

データサイエンスと滋賀大学の活動

竹村 彰通
(滋賀大学学長)

*編注

- ・司会進行関係の発言は文字化していない
- ・小見出しは当方で読みやすさを考慮して勝手に入れたものですので削除または適宜ご訂正ください
- ・スライドはタイトルスライドを1として20まで全スライドをナンバリングした。
- ・スライドは本文に、言及された箇所に挿入する形とした。
- ・全体的にリモートのせいか音声不明瞭の箇所が多々あったが可能な限り文字化した。



データサイエンスと滋賀大学の活動

竹村 彰通
滋賀大学学長

2022年9月15日
PHR協会Zoom講演会

はじめに

データサイエンス、そして滋賀大学のご紹介も少しさせていただきます（スライド1）。

スライド2のような内容でお話ししたいと思います。

- ビッグデータとデータサイエンス(DS)
 - 数理・DS・AIの動向
 - データサイエンスの3要素
 - データからの価値創造の流れ
- 滋賀大学DS学部と大学院の教育

2

ビッグデータとデータサイエンス (DS)

ビッグデータとデータサイエンス (DS)

データサイエンス:ビッグデータを対象として、そこから新たな知見を引き出し、価値を創造するための新たな科学



3

今はビッグデータの時代ということで、このスライド3はビッグデータが集る様子を示しているのですが、われわれは10年ほど前から使っていますのは、スマートフォンですが、その後も日常的に使うようになりました。デジタルデータとして集ったデータを利用して、新たな価値を生み出そうということで、それがデータサイエンスということです。そういうことがやられ、重要になっているということです。

例：ウェアラブルデバイスの進歩



4

今日の話はパーソナル・ヘルス・レコードということで、スライド4は個人の健康データですが、そういうものがテーマなのですが、それに関してアイフォンの新しい機種が出たりして、血中酸素が出たり、心電図アプリの機能なども普通に使えるようになるのかなと思っています。

このような時代で、データが資源で価値を持っているということは、しばらく前から、21世紀は石油ではなく、このデータが価値の源泉であるということで、「21世紀の石油」という言い方をされるようになっていきます（スライド5）。

21世紀の石油

The Economist 誌 2017年5月6日

“世界の最も価値ある資源はもはや石油ではなくデータだ”



インターネット巨大
企業を警戒する論
調も強くなってきた

5

それはスマートフォンを使っていると感じるのですが、データを生かすも殺すも企業、とくにアメリカのGAFA（「Google」「Apple」「Facebook」「Amazon」）などが世界を席巻すると言いますか、ブラックホールになって、日本は全体的に遅れていて、とくにデータを扱う人材が不足している状態だと思います。もちろんまだまだ日本はデジタルトランスフォーメーションが進んでいない状況だと考えています（スライド6）。

ビッグデータ＝新たな資源

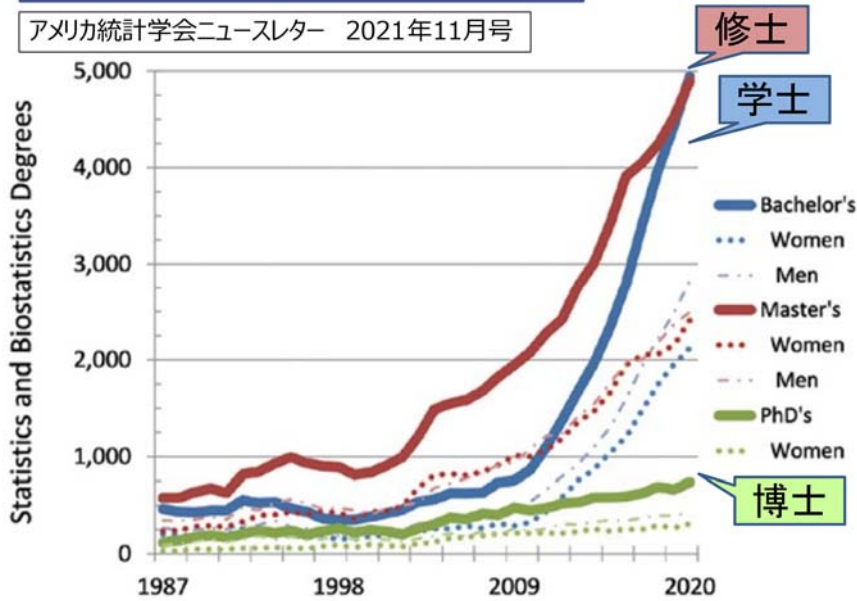
“21世紀の石油”

