

我が国の未来を担うデータサイエンティストの育成 —政策の動向と滋賀大学の挑戦—

須江 雅彦^{†1}

概要：本年(2017)年4月、日本初の「データサイエンス学部」が滋賀大学に誕生します。この学部は、現代社会に不可欠なデータ分析と価値創造を担うプロフェッショナル、データサイエンティストを組織的に育成するため学部です。インターネット商用開始後この四半世紀にわたる ICT の劇的な進化により出現した、いわゆるビッグデータを様々な価値創造につなげていく上で大きな役割を果たすデータサイエンティストが、我が国では少なく、人材不足は危機的状況にあると言われていました。一方、我が国ではデータ利活用により付加価値を生み出す新事業・新サービスの創出を柱とする第4次産業革命を実現するという方針が示され、様々な政策が動き出しています。これに向けた人材育成に関する取組みを中心に政策の動向や課題をみていくとともに、我が国初のデータサイエンス学部創設とその教育研究拠点を形成する滋賀大学の取組みなどについて、私見を交えながら述べていきます。

キーワード：データサイエンス、データサイエンティスト、第4次産業革命、データサイエンス学部、データサイエンス教育研究拠点、PBL 重視の滋賀大モデル、価値創造プロジェクト

1. はじめに

平成 29 (2017) 年 4 月、日本初の「データサイエンス学部」が国立大学法人「滋賀大学」に誕生します。この学部は、現代社会に不可欠で、我が国の未来のために極めて重要な価値創造のための新たな科学、「データサイエンス」に関し、本学が形成する日本初の教育研究拠点の中核をなすもので、平成 28 年本年 4 月に設けたデータサイエンス教育研究センターと協働し、関連する様々な教育研究活動とともに、データ分析と価値創造を担うプロフェッショナル、データサイエンティストの組織的育成を行うための学部です。

1.1 ビッグデータ時代の進化

インターネットの商用利用開始後、コンピュータ性能の劇的な向上を伴う ICT (情報通信技術) の飛躍的發展により、データ利用の高度化は大きく進行しました。今やタブレット端末やスマートフォンなどによってユビキタスネットワーク環境が整備され、いつでも、どこでも、だれでも、なんでもネットワークにつながる時代が到来しています。

こうしたネットとの接続などを通じ、日常業務の処理、POS など種々の電子機器の活用、様々なセンサーや IC カード等の利用拡大が進むとともに、スマートフォンとアプリケーション利用から SNS をはじめ行動履歴など個人の様々な情報も集積されるようになっていきます。音声・画像・テキストなどを含む多種多様で膨大な量のデータが日々刻々と生成され、その集積が進むビッグデータ時代は更に進化し続け、今日、社会のあらゆる領域においてデータは新たな開発可能性を秘めた豊かな資源となっています。

世界をけん引している先端企業のアップル、マイクロソフト、グーグルなどはいずれも、ビッグデータとそれを分

析・処理する情報関連技術を駆使して、様々なサービスを構築し、世界の名だたる製造業を超えるほどの莫大な価値を生み出しています。我が国においても、データ利活用による付加価値の創出は、SNS をはじめ、テラーメイド型の個人サービスやレコメンドサービスなど新サービス、マーケティングやビジネス運営の効率化、交通システムや電力供給システムの効率化などの社会インフラのスマート化、自然災害の予測や防災など、枚挙に暇がありません。データの利活用による新たな付加価値の創出が少なくとも 60 兆円に上ると推計されており、データの利活用が進むほど更に増大するものと期待されています[3]。

さらに先進各国では、公的機関が保有するデータの公開を進め、人々との共有を通じてその運営の効率化や新しい価値の創造等を実現するとの観点から、オープンデータへの取組みもなされています。

1.2 データサイエンス

現代社会は、こうしてビッグデータのような様々なデータの集積を開発可能性のある豊かで新たな資源とし、その分析を通じて新しい発見を行い、それに基づく最適化や利便性の向上、新事業の創出などが大きな価値の創造につながる時代となっているのです。

