皆様からもご指導くださいますようお願い申し上げます。

・高精度教育データの利活用のためには、データの管理と共有の仕組みも必要である。完全匿名化処理はもちろんであるが、学習や意識データは、日々アップデートされ続けている。そのようなデータを基に論文を書く場合には、どの時点で、誰を対象にしたデータであるのかを特定できるしくみも必要である。本プロジェクトでは、必要となる特定の学習・意識データを外部から抽出できる API は開発したが(別の予算により)、抽出されるデータのオリジナルデータを再構成できるしくみや、誰がそのデータをダウンロードしたのかを管理できる認証システムも装備することは避けて通れない。

以上、研究成果の報告とするが、多数の心理尺度に関して、長期間の縦断データを収集できる状況が生まれ出されたことは、本プロジェクトの最大の成果といえる。また、このような大量の学習・意識データを学術研究に活用できるようにするためには、新たに解決が必要となる様々な課題が存在することも明らかになった。収集された膨大なデータからはこれから多数の新しい事実が明らかになると考えられるが、それと並行して高精度教育ビッグデータを学術利用するためのルールやデータ共用のためのシステムの構築に力を注ぐ必要があることも明らかになるなど、様々な課題が明らかになった3年間であった。

(注)氏名欄は適宜増減してください。

* 字数の制限はありません。記述欄が不足する場合は、複数枚になっても構いませんので適宜行数を増やしてください。