

大学教育とビッグデータ：その可能性

美濃導彦[†]

情報通信技術が我々の社会生活の様々な場面に深く入り込んだ社会は情報社会と呼ばれている。情報社会では、ログデータ等として得られる様々な形式の膨大なデジタルデータが蓄積され「ビッグデータ」が出現している。このデータの解析を通じて新しい知識や価値を創造する研究開発手法が注目を集めている。一方で大学においても、教育の情報化のための環境整備が進みつつあり、センターサイドではコース管理システム・学習管理システムやeポートフォリオシステム等が、ユーザサイドではPC・タブレット・スマートフォン・クリッカ等が、様々な学問分野の多様な教育学習活動において利用されるようになってきた。その結果、様々な形式の膨大なデジタルデータを蓄積し、それらをつなぎ合わせ解析する「ラーニングアナリティクス」が注目を浴びている。しかも、MOOC (Massive Open Online Course) の登場により、これまでになかった形での膨大な教育学習記録の蓄積や利活用の可能性も見えてきている。本稿では、情報社会の特性を考えた上で、情報社会における教育の問題を議論した後、大学教育の情報化推進における問題点とビッグデータへの期待について述べ、今後の大学教育改革について考えてみたい。

Higher Educational Big Data and Its Possibility

MICHIHIKOMINOH[†]

In the information society, many systems are used for daily life, which results in generating big data. At higher educational institutions, there introduced many servers and terminals for supporting education. This equipment generates big data for education. In this manuscript, the characteristics and education of information society are discussed followed by the current status and challenges on big data for education, and learning analytics for improving teaching and learning at higher educational institutions.

1. はじめに

社会基盤の情報化が急速に進み、人々のコミュニケーションの方法が目に見える形で変化している。端末ではスマートフォンやタブレットが急速に普及し、それらを利用するコミュニケーション環境としてのTwitter、FacebookやLINEの利用者が増大している。このような急激な社会変化の中にあって高等教育をどう変革してゆくかは重要な問題となってくる。教育は次世代の人材を育てることがその目的であるが、育てた人材が活躍する社会環境が想像できない状況で、何を予め教えておけばよいかを議論することは極めて困難である。人間力や教養教育の重要性が強調されているのはこのような背景が一因であると考えられる。

一方で最近の学生はネティズンとかデジタルネイティブとか呼ばれている世代であり、物心ついたときからコンピュータが身近にあった世代である。学生にとっては教育も含めコンピュータを利用するのが当たりまえになっている。このような状況に対して教える側の教員は従来のやり方で講義がうまくいっているという自負があり、教員評価

が研究重視であることも相まって、教育改革に対する根本的な疑問がある。

この結果、大学が育成している人材が社会の要求に合わなくなってきている。大学改革の必要性が産業界を含めた広範な領域から叫ばれるようになってきたが、それに対して大学の動きが遅いのは教員側の意識の問題が大きい。大学は社会の動きに安易に追従してはならないという信念は必要であるが、単に改革を怠っていると社会から受け取られている。このため、国が主導して大学改革を進めようとしているが、これは大学にとっては自殺行為である。真の大学改革を、特に教育改革を大学が主体的に進めてゆかなければならない。

本稿ではこのような立場から、情報社会の特性と情報社会における教育の問題を議論した後、大学教育の情報化推進における問題点とビッグデータへの期待について述べ、今後の大学教育改革について考えてみたい。

2. 情報社会の特性

情報社会では、情報が社会の中心となる。既にWebには、膨大な量の情報が蓄積されており、今後のデータ量の増加は指数関数的であると予想されている。知識が大量に外化され、それが簡単にアクセスできる環境と個々人が携帯端末を持ち歩いていつでもどこでも他人とコミュニケーション

[†]京都大学情報環境機構長、CIO、学術情報メディアセンター
Director-General, Institute of Information Management and Communication, CIO
Kyoto University

ンできる環境は人間の活動を根本的に変える可能性がある。

人類が有史以来蓄積してきた情報は音声で伝えられてきた。文字と紙の発明により、情報の増加率は多少大きくなったが線形であり、そのまま情報社会を向かえる。情報社会ではコンピュータとインターネットにより、個人が情報を蓄積できるようになり、その増加率は指数関数的になった。ただし、蓄積されている情報の質は多様であり、正しい情報だけでなくさまざまなレベルの情報が混在している状況である。

