

Title	ラーニング・アナリティクスと教育データ
Author(s)	武田, 俊之
Citation	サイバーメディア・フォーラム. 2018, 18, p. 5-9
Version Type	VoR
URL	<a href="https://doi.org/10.18910/70426">https://doi.org/10.18910/70426</a>
rights	
Note	

*Osaka University Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

Osaka University

# ラーニング・アナリティクスと教育データ

武田 俊之（関西学院大学 高等教育推進センター）

## 1. はじめに

教育のデジタル化にともなって、教育機関の内外において生成、蓄積されるデータが増加している。そして、蓄積された「教育ビッグデータ」を、教育の理解と改善に活用するラーニング・アナリティクスが注目されている。

ラーニング・アナリティクスは、“the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs”と定義される[1]。その構成要素の関連を図1に示す。まず学習活動のモデリングがおこなわれる。このモデルにもとづいてデータが測定、収集される。データはモデルにもとづいて分析される。分析結果はモデルを参照しながら解釈されて報告・視覚化される。この結果は学習活動にフィードバックされる一方で、モデルの修正、洗練がなされる。

このようにラーニング・アナリティクスは学習活動を理解するための、データを中心としたプロセスである。本論文では（主に高等教育の）教育データに関連した課題について論じる。まず教育データの種類について整理をおこなう。次にデータの標準規格の概要とその制限について述べる。最後に倫理とプライバシーについて述べる。

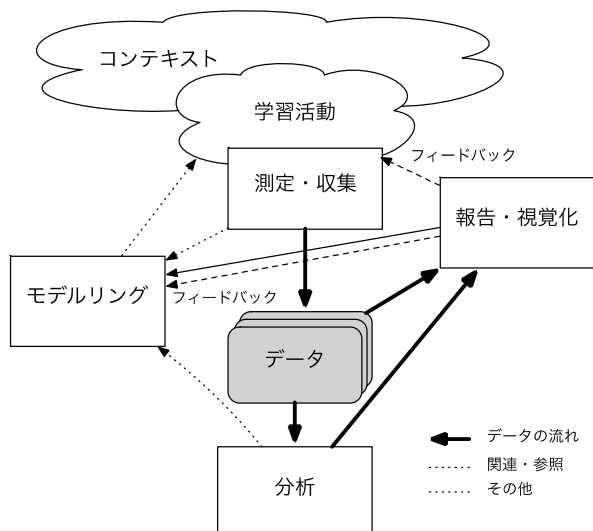


図1 ラーニング・アナリティクス概観

## 2. 高等教育の変容と教育データのかたち

高等教育は、教育のデジタル化と社会的要因の影響によって、大きく変化している。以下に、ラーニング・アナリティクスに関連した変化の一部を挙げる。

### 教育

- (1) 授業（フォーマル学習）だけではなく、授業以外の活動における学習（ノンフォーマル学習）や、大学外での学習（インフォーマル学習）を重視する傾向。
- (2) 授業での多様なシステムやデバイスの利用。
- (3) オープン教育資源（Open Educational Resources）、MOOCなど大学外の教育資源の利用。

### 組織

- (1) シラバス、カリキュラム・ポリシー、ディプロマ・ポリシーなど、各授業の内容や授業間の関連・位置づけの明確化への要求。
- (2) エビデンスベースの教育改善への期待と要求。

### システム

- (1) 単独システム（PC、LMS=学習支援システムなど）の授業への導入から、システム間の連携（学生データのひもづけ、API・データ連携など）への変化。
- (2) システムの大学内運用（オンプレミス）から、データセンターへの設置（リモート運用）。さらには、クラウドサービス利用への変化。
- (3) eポートフォリオのような、ノンフォーマル/インフォーマル学習を支援するシステムの導入。

### データ

- (1) システム、デバイスに蓄積されるデータの急激な増加、詳細化、精緻化（成績、アンケートから行動、生理データまで）。
- (2) 学術研究以外の、大学運営等におけるデータ分析への要請。

