

データサイエンスを育成する大学教育の 取り組みについて

竹村 彰通

滋賀大学データサイエンス学部長

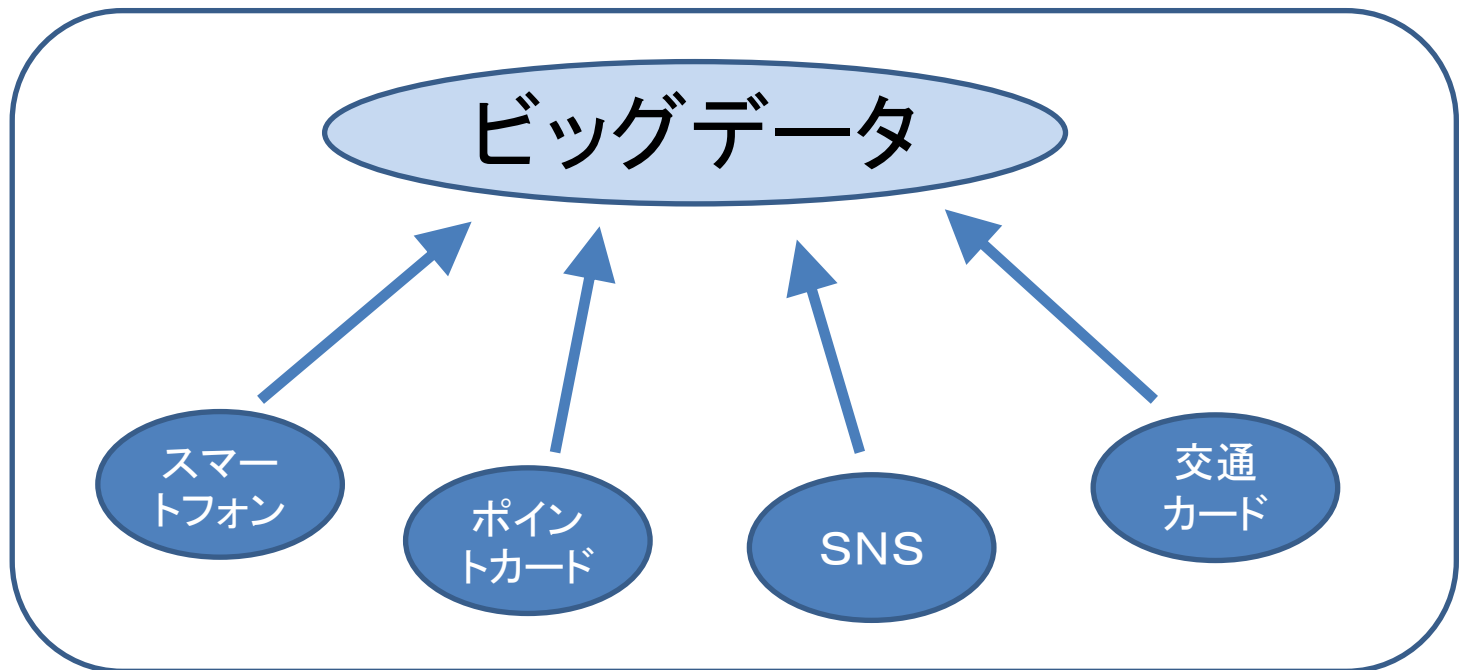
2018年3月14日

第9回産学連携人材ニーズ交流会

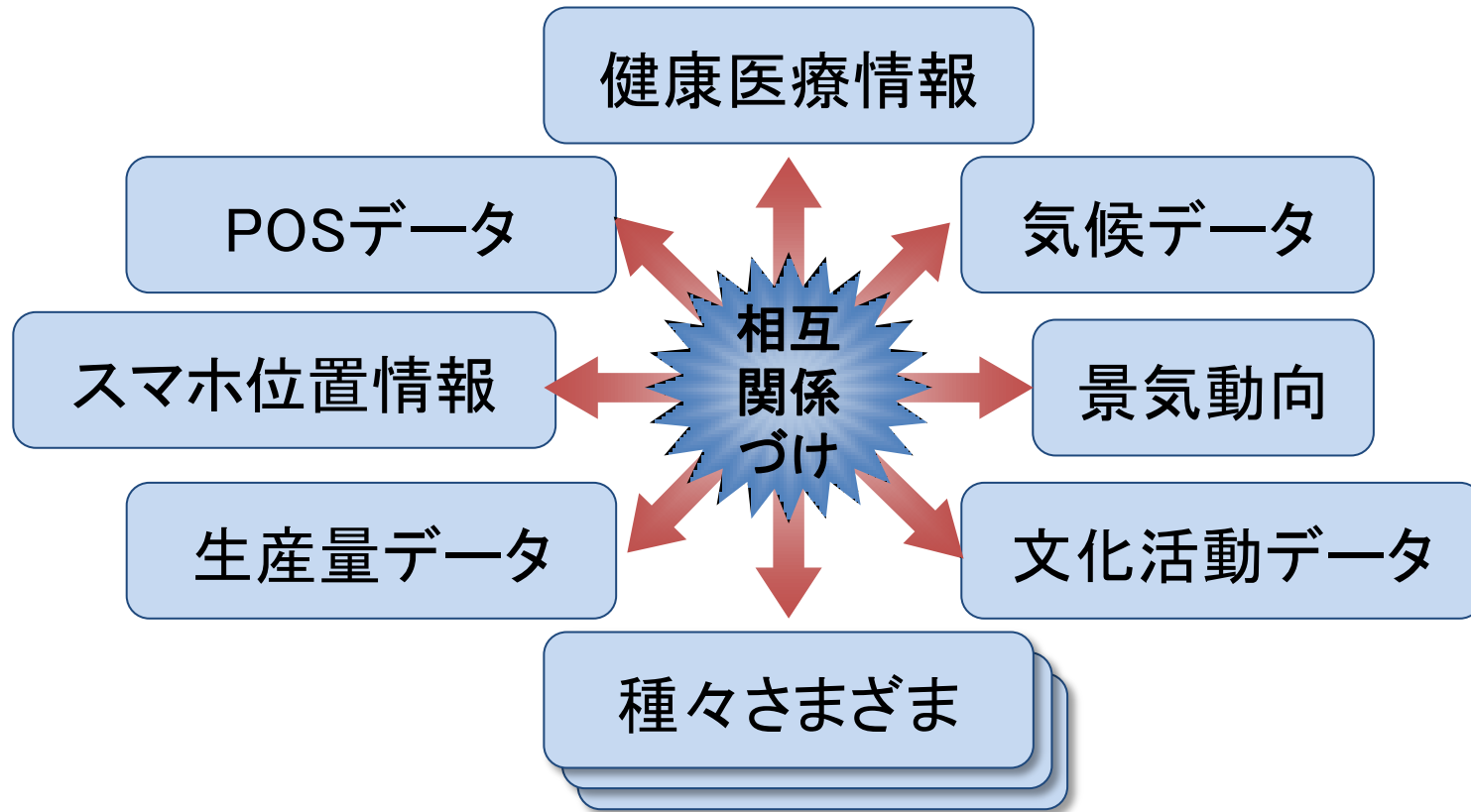
- データサイエンス(DS)とビッグデータ
 - スマートフォン普及の例
 - 新たな資源としてのビッグデータ
 - データの流通
- 海外の大学の動向
- 滋賀大学データサイエンス学部開設
- カリキュラムの概要と育成する人材像
- 企業や地方自治体との連携
- データサイエンスの手法、AI

データサイエンスとビッグデータ

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



ビッグデータから価値を生み出す



特に、人の行動履歴が直接とれるようになったことが重要
例：スマートフォンによる検索履歴

ビッグデータ時代の象徴：スマートフォンの普及

iPhone: 2007年アメリカで発売

- iPod、携帯電話、インターネットを統合
- マルチタッチ
- 2メガピクセルカメラ、ヘッドセットジャック、スピーカー、マイク
- 加速度センサ、環境光センサ