- 人々の行動履歴が直接観測できるようになったことが大きな変化
- 今後はモノや環境からも大量のデータ(IoT)
- この新たな資源を生かしたものが競争的優位に立つ
- データとそれを生かす技術の双方が必要
- データだけあっても、分析できなければ宝の持ち腐れ
- 日本は、データ自体を外国企業にとられ、活用もされてしまっている状況 ← ネットワーク効果

6

この人材の育成なのですが、実はアメリカではデータを分析する人の認知度が高いと言いますか、私の専門は統計学なのですが、スライド7にありますようにここ10年で学位の授与者が3倍ぐらいに増えているという状況があります。コンピュータサイエンス、さらに最近はデータサイエンスで、これは統計学の数字なのですが、修士は年間で5,000人ということで、それなりに専門性を持った人間が出てきて、そういう人たちが企業に入り、実はこのような人が増えている。簡単に言いますと待遇がいいということでこの分野で人が不足してしまっていて、いい職に就けるということでそういう現状もあり、増えていて、日本もやはりこの分野の人材を増やしていかないと、その内に「日本はどうなってしまうのだ」というような状況だと思います。

アメリカでの統計学の学位数



• 日本は基本的にゼロ。統計学教員の不足が深刻。

7

実際に文部科学省もそう思っています、とく今は「AI 戦略 2019」というものが発表されていて、これにしたがって文部科学省の政策が進められているのですが、認定制度が進んでいて、大学生に勉強をさせましょうというような計画が進められています。データサイエンスで人材が必要だということで、それに沿った政策が重要だと思います（スライド8）。

AI 戦略2019 (2019年 7月)

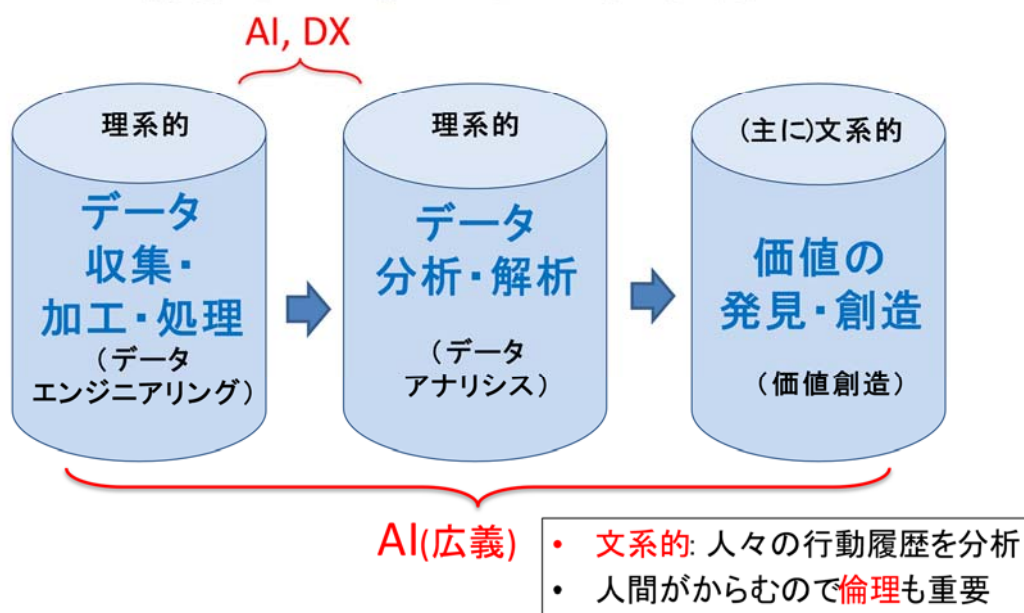


われわれの方も滋賀大学でデータサイエンス学部を創っていますが、そこでどういうことを教えているかですが、データサイエンスの3要素というように説明しています。データの「収集・加工・処理」をする。そしてそのデータを「分析・解析」をする。さらにそれにもとづいてデータから「価値の発見・創造」をするということで、実際に社会に役立つようなことをやっというこことです。こういう3つで、収集して分析し処理をする。こういうステップがあり、そのように流れて進んでいくということを説明しています。これはデータサイエンスの3要素と言うのですが、実はここはある意味では狭く、あとは技術的なことがあります（スライド9）。

