

高等教育アナリティクスのための階層モデル

武田俊之[†]

従来から存在するデータの電子化や学習支援システムやeポートフォリオなど新しいシステムから生じるログなどによって、高等教育機関において利用可能なデータが著しく増加している。これらのデータを活用した教育や経営管理の改善については、機関調査 (Institutional Research) やラーニング・アナリティクスの領域が立ちあがっている。本論文では、これらの領域における議論と高等教育機関に関連した多様な立場を整理した上で、データの収集方法、所在、所有利用の権利、分析方法などを整理・俯瞰するための階層モデルを提案する。

Multi-level Framework for Analytics in Higher Education

TOSHIYUKI TAKEDA[†]

Data available in higher education increases significantly; for digitized from legacy data and generated from new systems such as learning management systems and e-portfolios. New research and practice domain (institutional research, learning analytics and so on) aims to utilize these data for decision-making, helping learners in success and create a shared understanding about learners and the institute. This paper proposes a multi level model for analysis and synthesis that organize right data collection and its method, data location, ownership of data.

1. 背景

さまざまな政治経済的な背景から、根拠をもって大学の教育を評価することが求められている。しかし、知識や技術を学ぶことは長期的で多様な過程であり、教育や学習の評価は本質的な困難さをとまなっている。

さらに、さまざまな社会と技術の変化によって、多面的な教育と学習の評価を必要としている。変化の例としては、学生の変化、高等教育の国際化、市民や政府等からのアカウンタビリティへの要求、技術上の変化（コンピューターやセンサー技術の高性能化など教育に利用可能な技術の発展、インターネットやソーシャル・ネットワーク・システム (SNS) の発展）などがあげられる。

これらの変化のうち、インターネットやデジタル技術の進歩は、利用可能なデータの変化を質量ともにもたらしている。高等教育においては、入試、個人属性、履修、成績などの以前から電子化されてきたデータから、さまざまな調査アンケートや、学習支援システム (LMS) やeポートフォリオなどのログ、ICカードなどによる出席状況などがある。これらは今後も種類や量が増えることが予想される。

ビジネスにおいて、分析を駆使する組織は、成果と持続可能な競争優位をもたらす[4]。高等教育においても、データの分析を推進するために、機関調査 (Institutional Research: IR)、教育データマイニング、ラーニング・アナリティクス、アカデミック・アナリティクスなどの領域が立ちあがり、組織的、学術的な取り組みがおこなわれている[14]。アナリティクスとは、データ分析をおこない、結果からパターンを発見や洞察を導いて、予測や実行可能な計画につなげる活動である。

2. 高等教育アナリティクスの領域

機関調査 (Institutional Research: IR) の目的は、高等教育機関に関する様々なデータを分析し、効率的な運営を行えるようサポートすることである[16]。IRは、外部からの説明責任、サービスへの要求の変化と増加、教育の質保証、経営上の課題（たとえば、卒業率の向上など）などの要求の変化とともに発展してきた。典型的な分析の例としては、退学者数とデモグラフィックな属性、成績、経済状況と関連があげられる。アンケートによる学生調査はIRの変化にともない発展している。山田[15]は成績などの直接評価とプロセス評価としての間接評価結果の組み合わせによって、大学生活を通じての学生の成長に関する精緻な結果測定が可能であるとして、日本版大学生調査(JCSS: Japanese College Student Survey)や日本版新入生調査(JFS: Japanese Freshman Survey)を開発している。

教育データマイニング (Educational Data Mining: EDM) は統計学、機械学習、データマイニングなど、情報学および心理学の方法を利用した学習を研究する領域である。EDMの目標としては、学習者モデルの構築とそれに基づく学習行動の予測、コンテンツやインストラクションの最適化、学習ソフトウェアが提供する教授 (支援) 方法の効果の研究、計算モデルの構築を通じた学習および学習者に関する知見の獲得などがある。

ラーニング・アナリティクスは、立ちあがったばかりの領域で定義はさまざまであるが、1st International Conference on Learning Analytics and Knowledgeでは、"learning analytics is 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"となっている。EDM が学習者へ自動化された適応的な反応を志向しているのにくらべて、学習状況にあったテーラーメイドの改善を志向していると思われる[2, 6]。