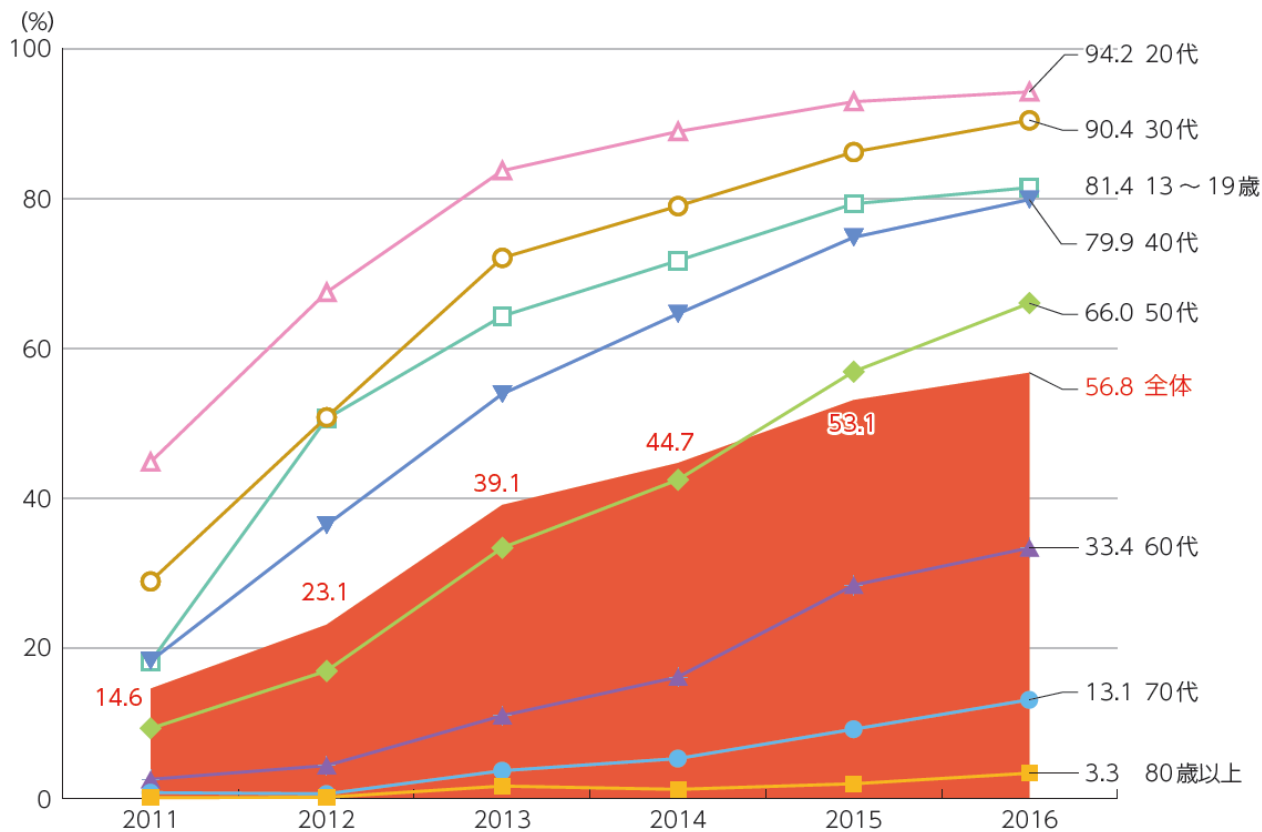
<http://japanese.engadget.com/2007/01/09/macworld-2007-jobs-keynote/>

現在のスマホは30年前のスーパーコンピュータ

- 1985年の Cray 2 のメモリ(RAM)2GB
- 2016年発売の iPhone 7 も2GBメモリ

日本でのスマートフォンの普及率

図表 1-1-1-2 スマートフォン個人保有率の推移



20代では
ほぼ100%

(出典) 総務省 通信利用動向調査

更に5GとLPWAの時代へ

5G

- 現在の4Gと比較して100倍高速
- 「多数接続」や「超低遅延」といった新たな特徴も持つ
- 2025年頃に普及

LPWA (Low Power Wide Area)

- 低速なもの、一般的な電池で数年から数十年持つ
省電力性
- 数kmから数十kmもの通信が可能な広域性
- 2020年頃に普及

平成29年版情報通信白書

ビッグデータ＝新たな資源

“21世紀の石油”

- 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
- 日本ではデータを分析する人材が圧倒的に不足

データを流通させる仕組み作りも重要

21世紀の石油

The Economist 誌 2017年5月6日

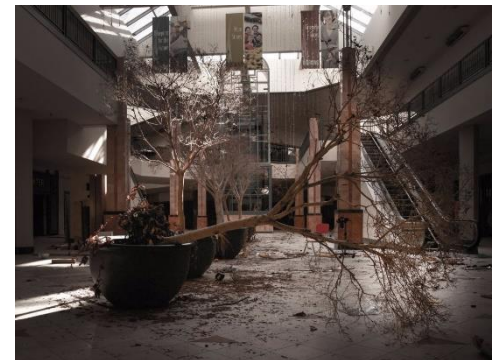
The world's most valuable resource is no longer oil, but data



Regulating the internet giants

アメリカでもショッピングモールの3割が消滅しつつある

by Seph Lawless
<http://sephlawless.com>



政府もデータの流通を後押し

改正個人情報保護法 2015年9月成立

- 匿名加工情報の利用・提供の促進 36条～39条
- 個人情報を適切な方法により匿名加工した情報は、本人の同意を得ることなく第三者に提供することができる
- ただし、匿名加工の実施、必要事項の公表、安全管理措置の実施等が要求される

官民データ活用推進基本法 2016年12月成立

第一条 この法律は、インターネットその他の高度情報通信ネットワークを通じて流通する多様かつ大量の情報を適正かつ効果的に活用することにより、…、官民データの適正かつ効果的な活用の推進に関し、基本理念を定め、国、地方公共団体及び事業者の責務を明らかにし、…、官民データ活用推進戦略会議を設置することにより、官民データ活用の推進に関する施策を総合的かつ効果的に推進し、…快適な生活環境の実現に寄与することを目的とする。