データサイエンスの3要素

文理融合分野

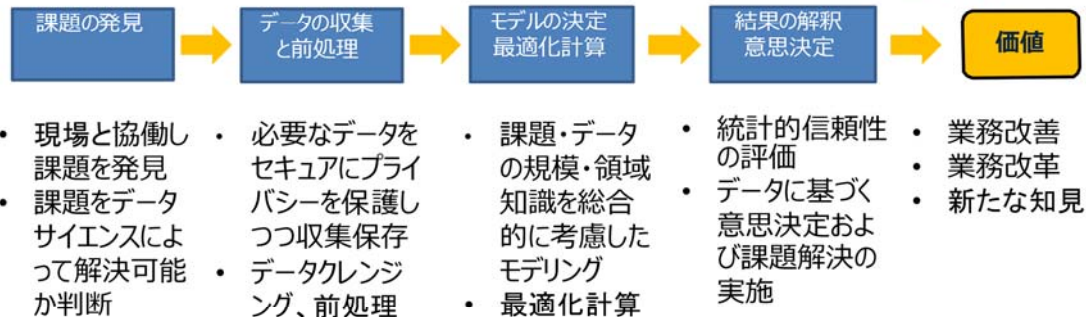
情報学 + 統計学 + 価値創造



その前後に、そもそもどういう課題を解決しようかという「課題の解決」があり、問題意識ですが、そうすることで、それをデータサイエンスで解決していこうということで、分析したあとは、どうやって社会に実現していけるか、実際のものにしていくかということで、分析結果を活かして、実際の改善につなげるということが言われています（スライド10）。

データからの価値創造の流れ

課題駆動型アプローチ

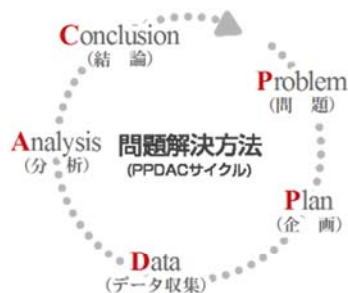


データ駆動型アプローチ: あらかじめ仮説や課題を設けず、データから知見を発見

10

スライド 11 は、課題を発見して分析し、解決したら、それで一連の流れはお終いなのですが、実は実際にやってみると次の課題が発見されますので、実は新たな課題の発見ということで、グルグル回っていく。PDCA サイクルということで、また戻るので、データサイエンスは PDCA サイクルと言われます。

PPDACサイクル (“PDCAサイクル”のデータサイエンス版)



• データサイエンスやデータサイエンティストはどこに関わるのか

- Problem: 課題、問題、指標の設定
- Plan: 仮説や分析手順の設定
- Data: データの収集
- Analysis: 分析
- Conclusion: 結論、ただし、さらに「実装」が必要