アカデミック・アナリティクスの定義[3]は、"Academic analytics can help institutions address student success (commonly measured as degree completion) and accountability while better fulfilling their academic missions. Academic systems generate a wide array of data that can predict retention and graduation. "とすることで IR に近いが、教育におけるコンピューターによって生じたデータを活用するために発展してきている点では、教育データマイニングやラーニング・アナリティクスと文脈を共有している[3, 8]。

3. アナリティクス領域の関連と課題

以上のような高等教育におけるデータ活用に関する領域を関連づけるためのいくつかの考察を述べる。

3.1 総括的評価と形成的評価

評価は総括的評価と形成的評価に分類することができる。総括的評価は、図 1 に示すようなプロセスにおいて、原則としてあらかじめ定めた目標について、主に数量的に測定された Input/Output/Outcome を利用しておこなう評価である。Output や Outcome として測定する指標は原則として事前に設定しておくものであり、評価の時点で追加することはない。

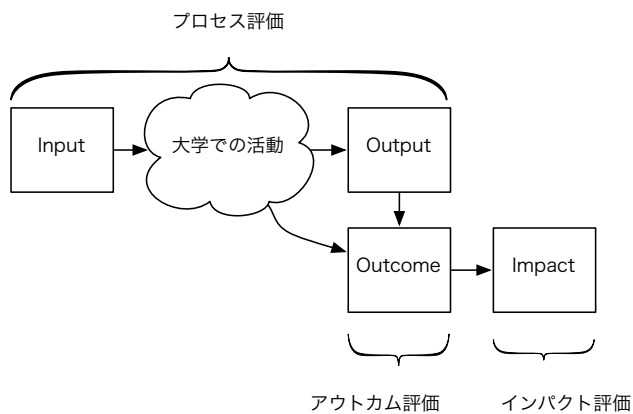


図 1 総括的評価

一方、形成的評価は活動の途上でおこなわれる評価である。たとえば、授業の目標について分析と構造化をおこない、授業における目標の達成状況を把握して、指導やカリキュラムの改善をおこなうための評価である。

データに基づく改善について考えた場合、ある形成的評価の中に、入れ子になって総括的評価が含まれていることが一般的であろう。

評価については、梶田[10]の指摘する以下の問題が存在する。いずれの評価も、活動の分析によって定めた目標に関連した指標を測定することによっておこなわれる。しか

し、教育のような多種多様な価値を含む活動は指標だけですべて記述できるわけではない。教育の評価は、目標に結びついているかどうか、測定できているかどうかによって、4セルのマトリックスに分類することが可能である。そして、目標に結びついていて、測定可能な指標だけが評価において利用される。評価の進歩によって、この割合は増加するが、活動の 100%把握できるほどになることはない。これは、評価にもとづいて意思決定をおこなう場合に意識しておかなければならない。

3.2 データの所有

さまざまなシステムが生成するデータは誰の者か、所有 (ownership) は自明ではない。特に、個人を特定する ID と関連づけてデータを結びつけることによって、プライバシーの問題が生じる可能性がある。入試から卒業、さらには卒業後までを申ししにすること、あるいは家庭環境などの属性とともに分析することによって、さまざまな分析結果が得られるとしても、それは倫理的にゆるされるかどうか。活動から生じたデータの所有者は誰であるか、また、誰がデータの利用を許されるかについては、明確な結論が得られるような問題ではない。しかし、学生や教員を含めたステークホルダーが納得するような一貫したポリシーを持つためには、明確な枠組が必要であると思われる。データの利用に関する倫理的な側面については後に考察する。

3.3 データの所在

データの所有に関するポリシーが明らかであったとしても、活動から生じたデータの所在がその当事者からは遠いレベルにあることがある。

たとえば、学習支援システム (LMS) の運用は、以前は教員個人や学部で運用されることが少なくなかった。しかし、利用者の増加や運用の負担増加にともなって、大学が運用する例が増えている。この場合、授業の教材や生成されたデータは、一括して大学に保持されていて、学部からアクセスが容易ではないことがある。