こうしたビッグデータから有意味な知見を導くための新たな科学あるいは学問領域が、「データサイエンス」(Data Science) です。データサイエンスは、ICT の発展を背景とする最先端の計算科学であり、「大規模データに語る」(データドリブン) という意味で帰納的な特徴をもっています。通常、①データエンジニアリングと②データアナリシスから成っており、前者は大規模データを加工・研磨・処理するための専門知識とスキルであり、後者はデータ駆

^{†1} 滋賀大学 理事/副学長 (連絡先: masahiko-sue@biwako.shiga-u.ac.jp)

動型の多様な分析・解析の専門知識とスキルです。

データサイエンスは、現代社会において日々集積される多種多様で膨大なビッグデータを可視化し、そこから新たな知見、価値ある情報を引き出し(発見)、価値を創造するための科学として、その最先端では、統計分析と情報学やコンピュータ科学が融合しています。

1.3 データサイエンティスト

現代社会に不可欠とされるデータ分析の専門家としてデータサイエンティストは、ビジネスや政策など様々な領域において、それぞれの課題を読み取り、関連する様々なデータをデータエンジニアリングとデータアナリシスの専門知識とスキルを駆使して分析を行い、引き出した発見を実際の意思決定に活かして最適化や効率化を行う、また、新事業創出などの課題解決に結びつけていくことが期待されています。データサイエンティストへの社会的要請は、データサイエンスのコアな専門知識とスキルに加え、領域科学分野の知識を前提とする「価値創造」の力量の持ち主としての役割にあるといえます。

平成 28 年 1 月に策定された第 5 期の「科学技術基本計画」では、第 2 章「未来の産業創造と社会変革に向けた新たな価値創出の取組」において競争力の維持・強化の観点から、「超スマート社会サービスプラットフォームを活用し、新しい価値やサービスを生み出す事業の創出や新しい事業モデルを構築できる人材、データ解析やプログラミング等の基本的知識を持ちつつ、ビッグデータや AI 等の基盤技術を新しい課題の発見・解決に活用できる人材などの強化を図る」とされています。

ビッグデータ時代の価値創造の担い手であり、こうした大変革の時代のイノベーションの牽引役であると同時に、データ活用の社会実装推進の役割も果たす「データサイエンティスト」に対する社会的要請は極めて高いのです[2].

2. データサイエンティストの深刻な不足

こうしたビッグデータ時代の新たな価値創造への期待、そして専門人材への期待といった社会的背景の盛り上がりとは対照的に、わが国ではデータサイエンティストの育成が現状では中々進んでいません。わが国では、欧米等と比較し、データ分析のスキルを持ち、統計科学を理解・応用のできる人材が極めて少ないのが現状で、科学研究のみならず多くの組織や企業でもこうした高度人材の不足が顕著で、危機的な状況にあると政府も指摘しています[1].

筆者は、データ分析人材育成の日本の著しい立ち遅れは、欧米と異なり、我が国の大学には統計系の学部・学科がないことが大きな要因だと考えています。

2.1 欧米等の人材育成

欧米(中国、韓国等も同様)には、従来から統計科学の独立した学部・学科が存在しています。独立した統計科学の教育研究部門があると、情報学、コンピュータ科学、数学も加えて、データサイエンスの教育プログラムが柔軟に生成されやすいと考えられます。実際、アメリカの学部教育では、統計学専攻のプログラムがデータサイエンスを意識した内容に変更されてきており、大学院では、各領域科学の専攻を中心に副専攻プログラムとしてのデータサイエンス教育が展開され、或いは企業の即戦力としての専門職育成のためにデータサイエンスを主専攻とする教育が行われたりしています。

これは ICT の進化に即して、米国の統計教育研究が、領域科学分野の研究者とのデータの分析・利用に係る交流を通じて複合領域化し、広がりを持った形で学部自体が進化していることを示しています。統計学部の学生はその発展の成果を含めて教育を受け、更に大学院で学ぶなどして、社会でデータ分析の専門家として活躍し、更に高度な経験を積み、統計学やデータサイエンスの社会実装に貢献しているわけです。