情報社会では、人間のコミュニケーションの方法が従来とは大きく異なっている。対面コミュニケーションが中心であった時代は会わなければコミュニケーションが不可能であった。これを変化させたのは電信や電話であり、これにより場所の制約はなくなったが、時間の制約が残っていた。ポケベルから始まった時間制約克服の流れは携帯メールにより完成し、この情報基盤のもとに Twitter、Facebook や LINE などの SNS サービスが普及してきた。これによりいつでもどこでも自分の都合のいいときにコミュニケーションがとれるようになっていく。従来よりも多くの人、嗜好の合う人々とコミュニケーションができる状況になっている。現状では、物理的に集まっても端末を操作している人が多く、対面コミュニケーションが軽視される傾向にある。この傾向はそのうちに改善されるであろうと考えている。

情報社会はさまざまな社会基盤が情報システムによって支えられている社会である。情報は共有されるので、人間は情報を創造するよりも編集する傾向にある。人の書いたものを簡単にコピーができるようになり、これが人間の知的活動に大きな影響を及ぼしている。また、至る所にある情報システムがサービスを提供しているので、人間はそれらを利用するが、その時に操作ログが収集される。このログは人間の社会活動を計測するいわゆるビッグデータとして集積されて解析される。その結果を利用すれば、社会活動のエビデンスに基づいて社会を改革してゆくことが可能になる。

このように情報社会にはこれまででは考えられなかった特性がいろいろとあり、これらの変化を理解したうえで、今後のことを論じてゆかなければならない。

3. 情報社会における教育と自己能力認識

情報社会で生まれてきた子供たちに対してはどのような教育を行えばいいのだろうか？ 教育とは社会で生きてゆける能力を身に付けさせることが目的であるので、その内容を考えるにあたっては情報社会で必須となる能力について考える必要がある。

情報社会では、老若男女を問わずほとんどの人が携帯端末を持ち歩いていると想定できる。携帯端末には、電卓機

能、メモ機能、ワープロ機能など日常必要な能力を支援する多くの機能が搭載されている。計算機能が低コストで実現されている現在、ほとんどの人にとっては日常生活において単純な四則計算すら必要でない。従来からも微積や幾何学などの数学は社会生活に不用でないかとか漢字を書かせる必要はないのではないかなどの議論がなされてきている。これからは人類が蓄積している情報が急激に増加してゆくので、教えるべきことが従来に比べて極端に増加している。情報社会の特性は従来にはない新たな情報リテラシー教育、情報倫理教育、情報関連法律に関する教育や情報検索、情報の利活用法に関する教育を要請しており、子供たちの教育時間を考えれば現在行っている教育の1部を取りやめなければならなくなる。人間の教養とは何かを含めた広範な議論が必要であろう。

大学の教育現場ではすでにさまざまな徴候が現れている。学生のレポートは氾濫している情報を編集したものが多い。いわゆるコピペである。レポートを読んでいると論理が飛躍する部分や文体が変わる部分があるので、コピペはすぐにわかる。学生は提出する前にレポートを読み直したのか、読み直したとしたら論理的な破綻や文体の変化に気づかなかったのかという点が疑問である。自分の能力や成果を客観的に評価できていないというか、そういうことを教えられていないと想像できる。

コピペをするためには、膨大な量の情報からどこを対象とするかを選択しなければならぬ。情報の検索には検索ツールが利用できるが、検索結果のどれを選ぶかは重要な判断である。人間は分かったことしか発言できないという原則を仮定すると検索結果の中から自分が分かるものを選択するのが最も簡単である。学生のレポートは、キーワード毎に検索して自分が分かった部分のみを並べたものであると考えるとそれなりに納得できる。大量に情報が存在する社会での人間の取るべき最適戦略は、自分が分かる情報のみを利用する、言い換えるならば分からない情報を捨ててしまうことである。これはある意味で難解なことを考えないという態度につながる。

もう1つ気になる傾向がある。学生サービスとして単位認定に対する不服申し立て制度が創設された。この制度で不服を申し立ててくる学生の成績を調べてみると、ストレスで単位を落とした場合だけでなく全然できていないケースが多々みられる。試験やレポートで自分ができたかどうか分かっていないのではないかと疑いたくなる。自己の能力に対する自己認識は分からないことを理解しないと身につかないと思われるが、分かることのみを選択している環境ではこのような能力を身につける機会が減っていると考えられる。