次世代医療基盤法 2017年4月成立

- 国民の医療情報を匿名加工して大学や製薬企業の研究開発などでの活用を可能にする
- オプトアウトにより患者本人に拒否されない限り第三者(認定事業者)に提供できるようにした

データが流通しても、分析する人材が不足

政府もデータサイエンスの重要性を強調

- 「ビッグデータ時代を迎え、データの利活用により付加価値を生み出す新事業・新サービスの創出が重要、第4次産業革命を支える基盤技術：AI、ビッグデータ、IoTなど」（日本再興戦略2016等）
- 「欧米等と比較し、データ分析のスキルを有する人材や統計科学を専攻する人材が極めて少ないという危機的状況」（科学技術イノベーション総合戦略2015）

現実にはこの分野で日本は著しく立ち遅れ

← 一つの要因として、統計学部・学科がなかった

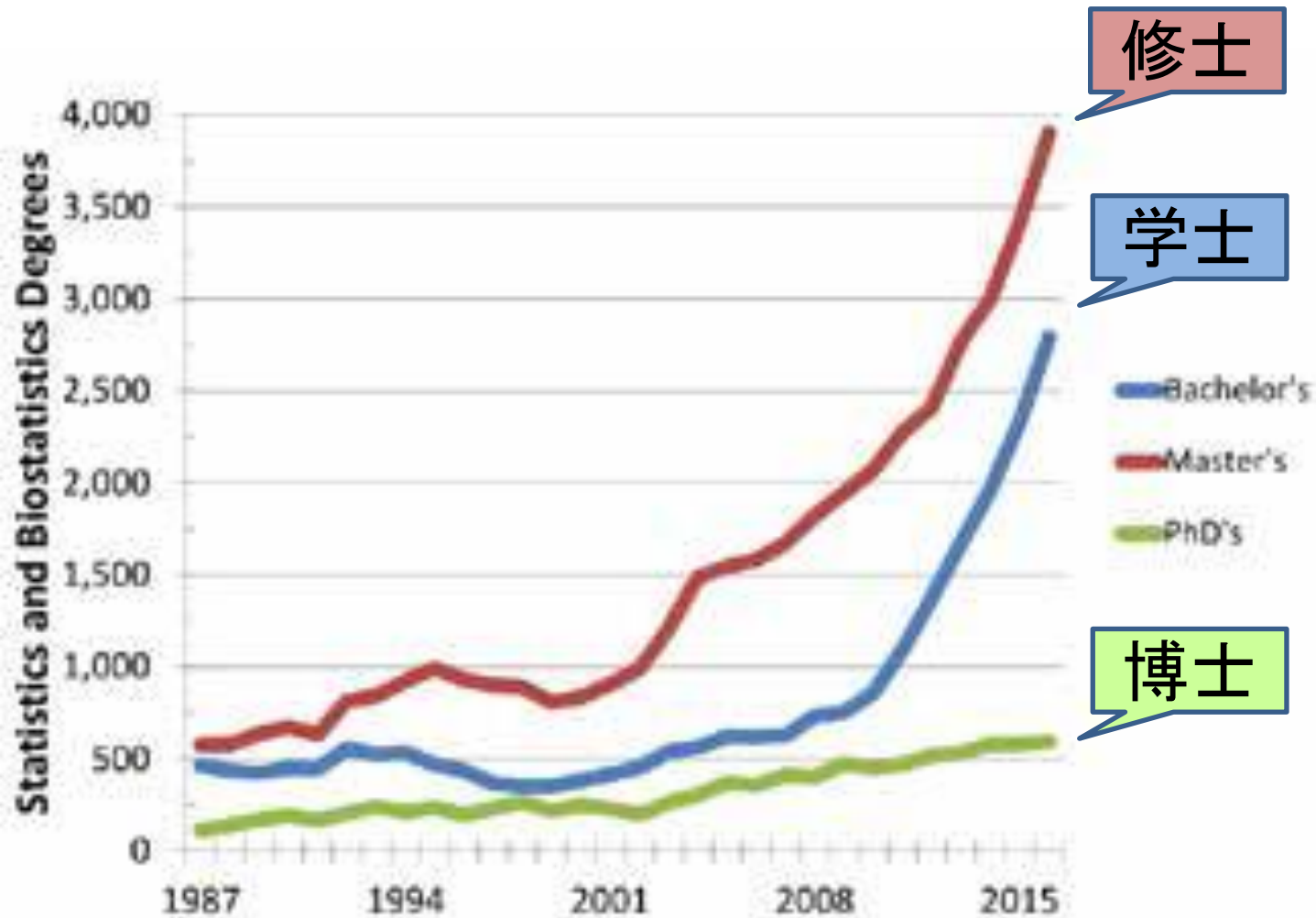
諸外国での統計学部・学科の数

- アメリカでは100以上。さらに大学院レベルでは生物統計専攻がたくさんある
- イギリスでは50程度
- 韓国にも50程度
- 中国では300以上あり、さらに増えている