これらの考察を含めて、大学教育において学習活動に関するデータの主なソースには以下のようなものがある。

## 学習の種類

フォーマル (formal) / ノンフォーマル (non formal) / インフォーマル (informal)

## データ・ソースのレイヤ

個人 / インタラクション / 授業 / カリキュラム / 学部 / 大学 / 行政

## システム運用の形態

オンプレミス / リモート / クラウド

## システムの種類

LMS/SIS (Student Information System) / eポートフォリオ / 評価 / その他内部 / ソーシャルメディアなど

## データの観測方法

仮説にもとづく測定 (measured) / システムが生成したデータの収集 (collected) / 外部からのインポート (imported) / 分析・推定値 (inferred or computed)

## データの種類

成果 (成績等) / アウトプット (テスト、レポート、作品等) / アンケート / 行動履歴 / 生理指標 / 学習環境

## 個人データ

デモグラフィック / プロフィール / 事前の経験 / パフォーマンス / 心理調査 / 生理指標 / その他指標

## データの型

表 / 数値 / テキスト / Boolean / 複合型

このようにデジタル化された学習環境から得られる教育データはきわめて多種多様であり、今後はさらに詳細なデータが得られるようになるであろう。

## 3. 学習履歴データの標準規格

前節で述べたように、LMS等のシステムにおいて、クリック等の学習者の行動に関する詳細なデータが大量に生成されている。この学習履歴データを教育の理解と改善に用いられることが期待されている。課題の一つはシステムごとに異なる蓄積されたデータの形式である。そこで、学習履歴データ互換のために Experience API (xAPI)、IMS Caliper (以下 Caliper) などの標準が策定されている。

## 3.1 IMS Caliper の概要

xAPI と Caliper は学習の記録 (learning record や tracking log と呼ばれる) に関する規格である。xAPI は Advanced Distributed Learning (ADL)、Caliper は IMS Global が、それぞれ標準化をおこなっている。これらの規格の目的は学習活動中の行動をそのコンテキストを含めて記述することである。どちらも主語-述語-目的語の3要素からなるステートメントを核に記述する点が共通している。また、いずれの規格もシステム (LMS など)、デバイス (PC かモバイルか)、組織などの制約を越えた横断的な集計や比較を志向している。一方、詳細に異なる点も多く互換ではない、以下では Caliper 1.0<sup>2</sup>を例に説明をおこなう。

Caliper の要件は以下の通りである<sup>3</sup>

- 学習活動を観測する一貫した手段を確立する
- 学習データに関する共通言語を定義する
- 学習の活動と効果を測定する標準的手法を提供する
- データサイエンス、標準規格、技術をレバレッジする
- IMS Global の標準規格上に構築する

図2は Caliper 全体の概要である。Caliper は LMS などのデジタル学習システムにおけるさまざまな学習活動 (Learning Activity) をあつかう。Caliper では学習活動について、Assessment、Reading、Media といったタイプが定義されており、それぞれのタイプごとにメトリック・プロフィールが Entity、Action、Event の仕様を定めている。Entity は Person、Assignment といった学習活動の主体やリソースである。Action は各タイプにおける行為の種類、Event は Entity と Action によって記述される学習行動であり、図3のように JSON、JSON-LD が用いられる。Caliper 1.0 のメトリック・プロフィールは表1の通りである。

学習活動における行為 (action) は、システムに埋め込まれた Sensor によって Caliper の Event として捕捉される<sup>4</sup>。そして、Event は Envelope で梱包された上で Event Store に送られて、保存、処理される。表2は LMS 上のあるコースにおいて、課題として出題されたクイズに学生が回答する場合に生成される Event の例である。LMS 上のクイズモジュールが quizApp、クイズの自動採点モジュールが grading engine で最終的に grade (成績) が生成される。

<sup>1</sup> 両団体は規格の共通化を開始している。  
<https://www.imsglobal.org/initial-xapicaliper-comparison>

<sup>2</sup>現在 IMS Caliper 1.1 が策定中である。

<sup>3</sup> <https://www.imsglobal.org/caliper/caliperv1p0/ims-caliper-analytics-implementation-guide>