4. 大学教育におけるエビデンスの収集

4.1 テーラメイド教育が可能になる？

情報社会で生まれ育った学生は、コンピュータの扱いに慣れており、小学生からタブレットを学習に活用しようとする試みも多く行われている。情報機器や情報システムを利用した教育（電子教科書や電子教材などを含む）を行えば、個々人の学習過程が観測できる。従来は、日常的な学習態度や学習過程は教員が主観的に観測し評価してきたが、これが客観的な観測データとして収集できるという点が情報技術を利用した教育の一つの利点である。情報技術を利用した教育はサイバー空間や仮想空間で行われる教育と呼ばれている。図1に示すように、学生は実世界でもサイバー世界でも継続的に学習するので、サイバー世界の学習過程だけでなく、実世界でも類するデータを収集し、これらのデータを統合して個人の学習過程のデータを時間軸に沿って収集することで、個々人の客観的な学習履歴が構築できる。



図1 物理世界と仮想世界での学習履歴の統合

Figure1 Integration of the learning data obtained in physical world and virtual world

収集された個人学習履歴は e-portfolio という形でまとめられる。新たな学習を始める場合は、これまでの個人学習履歴を参照して、どのような科目を取得することが望ましいかの相談に個別に乗ることができる。これは自分が受けた医療履歴を管理してそのデータに基づいて最適な医療を提供するのと同じ考え方である（図2参照）。



図2 個人学習履歴と個人医療履歴

Figure 2 Personal learning record and personal health record

個人の学習履歴に基づいた教育をテーラメイド教育と呼んでいる。教育は教室で行うことが現在では一般的であるが、江戸時代の寺小屋教育は個別指導であった。人間は、個人によりその能力がさまざまであるので、その能力にあった教育を行うことが理想である。情報技術はこのような教育を実施する基盤を提供できる。

就職活動においては応募してきた学生が今までどのように学習してきたか、他の学生と比べての強み弱みはどこにあるかなど学習過程に関する詳細な情報が記録された e-portfolio を企業が要求するという例もアメリカでは広まりつつある。日本においても企業は新規に採用した人材が短期間で転職するという問題に直面しており、今後このような動きが出てくる可能性がある。こうなると大学側も学生サービスの一環として教育関係の情報を収集し学生に提供しなければならない。

このように教育系のビッグデータは、学生サービスを充実させてその恩恵を学生個人に還元するという側面も強い。ビッグデータを解析して大学教育をエビデンスベースで改革する視点だけではないのである。

4.2 大学における教育の情報化

前節で述べたような教育改革、学生サービスを実行するための前提条件は、大学内で教育の情報化を進めることである。日本の大学の状況はアメリカと比べて10年遅れていると言われている。この要因の一つは、日本の大学における教員採用が研究業績に偏っており、教員が教育を重視していないことである。教員にとっては研究が本務であり、教育はどちらかという雑用となっている。

また、教育に情報技術を導入することの意義が情報技術を利用した教育の効率化と捉えられ、一つの選択肢と考えられている。したがって、行うかどうかは個人の問題である。この文脈では、教員自身の成功体験、すなわち自分が

受けた教育が最良であるという信念が優勢になり、従来通りの教育を行う根拠となっている。

文科省が毎年行っている学術情報基盤調査では多くの大学が CMS を導入していると回答している。同時に教育の情報化を進める基盤設備としての無線 LAN もほとんどの大学に導入されている。統計的に見れば情報環境は整っており、あとは教員が利用するだけという調査結果である。現場にいる情報系の教員としてはこの調査結果は納得がいかない。情報関係の部署や大学の本部事務としては、自分の大学が他大学に比べて遅れているとなると執行部から批判される。うまい具合に質問が「大学で CMS を導入して利用していますか」という形であるために、学内のどこかの学部にいる熱心な教員が個人的に立ち上げて利用しているだけでも利用していると返事をする。統計になればこれが大学として導入されているということになり、実態とかけはなれた調査結果になっていると思われる。本音と建前を使い分ける日本の慣習がなせる業である。

このような背景で問題点が発覚しないまま 10 年以上の歳月が流れたのであるが、ここきて教育系のビッグデータが議論されるようになってきた。教育の情報化が進んでいるのにビッグデータが存在しないという状況が明らかになりつつある。これに伴って、エビデンススペースという言葉が重要視されてきており、日本の教育の実態が明らかになってきたというのが現状であると認識している。