近年米国では統計学関係の学位は人気も高く、取得者が急増しています。

アメリカ統計学会ニュースレター 2016年10月号

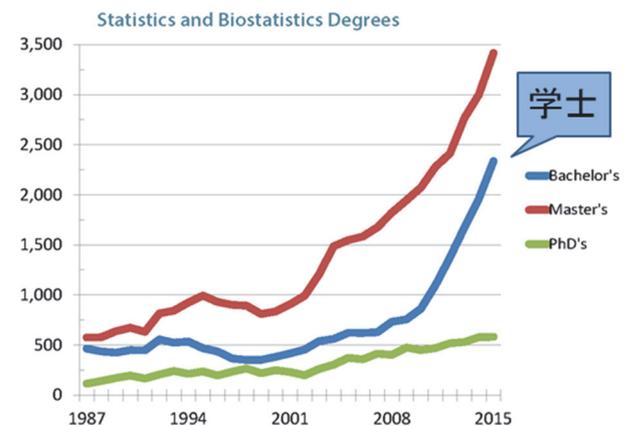


図1 米国における統計関係学位取得者の推移

さらに欧米では、データサイエンスの学部教育プログラムを整備し、独立した学位を授与する大学も登場しています。(イギリスのウォリック大学、アメリカのオハイオ州立大学、ロチェスター大学等)

2.2 日本の状況

我が国は、今現在、ビッグデータ時代に対応したデータ分析の専門家育成のための本格的な学部教育プログラムがありません。(来年滋賀大にできるデータサイエンス学部が

日本初の本格的な学部教育プログラムを提供)

また、わが国における統計科学の教育研究は、専門の学部や学科を設置せずに、経済学、理学、工学、医学、生物学、心理学などの「分野点在型」で行われてきました。これには固有領域データの取扱いの点で実践的なメリットもありましたが、現在のようにICTの進化によって必要とされる技術スキルが急速に進化・変容していく時代には対応が遅れがちにならざるをえず、特にオープンデータを含め、様々な種類のデータを、手法を駆使し同時に分析し新たな発見を行うといった面での人材育成には適さないと考えられます。データサイエンティストのような高度なデータ分析力を持つためには、コンピュータ科学や情報学だけでは不十分で、統計学に関連した幅広い多様なデータ分析手法などデータアナリシスの十分な能力を理論面・実践面から身に付けることが必要なのです。

また、そもそもこの分野点在型の方式では、大学において本格的なデータ分析人材の育成を組織的に行うことは極めて難しいと言わざるを得ません。

3. データサイエンティスト育成策の展開

3.1 統計教育の改革

米国では、1980年代の日本研究、特に統計的品質管理(KAIZEN)から学び、数学や統計学の能力を国民に身に付けさせるために、初等中等教育での統計教育の充実と、高等教育での産業界との連携が進められ、インターネットの普及とコンピュータ技術、性能の飛躍的発展を背景に生み出されたデータを、ビジネスや行政の様々な場面で活用し、価値創造につなげていきました。その過程で米国に情報産業は飛躍的発展を遂げ、今日の情報社会のプラットフォームをいくつも担い、大きな利益を上げています。

日本でも、産業界や統計学会関係者などの尽力により、初等中等教育に統計教育が改めて位置づけられ平成22(2010)年から順次再導入され、その教育を受けた高校生が平成27(2015)年から卒業しています。米国と比較し約20年遅れながらも、初等中等教育は動き出し、プログラミングなどの情報教育も始められようとしています。

しかし、高校で統計を学ぶ機会があっても、大学の学部がないという状態は解消しておらず、滋賀大学が来年開設するデータサイエンス学部が初の統計系学部となります。

また日本統計学会が中心となり平成23(2010)年から「統計検定」が開始されており、平成24(2011)年からは文部科学省の支援を受け統計教育大学間連携(JINSE)の活動も進められ、データに基づく課題解決型人材育成に資する統計教育質保障に向けて、様々な取り組みがなされています。

一方、MOOCという最新の教育手法を活用した「統計学」の講座や「社会人のためのデータサイエンス」講座などが

日本統計学会や総務省統計局から提供され、データ分析を学ぼうとする意欲ある学生や社会人が多く受講しており

双方の累計で10万人近くに上るものと見込まれています[4]。

こうした意欲を持つ学生・社会人に対して、大学はどのようにデータサイエンスの専門知識とスキルを学ぶプログラムを提供し、社会に有用な高度な人材に育成していくのか、対応を真剣に考え関係者は行動する必要があります。