11

滋賀大学 DS 学部と大学院の教育

以上ビッグデータとデータサイエンスについてお話ししましたが、次に滋賀大学の方でどうしているかをお話しします。

滋賀大学は日本初のデータサイエンス学部ということで、2017年に、5年ほど前になりますが、学部を創りました（スライド12）。

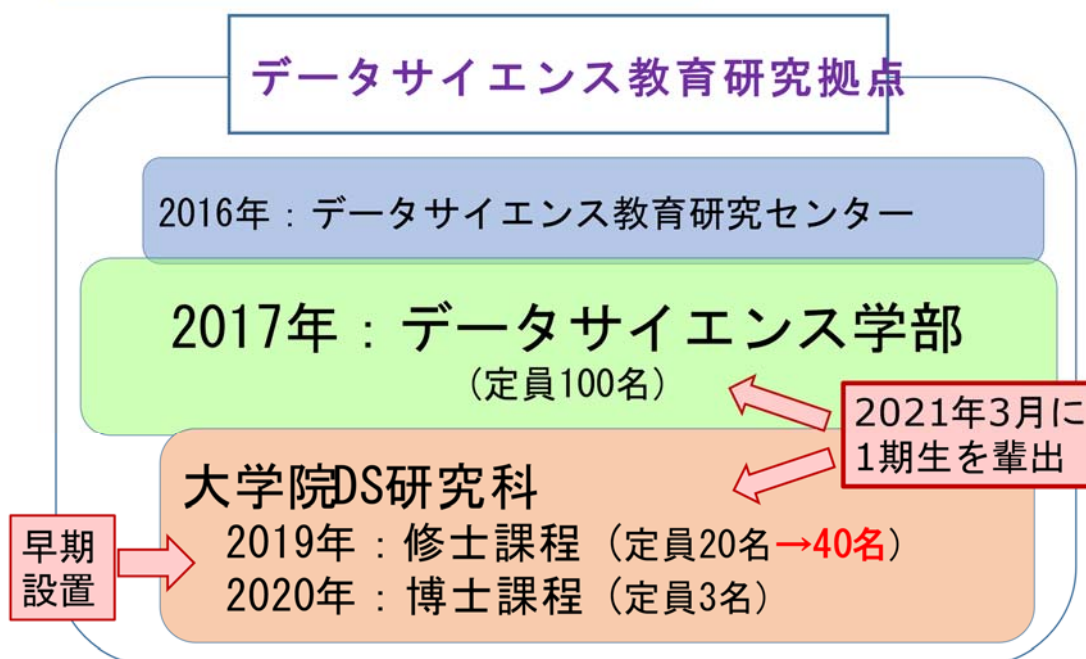
滋賀大学DS学部と大学院の教育

- データサイエンス拠点
- 学生や修士院生の現状

12

学生は、今は100名で、昨年3月に1期生を輩出しました。実はその内20名は大学院に進みました。大学院は学部ができてから2年後に創りましたが、早期設置で学生がまだいない段階で20人で創り、1期生が出る段階で20名を引き受けるために定員を40名にしました（スライド13）。

日本初のデータサイエンス教育研究拠点

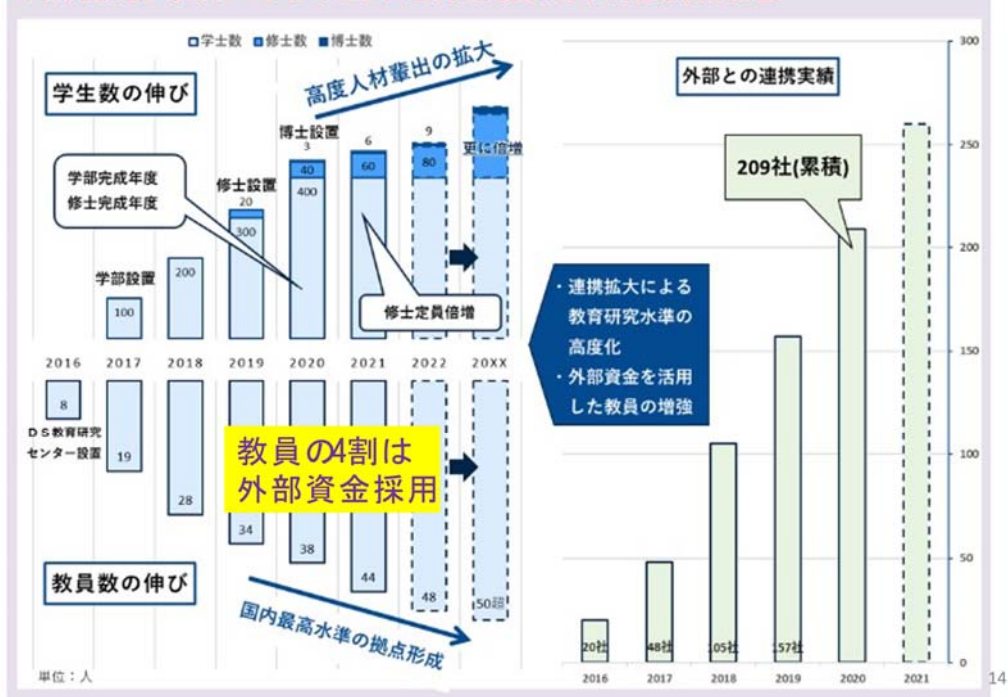


13

1 期生は、最初は医療保健の方が多く、そういう社会人を引き受けるために大学院を創りました。滋賀大学はデータサイエンス分野では人材を輩出してしまして、今ではデータサイエンス教育研究拠点の性格を強めています。医療連携も進めていて累積では 209 社になっています (スライド 14)。