これに対して日本はこれまでゼロ

アメリカでの統計学の学位数

アメリカ統計学会ニュースレター 2017年10月号



滋賀大学：昨年4月に日本初の「データサイエンス学部」を設置

- 募集定員：1学年100名
- 2年目の出願状況：
前期日程 3.1倍、後期日程 7.0倍
- 1期生110名の分布
男子：女子 = 8：3、 理系：文系 = 6：4
- 実践的教育を目指して、多くの企業と教育・研究に関する連携推進

日本初のデータサイエンス学部

➡ DS教育研究拠点

データサイエンス教育研究拠点

データサイエンス教育研究センター (H28)

データサイエンス学部 (H29)

大学院DS研究科 (早期に実現目指す)

社会人スキルアップ
の需要に対応

DS基盤研究

機械学習、最適化等の先端研究

DS教育開発

データ駆動型PBL演習教材の開発

DS価値創造プロジェクト研究

企業や自治体のデータ利活用

DS調査・情報発信

海外動向、育成人材像調査

「数理・DS教育に係る教育強化」拠点大学に選定

(2016年12月)

6拠点選定

No	大学名	事業名
1	北海道大学	数理的データ活用能力育成特別教育プログラム ～データサイエンスセンター(仮称)の設置～
2	東京大学	数理・情報教育研究センターの設立
3	滋賀大学	データサイエンス教育の全学・全国への展開 ーデータリテラシーを備えた人材の育成に向けた カリキュラム・教材の開発ー
4	京都大学	データ科学イノベーション教育研究センター構想 ー21世紀イノベーションを支える人材育成ー
5	大阪大学	数理・データ科学の教育拠点形成
6	九州大学	九州大学「数理・データサイエンス教育研究センター(仮称)」構想