同様に、学部の責任において実施されたマークシート式の授業評価アンケートを大学で取りまとめて、一括して分析することがある。この場合、アンケートの結果は各項目に関する平均や偏差程度のものであろう。

いずれの例も問題となるのは、授業に責任を持っている教員、そして、その授業をカリキュラム上に位置づけた学部がデータへ直接アクセスできないことである。実際には必要な人がデータへアクセス可能な仕組みをシステムあるいは組織として用意すればよい。しかし、所在と所有、データの流れを示すための枠組が必要である。

3.4 データマップ

高等教育におけるデータマップとは、データとそれに関するビジネス・プロセスのスキーマを指す。ワシントン大学では、各システムで利用生成されるデータを、マスターデータを中心に、研究、教育、人事、財務、サービス、校

友の領域に分類している。ワシントン大学におけるデータマップ^aの目的は以下の通りである。

- 情報管理の課題について、ビジネスと技術の関係者が共通の基準を共有する
- データ管理についての意思決定をガイドする
- システム間とビジネス・プロセス間の差異の境界と依存関係を明確にする
- 情報のグループ化と普及のために、共通の基準を作成することによって、効果的に情報を共有する
- 情報システムと分析レポートの構築、拡張のための戦略的計画のプロセスを支援するフレームワークとして用いる

日本においては主に経営的な観点から、機能や組織によって分類されたデータを関連づけるデータマップの作成がおこなわれている [7, 9]。藤原ら [7] は、立命館大学における教学分野のデータの活用状況を可視化するために、主要データ 54 項目を「収集」「報告」「調査」「分析」の区分で分類したデータマップを作成している。

3.5 アナリティクスの分類

Barneveld ら [1] は、データの分析と、そこから得られる洞察や改善にかかわるアナリティクスに関する定義を収集・整理している。そして、それらを整理して作成した概念的なフレームワークを、Scholarship of Teaching and Learning (SoTL) [9] と関連づけた。SoTL は、教員による授業実践に関する学術的探究を通して教授・学習過程を改善する試みである。また、Long ら [11] は、表 1 のように分析の対象とステークホルダーからアカデミック・アナリティクス、ラーニング・アナリティクス、教育データマイニングについて整理している。

Long らはアナリティクスの意義を、IR が担ってきた経営管理における意思決定や資源配分のようなトップダウンの活動だけではなく、個々の学生の学習パフォーマンスや卒業に向けたリスクの提示などのボトムアップの活動の両方に見ている。これらはどちらが重要というものではない。さまざまなレベルで生成されたデータとその分析と、その分析を洞察に結びつけた組織的な対話によって、トップダウンとボトムアップの両方が豊かな結果をもたらすと考えられる。

4. アナリティクス階層モデル

以上の課題を検討しながら、Barneveld ら [1] と Long [11] らの研究を元に、高等教育において生成・収集されるデータと分析・利用に関して、実践と研究において利用可能なアナリティクス階層モデルの作成を試みた。表 2 は、フォ

ーマルな授業の例である。階層は、① 学生、教員などの Individual ② 授業 (Course) ③ 学部など組織 (Organization) ④ 大学などの機関 (Institute) である。これらはさらに階層が分割されることがある。教育の場合、組織はカリキュラム (Curriculum)、学科 (Department) などに分けることができる。

このモデルでは、経営を上位に置いた一般のピラミッド型の階層とことなり、Individual を上に置いている。表の上では上位下位と呼ぶが、分析においては相対的にマクロとマイクロで表現する。

階層モデルの主な目的は以下の通りである。

① データの所在と所有の明確化

先に述べたように、それぞれのデータの所在と所有を明確にすることが、アナリティクスにおいて重要である。ワシントン大学のデータマップは現在の所在と利用の権限は明らかにするものである。それにくわえて、システムや組織の変更による所在や所有への影響を示すためには、組織や活動の階層の中に位置づける方がよい。