拠点として急速に整備

◆国内最高水準のデータサイエンス教育研究拠点の状況と今後の展望



他の大学を見ますと、2017年のわれわれが日本初なのですが、その後、横浜市立大学、武蔵野大学等で毎年できて、来年の4月には4大学が新たにできる予定ですが、このスライド15には書いていないのですが、順天堂大学に健康データサイエンス学部があり、健康・医療分野ではそちらが最初です。一橋大学、名古屋市立大学などにもできていて、来年4月には4大学で新しくできて、全体で11大学になります。これで学部生が1大学で100人ほどですから、全部で800人ぐらいでしょうか。大学院はまだできていないところもあり、学士も修士もアメリカ統計学のようにはまだまだなっていませんが、いちおうこのくらいはあるということです。

相次ぐデータサイエンス系学部の創設

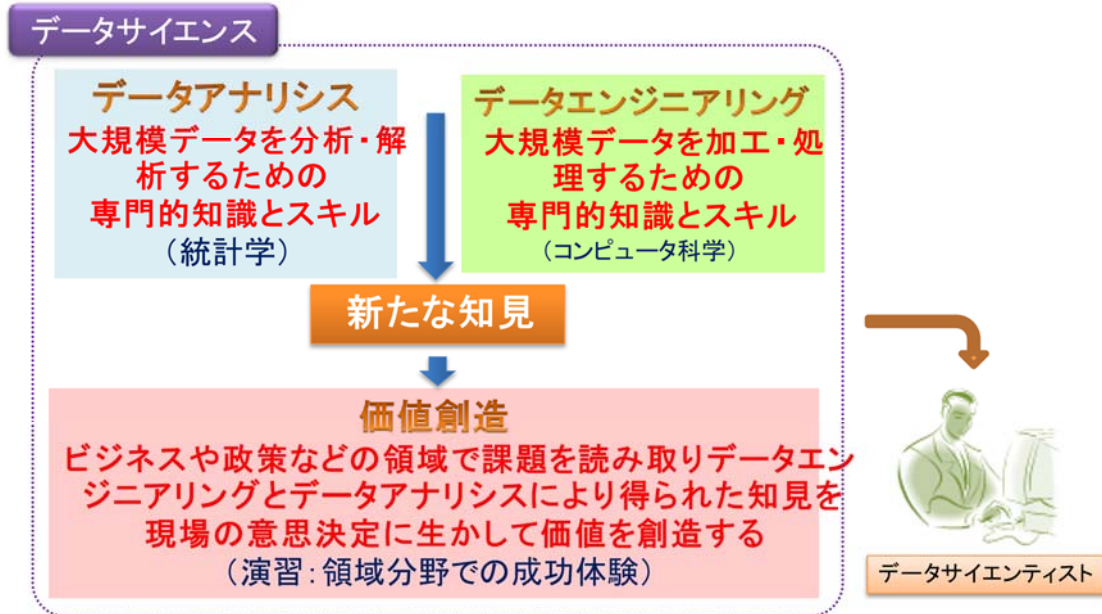
- 2017年 滋賀大学データサイエンス学部
- 2018年 横浜市立大学データサイエンス学部
- 2019年 武蔵野大学データサイエンス学部
兵庫県立大学社会情報科学部
- 2020年 長崎大学情報データ科学部
- 2021年 立正大学データサイエンス学部
群馬大学情報学部
- 2023年 一橋大学ソーシャル・データサイエンス学部
名古屋市立大学データサイエンス学部
京都女子大学データサイエンス学部
大阪成蹊大学データサイエンス学部