3.2 第4次産業革命実現に向けての人材育成

平成28年6月に策定された政府の「日本再興戦略2016」においては「第4次産業革命」の実現が大きな政策課題となっています。第4次産業革命としてデータ利活用により付加価値を生み出す新事業・サービスの創出が重要であるとの考えが示されており、この革命を支える基盤技術として、AI(人工知能)、ビッグデータ、IoTが位置づけられています。これらはいずれもデータ解析力が重要であることを示しており、そのことがわが国の経済成長力あるいは国力を大きく左右することになるという認識といえましょう。

政府は第4次産業革命に向けた人材育成のための総合イニシアチブとして「未来社会を創造するためのAI/IoT/ビッグデータ等を牽引する人材育成総合プログラム」を提示しています。

この包括的なプログラムでは、基盤技術であるデータサイエンス等の人材育成・確保に資する施策を、初等中等教育、高等教育から研究者レベルまで体系的に実施し、様々な施策によってデータサイエンス人材を産業界に輩出する考えを示しています(図2)。

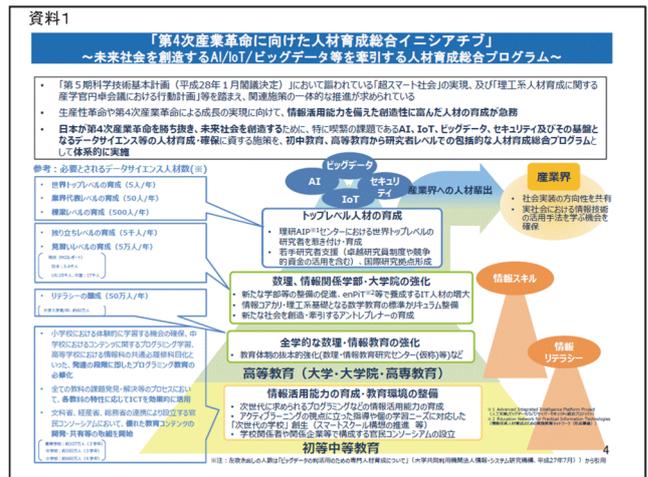


図2 第4次産業革命に向けた人材育成総合イニシアチブ (出典：文部科学省)

その主な内容は次の通りです。

- ① 初等中等教育においては次代に求められる情報活用能力の育成・教育環境の整備を行う。特にプログラミング教育については2020年から初等中等教育の発展段階に即し必修化を行う。その際、民間のノウハウ・人材等を活用したコンテンツ本位の学校情報化を推進
- ② 高等教育においては、まず、教養教育での全学的な数理・情報教育の学修の強化を図り、全大学生50万人にリテラシーを醸成 等
- ③ また、データサイエンス専門人材育成のため、滋賀大学が来年開設するデータサイエンス学部のような新たな学部整備、大学院の強化の促進。数理・情報分野の専門教育への重点支援：実践教育を行う産学連携ネットワークの構築 等
- ④ ハイレベルの人材育成として、理研AIPセンター(人工知能/ビッグデータ/IoT/サイバーセキュリティー統合プロジェクト)におけるトップレベルの研究者育成、ポスドク等の若手人材に対するデータサイエンス等の高度なスキルセット獲得機会の提供、キャリア開発支援

これらの施策はデータサイエンティストの育成にも意欲的に取り組むためのものであり、今後の展開が大いに期待される所です。

ただし、現時点で産業界が欲している問題解決力を持つデータサイエンティストの不足分をどのように質的・量的に確保できるのかについては、必ずしも明確ではありません。

3.3 足元の不足への対応

従来からデータ分析人材の必要性は繰り返し表明されてきましたが、大学教育ではそれに対応した学部がないことから、企業は、今まで、専門外或いは近い専門分野の社員について、必要に応じ企業内教育などを中心に人材養成を行ってきたという側面が強いのですが、今日のように急速な技術変化の中では社内教育だけで十分なスキルアップを図るということには相当困難な面があると思われます。現実の不足を部分的にせよ速やかに補うという点ではやはり、企業内のデータ分析人材の高度化が必要ですが、これに関して専門の大学や企業等との連携による高度再教育プログラム構築が喫緊の課題であり、産学双方の連携体制の整備と企業側からの投資を急ぐ必要があります。