<sup>4</sup> xAPI の statement に相当する。

### 3.2 学習履歴データの課題

Caliper は LMS など情報システム側からの学習活動の記述である。学習者の記述は Entity として、OneRoster / LIS<sup>5</sup>（教育機関における登録リソース等を記述する規格）を利用する。現在のところこれらは学習活動の記述、分析のためのボキャブラリとして十分とはいえない。記述力を豊かにすることによって、分析対象となる学習活動の領域を広げることが Caliper 等の課題であろう。

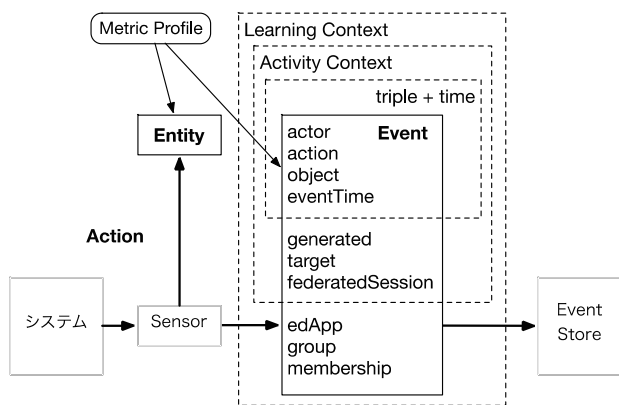


図2 IMS Caliper 概要

```
{
  "sensor": "https://example.edu/sensor/001",
  "sendTime": "2015-09-15T11:05:01.000Z",
  "data": [
    {
      "@context": "http://purl.imsglobal.org/ctx/caliper/v1/Context",
      "@type": "http://purl.imsglobal.org/caliper/v1/Event",
      "actor": {
        "@id": "https://example.edu/user/554433",
        "@type": "http://purl.imsglobal.org/caliper/v1/lis/Person"
      },
      "action": "http://purl.imsglobal.org/vocab/caliper/v1/action#Viewed",
      "eventTime": "2015-09-15T10:15:00.000Z",
      "object": {
        "@id": "https://example.com/viewer/book/3#epubcfi(/4/3)",
        "@type": "http://www.idpf.org/epub/vocab/structure/#volume"
      }
    }
  ]
}
```

図3 JSON 形式による Event の記述例<sup>6</sup>

表1 IMS Caliper 1.0 Metric Profiles

Profile 名	説明	主な action
Session	セッション	logged in, logged out, timed out
Navigation	ページ遷移など	navigated to
Reading	テキストの閲覧	searched, viewed, navigated to
Annotation	アノテーション	attached, bookmarked, commented, highlighted
Assignable	割当可能	started, completed, reviewed
Assessment Item	Assessment のアイテム	started, paused, restarted, submitted
Assessment	評価可能	started, viewed, skipped, competed
Media	音声やビデオの視聴	started, forwarded to, paused, rewind
Outcome	採点可能	graded

表2 課題（クイズ）自動採点の例<sup>7</sup>

	学生の行動	Profile	Actor / Action / Object 他
1	課題を開く	Navigation	student / navigated to / assignment
2	課題を開始する	Assignable	student / started / assignment generated: attempt
3	クイズを開始する	Assessment	student / started / quiz app: quizApp
4	問1の回答を開始する	AssessmentItem	student / started / question 1
5	問1の回答を完了する	AssessmentItem	student / completed / question 1
6	4-5 繰返し		
7	課題の提出	Assessment	student / submitted / assignment
8	課題が採点される	Outcome	grading engine / submitted / assignment (attempt - result) generated: grade app: grading engine

<sup>5</sup> <http://www.imsglobal.org/activity/onerosterlis>

<sup>6</sup> <https://www.imsglobal.org/caliper/caliperv1p0/ims-caliper-analytics-implementation-guide#4.2> を改変

<sup>7</sup> <https://www.imsglobal.org/caliper/caliperv1p0/ims-caliper-analytics-implementation-guide#5.3> を改変