15

滋賀大学では先ほどの3要素を重視して、統計学、コンピュータ科学などを教育します。実際のデータ分析では現実の課題を採り上げます（スライド16）。

実際の課題を採り上げてやることは大学の力だけでは難しく、実は多数の企業・自治体との連携をしています（スライド17）。

人材育成の滋賀大モデル



16

スライド 18 は、1 期生の卒業研究です。提供元企業と受け入れゼミが出ています。

学部1期生の卒業研究

発表者 (指導教員)	発表タイトル
森本 滯二 (佐藤 智和)	実世界の仮想化に基づく高臨場VR型防災教育システムの開発
水口 綾乃 (市川 治)	事前学習済み分散表現を利用した学部オープンキャンパス向け質問応答システムの構築
江口 公基 (加藤 博和)	彦根市を目的地とした観光交通における鉄道利用促進のための機関選択分析
高田 拓弥 (松井 秀俊)	関数データに基づく回帰モデルと農業・化学分野への応用
森口 翼 (河本 薫)	テナント型商業施設における会員用スマホアプリのログデータ分析による離反防止策の検討
田室 建志 (河本 薫)	自動車部品工場における機械学習を活用した異常検知モデルの構築
上田 知展 (清水 昌平)	平和堂のID付きPOSデータを活用したモバイルクーポンの改善
小西 秀明 (清水 昌平)	ID付きPOSデータを利用したモバイルクーポンの仕様改善に関する施策提案と効果検証
谷口 友哉 (清水 昌平)	購買履歴データを用いたモバイルクーポン配信の最適化
仲北 昌大 (和泉 志津恵)	2018年7月の西日本豪雨災害のアンケートデータから分かる発見と問題点

(提供元企業 (一部))

- | | | |
|---------------------------|---------------------------|-----------------------|
| ●アイシンAW (河本ゼミ) | ●滋賀大学施設管理課 (松井ゼミ) | ●平和堂 (清水ゼミ) |
| ●インタージ (河本ゼミ) | ●東京大学CoREF (市川ゼミ) | ●堀場製作所 (松井ゼミ) |
| ●株式会社ワイヤ・アンド・ワイヤレス (横田ゼミ) | ●トヨタファイナンス (松井ゼミ) | ●マクロミル (清水ゼミ) |
| ●京都大学防災研究所 (佐藤ゼミ、和泉ゼミ) | ●ドコモ・インタージ・マーケティング (横田ゼミ) | ●山口県農林総合技術センター (松井ゼミ) |
| ●滋賀県無料Wi-Fi整備促進協議会 (横田ゼミ) | ●日本経済新聞社 (笛田ゼミ) | ●楽天データセット (市川ゼミ) |
| | ●野村総合研究所 (姫野ゼミ) | ●統計数理研究所 (和泉ゼミ) |
| | ●パルコ (河本ゼミ) | ●東京大学社会科学研究所 (和泉ゼミ) |

18

先ほど大学院を3年ほど前に創ったと言いましたが、その1期生の写真がスライド19です。その後も年間15人程度の企業派遣院生を受け入れています。学位は2年間なのですが、1年目には、この分野の手法を集中的に勉強して毎日課題に取り組みますが、2年目には各企業の課題と、それを解決するためにどうするかを企業に戻り、実際のデータで分析し、修士論文を書きます (スライド19)。

修士課程の状況

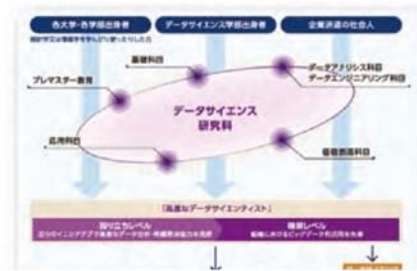
大学院データサイエンス研究科（修士課程）

社会ですぐ活躍できる一気通貫型のデータサイエンティストの育成！



期待の新星！ DS 修士課程第1期生たち

の共同研究に参加し、高度な分析結果を意思決定につなげていく基礎的力量を実践的に養います。



- 企業派遣院生を年間15名程度受け入れ
- 1年目はインテンシブな講義・演習
- 2年目は修士論文執筆（企業に戻ってデータ分析）

19

共同研究の分野としては、スライド 20 にあるような業界で、マーケティングと言いますか、そういうところで研究する一方で、製造業の技術管理や建設などの道を選ぶところもあります。健康・医療の方のスタッフも滋賀大学の方にもいまして、研究科を創り、色々な企業から相談があり、大学院生の派遣も様々なところから来ていただいています。健康・医療に特化しているわけではありませので、スタッフもおられますので、時間があれば見学していただければと思います。

業界別課題に関する共同研究分野