4.3 コンテンツデザイン

どこの大学にも教育に熱心な教員の方は多くおられる。これらの先生方は、情報技術を使った教育に興味の重要性に気づいておられ、いろいろと新たなやり方を模索しておられる。ところが、大学としてこれらの教員を支援する組織や制度が整っていないことが多い。大学の情報関係部署は、大学内の情報基盤整備をその役割にしている。いわゆる情報環境の提供である。京都大学でも CMS は全学的に導入している。情報関係部署の役割はシステムの使い方は説明するが、それを使ってどう教育するかまでは相談には乗れないし、その活動を支援することもできない。

情報技術を利用した教育を推進するために必須なものがデジタルコンテンツとしての教材である。電子教科書は、従来の教科書の執筆と同じ作業で作成できる。これは単に教科書を電子化したものであるが、それだけでも学習履歴を取得できるし、どの程度教科書が参照されたかもわかる。ここでいうデジタルコンテンツは、教材を提供するだけでなく、コンピュータを相手にした対話処理、学習過程における教員や TA とのコミュニケーションや小テストなども含むものである。教材に複数のストーリーを提供し小テストの理解度に応じて提示内容を変えていくという教材を作成すれば、ある程度のテーラメイド教育が個別の科目で実現できる。

このようなデジタルコンテンツは、図 3 に示すように、教員だけでなく教育法や教授法の専門家（エデュケイショナルストラテジスト）、プロジェクトマネージャの役割をもつインストラクショナルデザイナー、メディアデザイナー、情報（ICT）技術者の共同作業の結果として作成される。

この作業をやる気のある教員一人でやらせようとしているのが日本の大学の現状である。教員一人ではこのような作業は不可能なので、やる気のある教員でもあきらめてしまう。この状況を打破するためには、大学に教育支援組織を作ることである。ここでは、コンテンツデザインだけでなく、授業支援、チュータ教育なども行う。大学が効率化を求められている現状では、大学の努力だけではこのような組織を作ることは不可能であり、国の支援が必要である。研究力強化のために URA という制度を国が大学を支援して作らせたように、教育支援組織を大学に作らせて教育の情報化を推進しなければ教育改革は進まないと考えている。

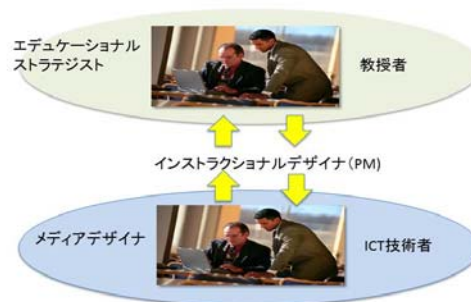


図 3 デジタルコンテンツの作成過程
 Figure3 Content creation process

4.4 MOOC

MOOC(Massive Open Online course)の概念は古くから提唱されていたが、大学で情報技術を使った教育が普及して現実味を帯びてきた。すなわち、日常行っている大学での情報技術を利用した授業を一般に公開した MOOC である。手間をかけて作成した教材を授業に出席している学生だけにしか提供しないことは、情報社会における情報共有の概念に反しているため、これは情報社会においては全く自然な発想である。ただ、一般公開して多くの学生が受講した場合、質問の対応や試験などは大規模になり、教室内の授業に比べてコストがかかる。それにもかかわらず物珍しさや社会的貢献の意識から多くの大学が MOOC を始めるようになってきている。

MOOC が多くの受講生を集められるというのは、情報社会においては利点になる。広告モデルが成立するからであ

る。様々なビジネスモデルが考案され、コーセラや edX など世界的に注目される MOOC プロバイダーが成長してきた。商業的にビジネスモデルは成立する可能性はあるが、大学が主体になる場合はコスト負担の問題があり、このままでは長続きしない。やる気があり能力のある学生が発掘できるという魅力はあるが、面倒見の良い日本の教育という意味では成立はしない。

edX はトップクラスの大学の教育コミュニティを目指すという理念があり、大学連携に重点が置かれている。情報社会における教育を世界的に議論してゆくことを重視しているが、運営主体は NPO 法人であり、必要経費をどう負担してゆかが今後は問題になってくる。

日本でも JMOOC が設立され、講義が始められているが、独自のビジネスモデルが必要である。企業がスポンサーになり、企業研修などの実用的なプログラムを実施することが議論されているのはいいが、大学の役割を明確にして継続性のある体制を作らなければならない。