## 4. 倫理上の課題

多様かつ詳細な学習者のデータの蓄積と分析によって、学習活動の理解と改善が期待されている。本節ではこれにともない生じる倫理面の課題について述べる。

### 4.1 プライバシーと個人情報保護

長期間にわたる詳細な学習履歴は、そのデータの主体にとってのプライバシーそのものである。2017年5月に改正された個人情報保護法は、EU、北米におけるプライバシー保護の考え方にも沿ったものである。この法律は民間が対象で、独立行政法人と行政機関については別に法律が定められているため「教育」の実践すべてに関わるものではない。日本の教育データのプライバシー保護について、教育業界の認定個人情報保護団体が立ち上がるまでは、プライバシー保護に関する法律家や研究者の書籍(たとえば中川[2] や日置ら[3]、個人情報保護委員会(個人情報保護に関する第三者機関)のガイドライン、先行する欧米の研究や実践が参考になるであろう。

個人情報保護法ではデータを匿名加工することによって、データ主体の許諾なく第三者による分析が可能とされている。匿名加工は単にID、氏名など個人を識別する符号の削除や暗号化では不十分で、個人を特定不可能になるように行動履歴等を加工しなければならない。一方、匿名加工情報が特定の個人を完全に識別できないことを判別、証明することは困難である。匿名加工情報が分析の精度を劣化させるという主張があるが、研究上の知見が得られているわけではない。匿名化された教育データの有用性の研究は今後の課題である。

パーソナルデータの利用で先行する医療分野ではリスク評価あるいは基準を設けて手順を透明化することがおこなわれている[4]。また、統計処理の結果は多くの場合個人を識別することはできないため、プライバシー保護技術の一種と考えることもできる。

データ主体から同意を得ることも個人情報保護法で許されているデータの第三者提供の方法である。大学の教職員が教育データを直接的な教育改善(その授業やカリキュラム)に用いる場合には許諾は必要ない。しかし、教育機関は学生に対して強い権限を持っており、多くの場合に学生が個人情報の提供を拒めないことは留意しておく必要がある。拒否できない学生の同意をもって、教育と直接関係のない目的外のデータ利用(たとえば研究や長

期的なデータ分析)に用いることは望ましいとはいえない。大学運営でデータを追加した長期的な分析が必要ならば法令遵守と個別の許諾を得ることによって、また、研究で用いる場合は研究倫理審査委員会(IRB)の審査とともに個別の許諾を改めて得ることが必要となるであろう。

### 4.2 アナリティクスの担い手の変化

従来、教育データをあつかう倫理は「研究倫理」であった。しかし、先に述べたように、データを用いた大学改善が求められるようになり、教育改善等大学運営のための教育データ分析のニーズが増加している。データを扱う専門職といっても、研究者と大学運営のためのデータ分析者では、それぞれの倫理に異なる点がある。

人を対象とする研究において、個人から行動、環境、心身などに関するデータを取る場合、事前に研究倫理審査を受けた上で、取り扱い方法を明示して提供者の同意を得ることが必要である。しかし、大学運営のための分析プロジェクトについては、データの収集、分析、利用について、ガイドラインの策定、第三者委員会による審査がプロセスに組み込まれることは少ない。このような組織的なデータの取り扱いの枠組みは今後の検討課題であろう。

Cormack[5]は学術研究ではないアナリティクスについて、法令遵守を含めたフレームワークを提案している。日本でも、このようなフレームワークに、改正個人情報保護法とそのガイドラインを含めることは考えられる。またDELICATEチェックリスト(後述)は教育データのステークホルダーが安心できるアナリティクスのためのチェックリストである。

### 4.3 データの所在と所有の問題

デジタル学習環境が大学からクラウドに移行している。大学が運営するサービスにおいては、大学がデータの取得と管理をおこなう。この場合、データの利活用は、大学と学生などのユーザー間の合意に基づくと考えられる。また、emailなどのサービス運用をインターネット上の外部サービスへ大学が委託することは、契約の性質から大学が運営する場合と同様に大学とユーザーの合意に基づくと考えることができる。しかし、この場合には大学側でのデータ利活用は可能であるが、委託先の事業のためのデータ利活用は不可と考えられる。