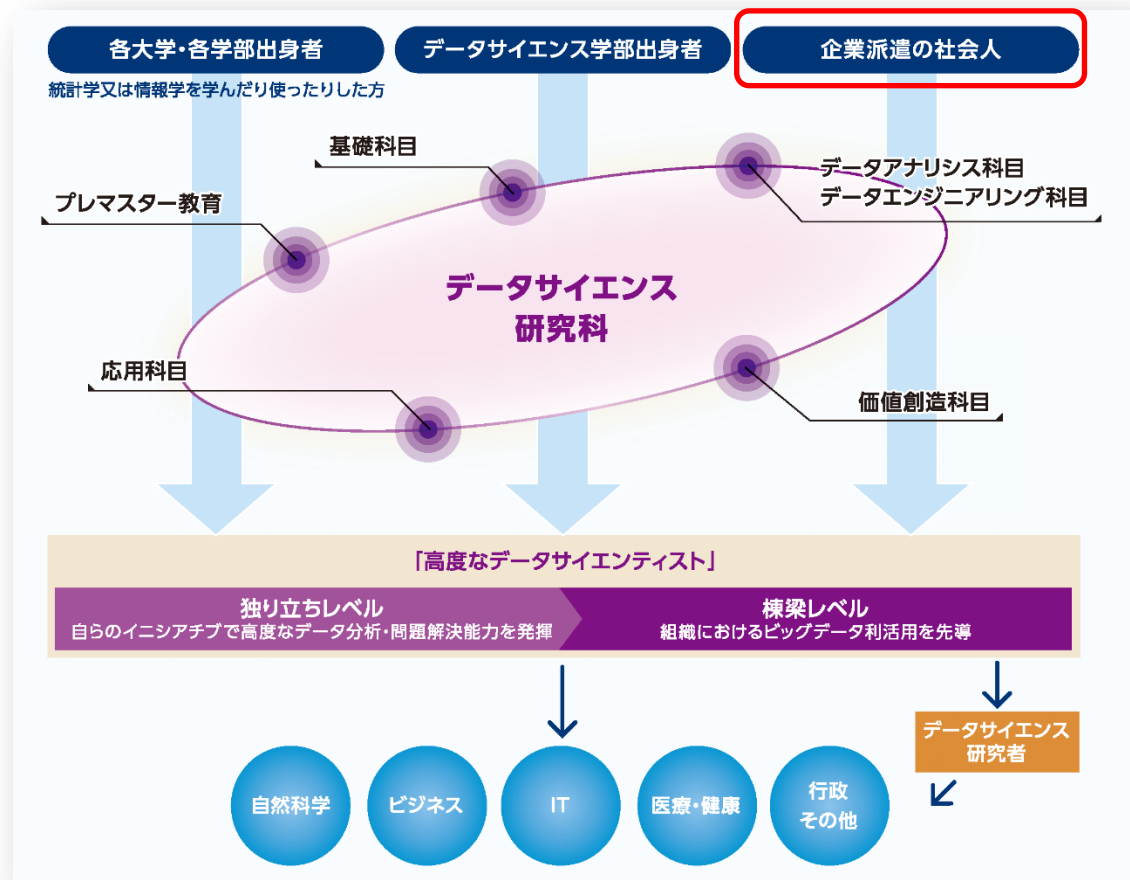
滋賀大学DS大学院研究科構想（2019年4月設置構想中）

DIVE TO THE FUTURE

数理・統計/コンピュータ×現実社会

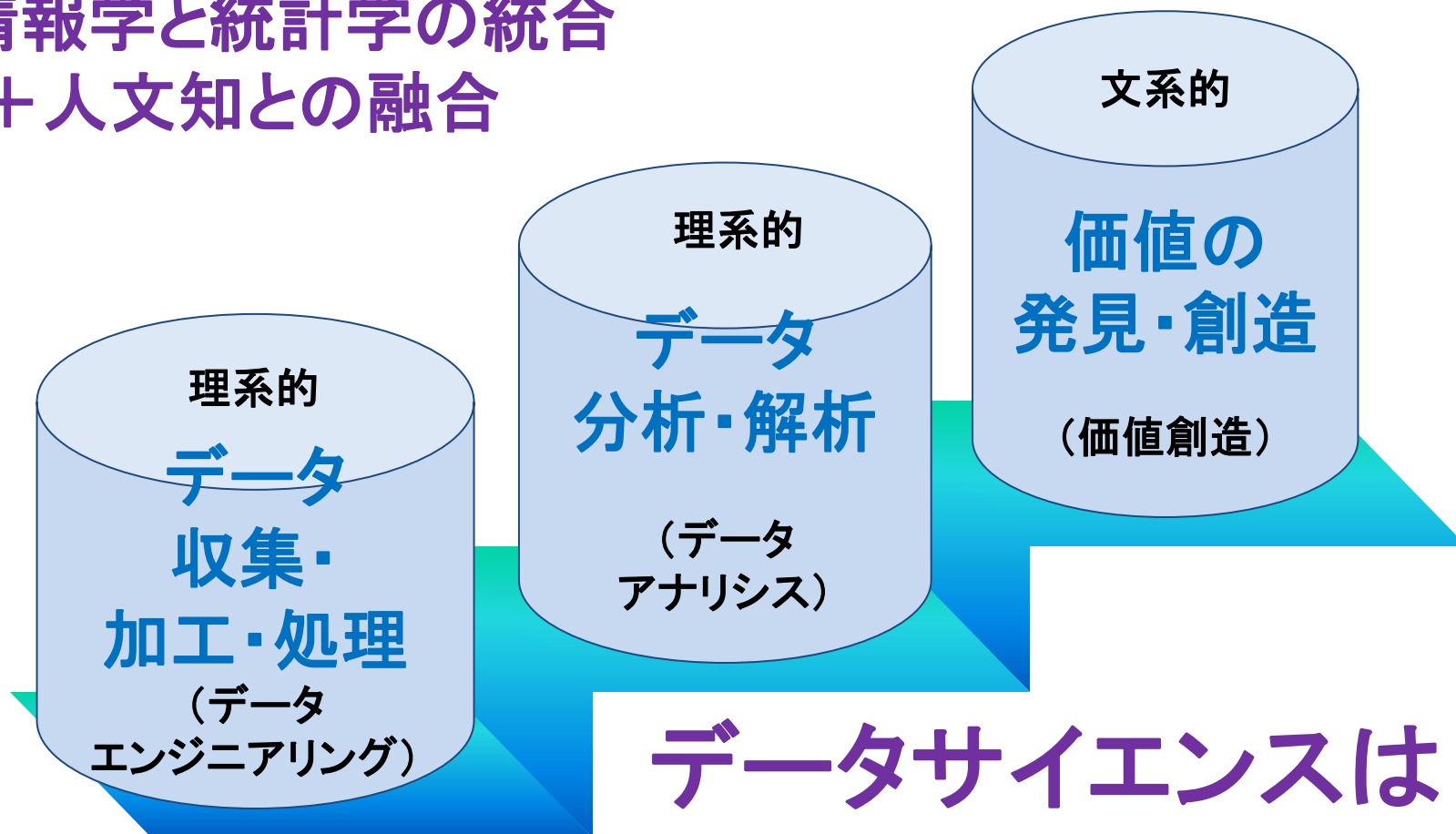
- データ科学を探究し価値創造を担う「高度なデータサイエンティスト」を養成

早期開設により、企業DS人材の高度化に貢献



データの処理・分析から価値創造までー滋賀大モデル

情報学と統計学の統合 ＋人文知との融合



データサイエンスは 文理融合

データサイエンス学部における育成人材像

文理融合 逆Ⅱ型 人材

《多様な価値創造のフィールド》

- マーケティング
- ファイナンス、保険
- 企業会計
- ビジネスエコノミクス
- 医療・健康・福祉
- バイオ、製薬
- 環境、防災、気象
- 教育
- 公的統計
- 社会心理
- 地域文化情報