またいわゆるポスドクのデータサイエンスのキャリア開発支援による活用といったことも考えられていますが、米国の事例 (Insight Data Science Follow Program) のようなものとなるかは実務面で課題が多いと思われます。いずれにしても、一日でも早く、データからものごとを捉え、課

題解決に結び付ける分析能力の高い者の育成を使命とする専門学部の設置、充実を進めることが、我が国の関連人材供給のベースを広げる上で最も望まれます。

なお、我が国では専門職としてのデータサイエンティストについて、一般企業や行政では内部での位置づけが処遇を含めてまだ明確でないと思われませんが、その能力を活用していくためには、こうした処遇面での改善も必要となります。

3.4 統計分野の研究者の確保

データ分析手法は多様であり、一定規模の研究者集団を意図的に形成していかないと、組織的な専門教育や研究水準の更なる高度化は難しいと思われれます。その際、統計分野の専門家の数が少ない現状をどう乗り越えていくかは重要な課題となります。専門学部がない中では統計研究者の量的育成も進んでいないことにも留意する必要があります。米国では毎年600名以上の統計学分野の博士が授与されますが、我が国では年間数名程度に過ぎません。

3.5 人工知能, AIに関連して

近年注目を集めているAI、いわゆる人工知能も、ビッグデータ解析に基づくものであり、データ解析の専門性を有する人材がその研究開発、性能向上、社会実装において重要な役割を果たすと考えられます。

AIに関しては、平成28(2016)年4月に政府に人工知能技術戦略会議が設けられています。その下にある「研究連絡会議」と「産学連携会議」両者とも人材育成を議論の対象としていますが、後者の人材育成タスクフォース (TF) ではAI時代の即戦力育成としてデータサイエンティスト等について検討が進められています。

また、平成28年12月からは別途「第4次産業革命人材育成推進会議」が政府の構造改革徹底推進会合の下に設けられ、有識者を交えた議論が行われています。

AIに関しては、報道等ではあたかも「あらゆるものがAIで自動化、判断も」といった風潮が強く、「AI=ロボット」ように感じている人々も多い状況があり、こうした局面で、これからの会議がどのような方向性を打ち出してくるのか注目されます。

3.6 EBPM : Evidence Based Policy Making から

データサイエンスの社会的応用範囲はあらゆる経済社会活動に及ぶものであり、第4次産業革命の文脈とは少し異なりますが、平成28(2016)年12月の経済財政諮問会議においても、GDP統計の精度の議論に際し、エビデンスベースドポリシーメイキング (EBPM) の環境整備等の観点から「官庁データサイエンティスト」の育成が必要といった議論も行われています。地域創生に関連して、地方自治

体でもデータサイエンティストが必要という声も高い。

データサイエンスの社会実装を進め、社会運営の合理性・効率化高めるためにも本格的な人材育成の量的拡大を着実に進めることが必要です。

4. 滋賀大学データサイエンス学部創設のチャレンジ

4.1 日本初のデータサイエンス学部

滋賀大学は、1節、2節で述べたような現代社会に不可欠であるにも関わらず我が国で大きく不足しているデータサイエンスに係る人材を供給するという社会的な要請に応えるため、早急に本格的な学部教育プログラムを整備し、相当程度のスケールで教育を開始する必要があるとの観点から、平成26年(2014)秋、日本初のデータサイエンス学部を設置することを決めました。この取組みは、人文社会系大学の文理融合型大学への転換に向けた先進的な大学改革として高い評価を受け、文部科学省から大学改革強化推進補助金(平成27年度から)を受け、諸準備を進め、同年度末の平成28年3月に同省に学部設置申請を行い、同年8月26日承認されました。

平成29年4月に開設するデータサイエンス学部は1学年100名であり、①大規模なデータを加工、処理する情報技術(データエンジニアリング)と、②多様なデータを分析、解析する統計技術(データアナリシス)に加え、③ビジネスや政策など様々な領域における課題を読み取り、データ分析による知見を活かして様々な課題を解決していく価値創造スキルを身に付けた人材、データサイエンティストの養成を目指しています。