MOOC で集められたデータは、世界の様々な人々に関する学習過程のデータであり、地球規模のビッグデータであることが興味深い。地域ごと文化ごとにどのような特色があるか、日本の教育をその中で相対化し改善を進めていくうえでも重要なデータとなる。ただし、やる気があり能力のある学生だけのデータであるという点に注意しなければならない。

4.5 学生の個人情報

学生個人の学習履歴を収集するには学生から個人情報提供の同意を得なければならないというのが現在の法規制である。教員個人が授業で学習履歴をとるために学生から個人情報提供の許可を得ることは可能であるが、大学業務の一環として大学が学生の個人学習履歴を収集する場合の規則を明確にしておかないと、今後、教育の情報化は進められない。

社会は、人間が活動する場であり、社会を考えるに当たって人間の要素は排除できない。現在、国において個人情報保護法の改定の作業が始まっており、商業活動に利用するために匿名化した個人情報を活用しようとする方向で議論が進められている。人間の生涯は社会活動の準備期間である教育期間、社会活動期間、老齢期間があり、それぞれで個人情報の扱いを検討するべきだと考えている。法律でこの種のことを規定するのは困難であるので、自主的な規定を作って公表し、それを遵守するというソフトローの考え方を適用し、大学や介護施設が業務として行う場合の個人情報の扱いを自主的に定めていくことが重要である。これらの問題は、情報処理学会の理事会のもとに設けられた「個人情報利活用 WG」で議論されている[1]。

大学として教育の情報化を進める場合、個人情報の提供を学生が拒否した場合は教育が不可能になる。したがって、

大学に入学することは個人情報を提供することに同意したという契約を大学が学生と結ぶという考え方が必須である。ただし、大学が個人情報をどのような目的で利活用するのか、どう管理しているのかを学生に対して公表し、疑問やクレームがある場合はそれに対処する問い合わせ窓口を設置しなければならない。それぞれの大学において、この種の規則を設定することで教育関連ビッグデータを収集し、データ解析をしてエビデンスベースの教育改革を進めるべきである。

4.6 大学間でのデータ共有

学習履歴は個人ごとに収集される。大学という組織には多数の学生が在籍しているので、これらのデータを群として集計した統計解析ができる。

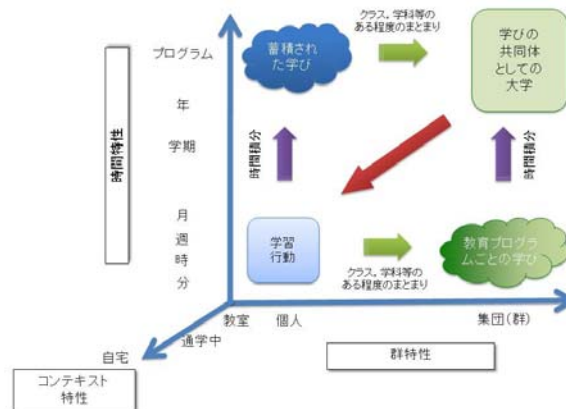


図4 個人学習履歴の集積軸

Figure 4 Coordinate to integrate personal learning records

図4に示すように個人学習履歴を集計するにあたっては、時間特性の軸、群特性の軸とコンテキスト特性の軸が設定でき、それぞれの軸にそったデータ解析が可能になる。

個人学習履歴の集積は1つの大学にとどまるだけでなく、大学間での情報集積、国全体で統計データ集積、さらには世界での統計データの集積を可能にする。データの記述形式を標準化してデータ集積の仕組みをクラウドサービスとして実現できる。クラウドの形態としては、アカデミッククラウドの調査研究[2]において述べられている拠点型が適するであろう。国内の地域ごとにデータを収集解析し大学連合を作って地域間で競い合う構図になるだけでなく、学生が物理的に動ける範囲の大学が連合して単位互換などの制度を作りやすくすること、学生の交流を深めることなどの利点がある。