価値創造の
経験とノウハウ
&
領域における
専門知識

文系的

理系的

データサイエンスの専門知識とスキル

データアナリシス
大規模データの分析・解析
専門知識とスキル
(統計学)

データエンジニアリング
大規模データを加工・研磨・
処理専門知識とスキル
(情報学・コンピュータ科学)

文理融合は、進
路指導の高校の
先生には好まれ
ない!

Ⅱ

領域を複数経験

領域①
価値創造の経験
とノウハウ

領域②
価値創造の経験
とノウハウ

データサイエンスの
専門知識とスキル

- 統計や情報は多くの分野で共通に用いられる横断的な技術（横串）
- 固有の専門領域は縦割り
- 理系と文系の区別など、日本の教育は縦割りの考え方が強い
- 統計などの汎用的な手法は「後から必要に応じて勉強すればよい」という考え方が強い
- しかし、最近では、横串の技術のほうがイノベーションに貢献
- **逆U型**：横串としての統計及びコンピュータを先に学んで、それを複数の領域に応用する



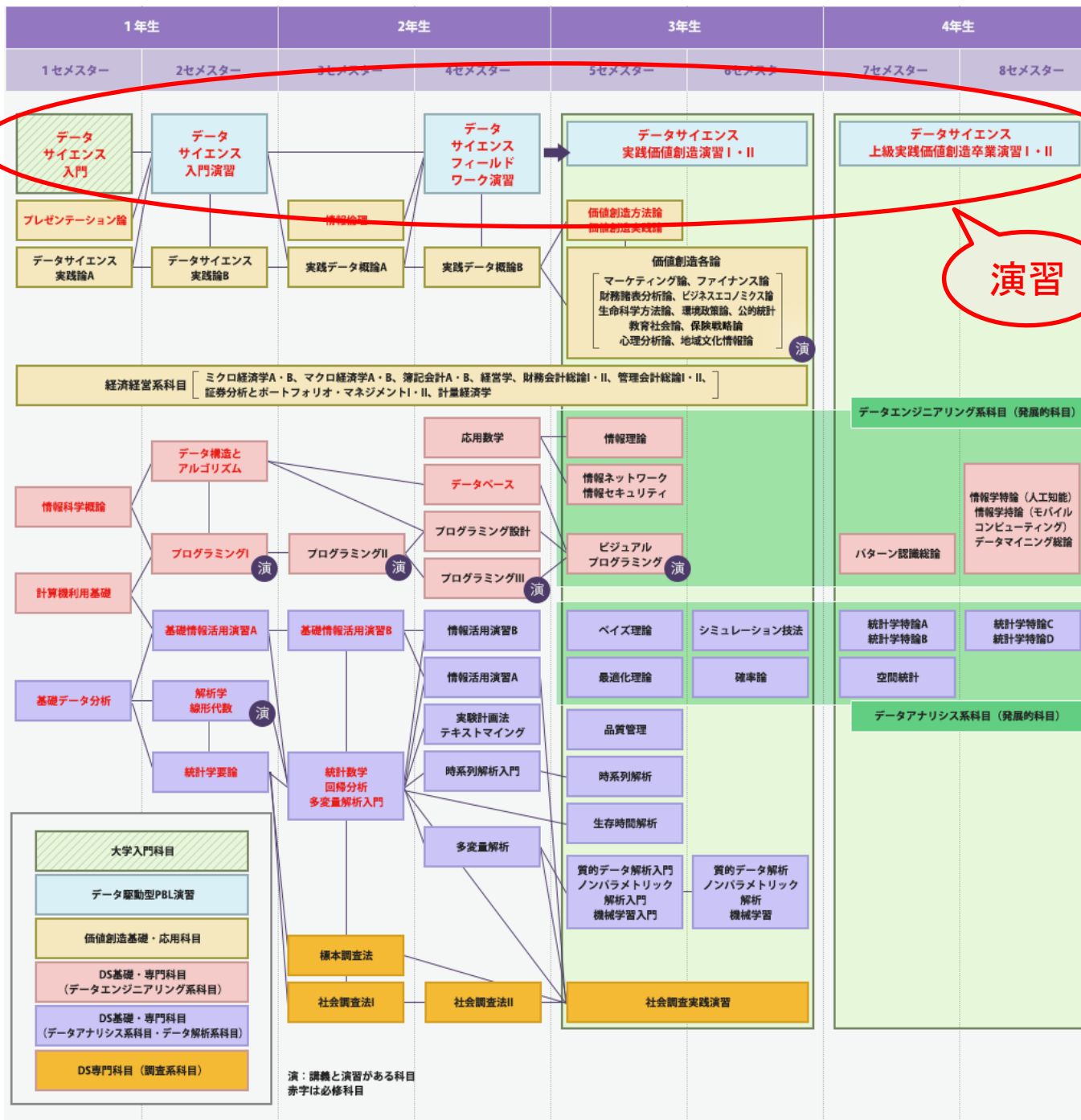
領域を複数経験

領域①
価値創造の経験
とノウハウ

領域②
価値創造の経験
とノウハウ

データサイエンスの
専門知識とスキル

カリキュラムツリー

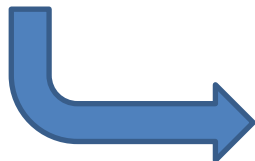


- **実際のデータ**を用いた演習重視
- そのために地公共団体や企業等との連携を進めている

演習

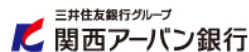
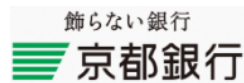
実際のデータを用いた価値創造教育

- 価値創造は講義では教えられない
- 現場のデータを利用した価値創造PBL演習での成功体験が重要
- DS教育研究センターにおける価値創造プロジェクトを通じて、企業・自治体等の現場とデータを提供
- 外部に開かれた実践の場でのコミュニケーション力やチームワーク形成力の鍛錬



多数の企業・自治体等との連携

多数の企業・自治体等との連携



(企業等)

あいおいニッセイ同和損害保険(株)
(株)アイディーズ
伊藤忠テクノソリューションズ(株)
(株)インフィック
(株)SMBC信託銀行
(株)オプトホールディング
(株)関西アーバン銀行
(株)京都銀行
(株)滋賀銀行
滋賀経済同友会
滋賀県商工会連合会