この価値創造スキルを身に付けるために、ビジネスや政策などの現場の実際のデータを用いたデータ駆動型演習を重視し、1年生から4年生までPBL演習教育を繰り返し行い、データを通じた問題解決の実践力を養う予定です。データサイエンスの価値創造は、ビジネス・政策・科学など多様なフィールドがあるため、複数領域のPBL演習を、フェーズを進化させながら繰り返すことにより、様々な手法を体験させ、分析能力の向上に役立たせます。

実際のビジネスや政策等の価値創造の現場で取り扱われるデータの多くは、いわゆる人間・社会・企業の3領域におけるデータであり、データサイエンスの専門知識とスキルを応用してデータ解析を行い、領域科学の知見を活かして価値創造に挑戦するためには、文系・理系双方の知見が必要です。この意味でデータサイエンスは「文理融合」領域といえます。このため、受験生に対しては、統計とコンピュータを社会的な課題に応用し、社会に貢献したいという文理融合的な意欲的な人材を求めると呼びかけています(図3、図4)。

4.2 データサイエンス教育研究拠点の形成

現在、学部には先立ち平成28年4月に設けたデータサイエンス教育研究センター(センター長:竹村彰通)において、既に様々な活動を開始しています。

同センターは、データサイエンスに関する基盤研究、企業等との価値創造プロジェクトの推進、教育プログラムやPBL演習教材・MOOCなどの開発、ワークショップや国際会議等を通じた研究交流・情報発信の4つの機能を有しています。

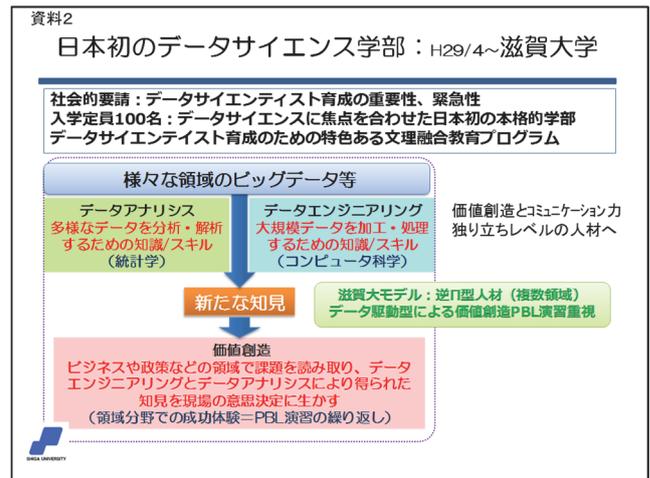


図3 日本初のデータサイエンス学部

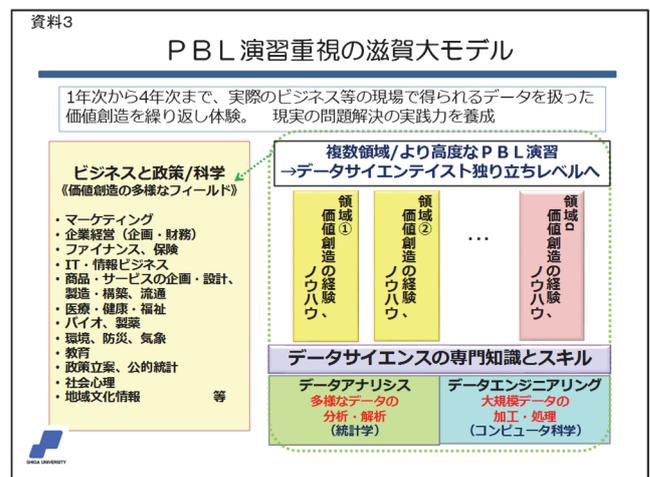


図4 PBL演習重視の滋賀大モデル

同センターの活動に当たっては、内外の大学や統計教育大学間連携ネットワーク(JINSE)をはじめ、統計数理研究所、総務省統計研修所、(独)統計センターなど国の機関、近隣自治体やデータサイエンティスト協会、様々な企業等との連携を重視しながら進めており、データサイエンスの学部教育を含めた日本初のデータサイエンス教育研究拠点を形成していきます。