② 関係者間におけるボキャブラリの共有

アナリティクスに関係しているさまざまな組織や学生を含む関係者は、それぞれ機能や目的がことなるために、それぞれ必要とするアナリティクスで使用するボキャブラリも異なる可能性がある。各アナリティクスを階層レベルに位置づけて関連を明確にすることによって、ボキャブラリを共有する。

③ 実践-研究間でのデータと結果の共有

認知科学、教育学、情報科学などにおいて、教育に関連する研究とデータが生みだされている。これらの研究は仮説検証型と仮説生成型の大きく分けることができる。これらの仮説とその分析結果を階層内に位置づけて、アナリティクスと関連づけることが期待できる。一方、実践のデータにもとづいて、予備調査や仮説検証の補助に用いることも考えられる。

④ メタ分析とデータフュージョン

同じレベルに位置づけられたデータは、統合が可能である。ここでいう統合は、ID によってひもづけることだけではなく、メタ分析や共変量を利用したマッチングなどのデータフュージョンを含む。

⑤ データの収集・分析に関する倫理面

倫理面の問題から、マイクロからマクロのデータは統計的な処理を通じてアクセスすることが望ましい。多くの倫理面の問題を解決するために、データや統計的手続きの組みあわせを整理する。

a

表 1 高等教育におけるアナリティクス

	アナリティクス	アカデミック・アナリティクス	ラーニング・アナリティクス	予測的アナリティクス
定義	データにもとづく意思決定として定義される全般的なコンセプト [1]	大学の行政、財務の意思決定を支援するためのデータに関連するプロセス [7]	学習者とそのコンテキストに関して、その理解と最適化のために、データの測定、収集、分析、レポートをおこなうこと [6]	行動やイベントを予測を目的とした、データ内の相関やパターンの発見する統計的分析に関する領域 [14]
レベル	各レベル	大学、学部	学習者、教員、学部	各レベル
サービス (例)	授業評価アンケート、学生調査、卒業生調査	経営ポータル	成績表、個人用ポータル、パフォーマンスの可視化、ダッシュボード、eポートフォリオ	成績の分析、LMSデータの分析

表 2 アナリティクスの階層とデータの例

レベル (組織)		学習者	教員	環境	システム (所有)
個別	属性	デモグラフィックな属性、過去の履修・成績	専門分野、教授歴、フィックな属性	デモグラフィックな属性	授業で利用する施設・設備の属性
	活動	出席、発言、課題の遂行、学習履歴	講義、演習の実施		
授業科目	各回	出席者数、課題遂行の状況	講義資料、課題、教授プロセス	各回で用いた設備	授業中レポート
	全体	出席状況、成績等の分布	授業全体のデザイン、リフレクション	授業で用いた施設・設備	
学部	カリキュラム	カリキュラムにおけるパフォーマンス	教員数、開講科目数、履修の分布	教室、設備の状況	
	学部全体	学部の授業におけるパフォーマンス			成績、出席、レポート
大学		出願者、受験者、入学者、退学者、休学者、卒業生			履修、LMS、eポートフォリオ

5. アナリティクス運用原則と倫理面の検討

データの所在がどのレベルにあるかは、サービス設計で決定される。このデータに対してどの活動や活動で役割をもった人がアクセスできるかについて運用の原則を決定しなければならない。ここで活動は組織の区分が一致するとは限らない。

5.1 分析と利用の原則

まず、アナリティクスおよびデータ運用の原則を検討する出発点として、マクロレベルからはマイクロレベルの一つ一つのデータにアクセスして利用することはできないこととする。マイクロレベルの評価は、マクロレベルで得られる統計量などの要約されたデータとしてのみ把握することができる。この規約によって、アクセス権限の適切な管理と、データの利用範囲の明確化を実現することができる。

一方、マイクロレベルの改善は、マクロレベルの評価指標にもとづいておこなわれる。マクロレベルに所在する同種の活動の評価指標は、改善のためのベンチマークとして利

用可能であろう。

図 2 に Individual/Course/Curriculum の部分の例を示す。見やすさのために、階層モデルと上下が逆になっている。左がある授業とそのデータである。データの生成と所在はレベルの境界にある。右は別の授業で、マクロレベルでは統計量として要約されている。