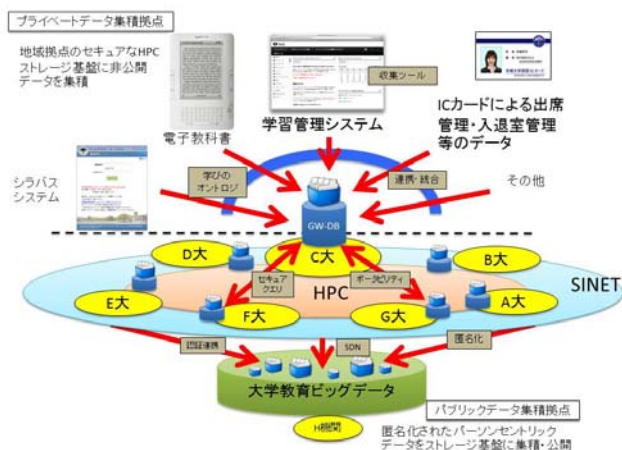


図5 アカデミッククラウドの実現形態
 Figure 5 A realization of academic cloud

具体的な実現形態としては、基盤センターが中心となり地域の大学を連合したシステムを構築し、それらが連携して日本全体の教育ビッグデータを収集する。これを国レベルでも解析し、国際比較できる体制を構築する。

この時に問題となるのが個人情報の第三者提供である。匿名性を担保して拠点や国に対して学生の学習履歴情報を提供することの安全性を系統的に担保したうえで、大学として学生の個人情報を提供することを決定しておく必要がある。このような枠組みは、大学が主導して構築する必要があり、大学が真剣に議論しなければならない。

5. ビッグデータで何がわかるか

前節までで述べた方法で収集される教育ビッグデータの内容を考えてみよう。一般的にはCMSで収集できるのはユーザの操作ログである。ログイン、ログアウトした時刻、閲覧したページとその順序、それぞれのページが表示されていた時間、教材側で行われた小テストの回答にかかった時間や成績などが含まれる。これに加えて、掲示板への書き込み、提出されたレポートやその評価、Q&Aなど教育内容に関わるデータも場合によっては収集可能である。e-portfolioのシステムを導入すれば、ルーブリックの情報とそれに対する学生の学習活動などが記録可能である。デジタルコンテンツの作成方法を工夫すればさらに多くのデータが収集できる。

PCやタブレットなどで収集できる端末のログデータも有用である。CMS以外のどんなソフトを利用したかがわかれば、教材関連のことを調査したのかどうか、どのようなソフトを起動して何を調べたのかなども場合によっては記録できる。ただし、端末のログは学習以外での活動の記録も含むので、個人情報の問題が微妙になってくる。

このようなデータを解析して何が知りたいのかを明確に

してゆくことが今後求められている。ビッグデータの研究は、宝探しのようイメージが進められることが多いが、このような形ではあまり成功しないと考えられる。学生に対してテラメイド教育を行いたいという抽象的な議論は理解できるが、それでは何がわかればテラメイド教育ができるのか、そのためにどんなデータを集めるべきかなど議論を具体化してゆかなければいつまでたってもこの分野の研究は進まない。海外の例を参照するのはいいことだが、日本の教育方法は海外とは文化的に異なっているので、現場で今必要なことを明確にしてゆく必要がある。

情報社会での必要な教育で自己認識能力の議論をしたが、この能力が身についているかどうかをビッグデータ解析で判定できれば大変有効である。操作ログだけではこのような能力の判定は困難であるので、提出されたレポートや小テスト、ディスカッションにおける発言などを統合的に解析しなければならない。自然言語処理を中心とするメディア処理技術を駆使して、この種のデータ解析の研究を進めてゆく必要がある。

6. おわりに

本稿では、教育ビッグデータに関する様々な問題やその解決方法について議論した。教育を改善するためには、エビデンスが必要であるが、どのように改善してゆくのか、どうあるべきかの議論が欠けているような気がする。教育の将来に対するビジョンを議論し、その方向へもっていくために教育ビッグデータを収集し、解析するという方法の重要性を強調した。

現在、議論されているイノベーションの本質は情報技術をいろいろな分野に導入することである。情報社会になり情報技術が社会のあらゆる分野に浸透してきた状況において、情報社会の本質を見据えた改革がイノベーションの源泉である。情報の研究者や技術者が中心になって改革するという意気込みで、多くの大学教員を引き込んで大学改革を進めていかなければ日本の教育に未来はないと思われる。今後の展開に期待したい。

謝辞 図を提供いただいた京都大学情報環境機構梶田正司教授に謹んで感謝の意を表する。

参考文献

- 1) 情報処理学会個人情報活用WG
www.ipsj.or.jp/event/sj/sj2014/itforum_kojinjochoWG_program.html
- 2) アカデミッククラウド環境構築に関わるシステム研究
www.icer.kyushu-u.ac.jp/ac