問題となるのは、委託ではないクラウドサービスの購入と利用である。この場合、大学はサービスの利用者(大

口ではあろうが)であり、ライセンス上さまざまな問題が生じる。たとえば、クラウド事業者は契約上制限されていなければ、改正個人情報保護法にもとづいて、匿名加工したデータを第三者に提供することができる。果たしてこれは問題がないのだろうか。

このように、データがどこにあるか(所在)と、データは誰のものか(所有)に注意しなければならない。データが存在するサーバーの管理主体は、契約上の制約がなければデータを処理することができてしまう。

所在と所有の問題は大学内にも存在する。教学の一次的な責任が学部にある大学において、全学の情報部門がLMSを運用しているとする。この場合データが存在する情報部門がデータへの自由にアクセスを許されるのか。望ましいデータ分析の手続きはどのようなものか。

#### 4.4 DELEGATE チェックリスト

Drachsler ら[6] は、ラーニング・アナリティクスの倫理とプライバシーについてレビューをおこなった上で、ステークホルダーが安心できるような実践のためのDELEGATE チェックリストを開発した。

**DETERMINATION (決定)** : 組織内でラーニング・アナリティクスの目的を決定する

**EXPLAIN (説明)** : データの収集と利用について目的を定義する

**LEGITIMATE (遵法)** : 法律を順守したオペレーションであることを説明する

**INVOLVE (関与)** : すべてのステークホルダーとデータ主体が関与する

**CONSENT (同意)** : 明快な方法によってデータ主体から同意を取る

**ANONYMISE (匿名)** : 可能なかぎりデータから個人を識別不可能にする

**TECHNICAL (技術)** : データへのアクセスをモニターしてプライバシー保護を保障する

**EXTERNAL (外部連携)** : 組織外との連携が必要な場合により高水準のセキュリティ標準を用意する

DELEGATE は有用なチェックリストであるが、今後の研究の進展や採用する組織の目的・目標によって修正されるべきものである。

## 5. おわりに—信頼できるラーニング・アナリティクスのために

これまで概観したように、デジタル環境の進歩によってきわめて大量かつ多種多様になった教育データは学習者をはじめとする教育の関係者の行動を捕捉していて、その分析と結果の利用が影響をおよぼす範囲は大きい。データ主体の保護を優先にデータをあつかうことは、教育改善、大学運営、学術研究のいずれの立場においても望まれる。しかし、このような手続きを面倒なものとして、データ利活用の足かせになるという意見もある。

Drachsler ら[6] は「信頼できるラーニング・アナリティクス (trusted learning analytics)」を実現するための倫理、プライバシー上の課題について検討している。データ主体を守るための必要な手続きをスキップすることなく、分析の結果を適切にフィードバックして教育の価値を高めることが、信頼されるラーニング・アナリティクスの発展のために望まれる。

### 参考文献

- [1] P. Long and G. Siemens, “Penetrating the Fog: Analytics in Learning and Education,” *Educ. Rev.*, no. SEPTEMBER/OCTOBER, 2011.
- [2] 中川裕志, プライバシー保護入門. 勁草書房, 2016.
- [3] 日置巴美 and 板倉陽一郎, 個人情報保護法のしくみ. 2017.
- [4] K. El Emam, L. Arbuckle, 木村映善(監訳), 魔狸(監訳), 笹井崇司(訳), データ匿名化手法 —ヘルスデータ事例に学ぶ個人情報保護. オライリー・ジャパン, 2015.
- [5] A. Cormack, “A Data Protection Framework for Learning Analytics,” vol. 3, pp. 91–106, 2016.
- [6] H. Drachsler and W. Greller, “Privacy and Analytics – it’s a DELICATE Issue A Checklist for Trusted Learning Analytics,” pp. 89–98, 2016.