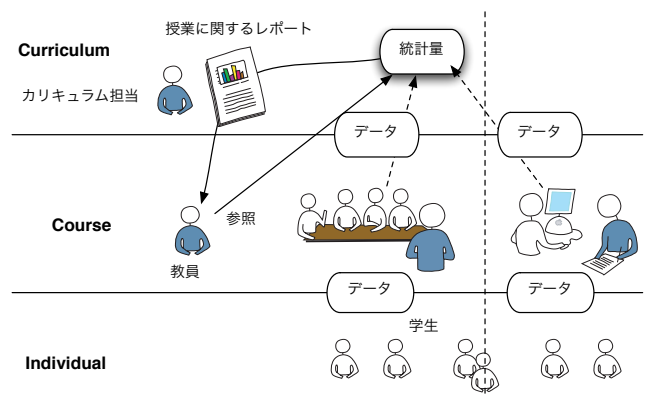


図 2 分析と利用の原則

データにもとづいて改善をおこなう場合に、マクロレベルの分析結果をそのままマイクロレベルの施策にそのまま使用しない方がよい。マイクロレベルのデータから次元を落とした要約として得られる分析結果によって、マイクロレベルの俯瞰と変数間の潜在的関係を知ることができる。これは、分析の結果、改善すべき点があった場合には、マクロレベルの施策を通じて、改善するのが原則である。また、分析から得られた結果を、個々の指導に利用することで問題が生じる可能性がある。

たとえば、LMS の利用時間と成績との間に正の相関があったとする。この場合、LMS の利用時間を増やすよう課題レポートを増やすようにしてもよい結果が得られるとはかぎらない。これは単純すぎる例であるが、あるレベルにおいて得られた分析結果は、間接的にその指標が改善するような施策につなげるのに限る方がよいと考える。

5.2 所有と利用の原則

サービスの運用者がかならずデータの所有者であるわけではない。この原則に従うと、たとえば、大学レベルに位置づけられる組織が学習支援システム (LMS) を運用していることは、その組織で管理されているデータを自由に利用する権限を保証しない。総括的・形成的評価において定められた評価項目 (たとえば、LMS の利用科目数や延べ利用者数) の統計は利用するとしても、さらにマイクロなデータの利用と分析はポリシーを設定によって権限が定められることになる。

一方で、ある授業において LMS で実行したログの利用権限が授業担当教員にあるならば、そのデータはシステム上のデータベースから分解 (decompose) されて教員に渡される必要があると思われる。また、その権限が学部のカリキュラム担当者にもあたえられているなら、その人 (役割) にデータが渡される。複数の役割に権限をあたえることができることは自然なことである。

5.3 倫理面の検討

先に述べたように、個人のさまざまな属性であるデータを集積、連結することによって、個人のプライバシーの問題が生じる。しかし、マクロのレベルの施策で用いるために、個人を特定することが可能なマイクロレベルのデータの分析を求めることが少なくない。その理由の一つは、強い倫理基準を求められる研究においてデータをあつかった専門家だけではなく、経営の側面から有効な施策のための根拠としてデータを必要とする人が増えたことであろう。

誰がどのデータにアクセスできるかを明らかにしておくことは、倫理面において特に重要である。ポリシーで定めることができる。それぞれのレベルの活動への参加者だけが利用できるのが最初の原則として適切であろう。

プライバシーを確保しながら、サービス連携を実現した例としては、名古屋工業大学のシステムがある。これは、

授業出欠システム、図書館利用、計算機ログインのどれも利用記録がない場合に学生相談室のみにアラートが知らされるようになっていいる。

6. 今後の展望と課題

教育において、生成されるデータは増加する一方である。経営的、社会的な期待にこたえるために、また、学習者の能力開発とその環境を向上させるために、データの高度な活用は必須である。このとき、データの適切な所有と利用の下で、学習者が単なる分析結果の利用者ではなく、自らのデータの分析者であるかもしれない。