特にデータサイエンス教育に関しては、我が国で先行事例のない最先端の教育プログラムを開発し、データサイエンス教育の標準化とPDCAサイクルによる教育の質保証システムを推進することとしています。データサイエンス教育の教材・教授法には未だ定型化された参照標準がなく、本学では、特に次の3つの観点で先進事例を創出し、データサイエンス教育を行う後続の大学が利用できるように努め、拠点としての機能を果たしてまいります(図5)。

- 標準的な「データサイエンス基礎教育」の教材と教授法を開発し、また内容に応じ順次eラーニングとして提供
- データサイエンス教育に固有の機械学習、最適化等の最先端の専門教育に焦点を合わせた教材・教授法を開発し、その教育効果を検証し改善
- データサイエンティストの育成では、実際のデータを利用した「データ駆動型価値創造 PBL 演習」での多様な成功体験の積み重ねが重要であり、このため、様々な領域のデータを活用した PBL 演習教材・教授法を開発

点から、以下のような企業等との多様な連携を重視しています。

- 教育面では：
 - 実務家によるデータ利用・解析事例に関する講義、PBL 演習への実務データの提供・共同開発、学生の実習・インターンシップ、データ利用/分析環境やプラットフォームの活用、企業等人材の高度化教育への参画 等
- 研究面では：
 - 企業等の価値創造プロジェクト(共同研究等)の推進、最新のデータ分析手法に関する助言、研究開発地域創生のためのデータ利用の研究 等

4.4 まとめ

滋賀大学のデータサイエンス学部を中核とするデータサイエンス教育研究拠点形成には、我が国経済の発展に貢献する取組みとして、政府や企業等からの多くの期待が寄せられています。

平成28年12月、文科省から高等教育におけるデータサイエンス教育強化を先導的に貢献する拠点大学として、東京大学などとともに選定されました。

滋賀大学は、専門学部を擁する唯一の先行拠点として、この未来志向のチャレンジを通じて、若きデータサイエンティストの育成はもとより、企業等人材の高度化、企業等との価値創造プロジェクトの推進、新たな分析手法の開発など、社会に貢献する大学を目指してまいります。

関係の皆様のご理解、ご支援をよろしくお願い申し上げます。

参考文献

- [1] 内閣府、「科学技術イノベーション総合政策2015」、内閣府(2015)。 <http://www8.cao.go.jp/cstp/sogosenryaku/2015.html> (2017年1月30日アクセス)
- [2] 日本学術会議情報学委員会 E-サイエンス；データ中心科学分科会、「提言 ビッグデータ時代に対応する人材の育成」(2014)。 <http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-22-t198-2.pdf> (2017年1月30日アクセス)
- [3] 総務省、「平成26年度版 情報通信白書」、総務省(2014)。
- [4] gacco, <http://gacco.org/> (2017年1月30日アクセス)

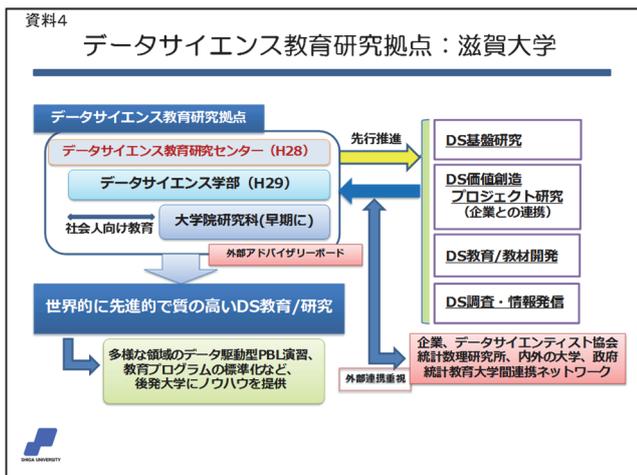


図5 データサイエンス教育研究拠点

4.3 企業等との多様な連携

滋賀大学ではこのデータサイエンス教育研究拠点の運営に当たっては、価値創造に貢献する教育・研究という観