一方で、学習者自身が望まない履歴データの利用を防止しなければならない。そのためには、6.1 節で述べたアナリティクス運用の原則を確立する必要がある。

いずれにしても、アナリティクスを積極的に活用するためには、それぞれの機関のそれぞれのレベルにおいて、提供しているサービスの価値にもとづいて、サービスのユーザーおよび提供者の目標を定義することが必須である。その目標は評価のために測定可能な形で定義される。

各サービスの評価は、上下または同じレベルの他のサービスの統計量などの分析結果を通じて、強化される。また、ポリシーを明確に定めることによって、Individual レベルのデータと結びつけて詳細な分析を可能にすることもできるであろう。

ダベンポート [4] は表のように組織が分析力を活用するステージを分析している。高等教育においては、経営や質保証のようなマクロレベルだけではなく、カリキュラムデザインや個々の授業、さらには学生自身による学習過程の可視化と改善のようなマイクロレベルにいたるまで、それぞれのレベルのプロセスにアナリティクスが埋めこまれて利用されることがステージ 5 と考えられる。

表 3 分析力を武器にする企業のステージ ([4]を元に作成)

ステージ	データ分析活用の現状
1. 分析力に劣る	ほぼゼロ、盲目飛行中
2. 分析力の活用が限定的	部分的または場当たりの
3. 分析力の組織的な強化に取り組む	統合的なデータ収集・分析に着手
4. 分析力はあるが決定打にいたりない	全社的に分析力は備わっており、活用もされているが、戦略の柱や絶対的な武器にはなっていない
5. 分析力を武器とする	全社でデータ分析が徹底され、成果に結びつき、持続可能な競争優位となっている

7. 謝辞

この研究の一部は、関西学院大学高等教育推進センター研究助成「Learning Analytics の研究」の補助を受けた。

参考文献

- 1) Barneveld, A., Arnold, K., and Campbell, J. : Analytics in Higher Education: Establishing a Common Language. EDUCAUSE Learning Initiative Jan:1-11 (2012).
- 2) Brown, M.: Learning Analytics: The Coming Third Wave. EDUCAUSE Brief Apr:1-4 (2011)
- 3) Campbell, J. P. and Oblinger, D. G.: Academic Analytics. EDUCAUSE Quarterly Oct (2007)
- 4) Davenport, T. H., Harris, J. and Morison: R. Analytics at Work: Smarter Decisions, Better Results. Harvard Business School Press (2010)
- 5) Eckersen, W.: Predictive Analytics: Extending the value of your data warehousing investment. TDWI Best Practices Report First Quarter (2007)
- 6) Ferguson, R.: The State Of Learning Analytics in 2012: A Review and Future Challenges. Technical Report KMI-12-01, Knowledge Media Institute, The Open University, UK. (2012)
- 7) 藤原将人, 近森節子, 浅野昭人, 吉井直宏: 教学分野の政策策定を支援する Institutional Research (IR) の構築. 大学行政研究 4:17-31 (2009)
- 8) Goldstein, P. J. and Katz, R. N.: Academic Analytics: The Uses of Management Information and Technology in Higher Education. ECAR Research Study 8 (2005)
- 9) Hutchings, P. and Shulman, L. S. : The Scholarship of Teaching: New Elaborations, New Developments. Change 31(5):10-15. (1999)
- 10) 梶田叡一: 教育評価 (第2版補訂2版). 有斐閣 (2010)
- 11) Long, P. and Siemens, G.: Penetrating the Fog: Analytics in Learning and Education. EDUCAUSE review, Sep (2011)
- 12) 小野宏: 関西学院大学における IR の現状・課題・展望. 関西学院大学高等教育研究 1:59-79 (2011)
- 13) Ravishanker, G.: Doing Academic Analytics Right: Intelligent Answers to Simple Questions. Research Bulletin 2, EDUCAUSE Center for Applied Research (2011)
- 14) 武田俊之: 高等教育におけるデータの活用と課題, 関西学院大学高等教育推進センター紀要 2:37-48 (2012)
- 15) 山田礼子: 大規模継続学生調査の可能性と課題. 大学論集, vol. 42, pp. 245-263 (2011)
- 16) 柳浦猛: 「アメリカの IR の本質」?. IDE, 2-3 月号, pp. 12-16 (2011)