滋賀中央信用金庫
(株)ショーケース・ティービー
第一生命ホールディング(株)
玉田工業(株)
(株)帝国データバンク
(株)デンソー
東レエンジニアリング(株)
(株)日本能率協会マネジメントセンター
トヨタ自動車(株)

PwCあらた有限責任監査法人

NPO法人ビュー・コミュニケーションズ

(株)堀場製作所

(株)堀場エステック

(株)堀場アドバンスドテクノ

(株)マクロミル

(株)三井住友フィナンシャルグループ

村田機械(株)

(政府研究機関等)

理化学研究所 革新知能統合研究センター

情報システム研究機構 統計数理研究所

総務省 統計研究研修所

(独)統計センター

(大学)

統計教育大学間連携ネットワーク(JINSE)拠点校

(東京、大阪、滋賀、総研大、青山学院、多摩、立教、早稲田、同志社)

数理及びデータサイエンスに係る教育強化拠点校

(北海道、東京、滋賀、京都、大阪、九州)

データ関連人材育成プログラム参画・連携校

(大阪、神戸、滋賀、和歌山、奈良先端、京都) など

(自治体等)

滋賀県、滋賀県警察、県内自治体

滋賀県教育委員会、大津市教育委員会、

彦根市教育委員会、彦根東高校、虎姫高校 など



2017年4月28日報道発表

- トヨタグループのエンジニアをビッグデータ分析の指導者(中核人材)候補として育成するための研修プログラムを開始
- トヨタグループ各社から選抜された46名のエンジニアを対象に5月から来年3月までの間、研修プログラム「機械学習実践道場」を開催。計11回
- 午前は講義、午後は参加者の課題学習
- 滋賀大学スタッフが講師。参加者の具体的課題の解決を指導。すでにいくつかの改善事例が得られている。

あいおいニッセイ同和損保との連携

- あいおいニッセイ同和損保／滋賀大学日本セーフティソサイエティ研究センタ（JSSRC）を設置して、本格的な共同研究をおこなう

研究組織

JSSRC センター長
(滋賀大学)

JSSRC 副センター長
(あいおいニッセイ同和損保)

研究プロジェクト(開設時)

- ・主任研究員A(プロジェクトリーダー・滋賀大学)
- ・研究員(あいおいニッセイ同和損保)
- ・研究員(滋賀大学助教)

研究プロジェクト(今後)

- ・主任研究員B(プロジェクトリーダー・滋賀大学)
- ・研究員(あいおいニッセイ同和損保)
- ・研究員(滋賀大学助教)

2017年春学期 土曜日丸一日を3回

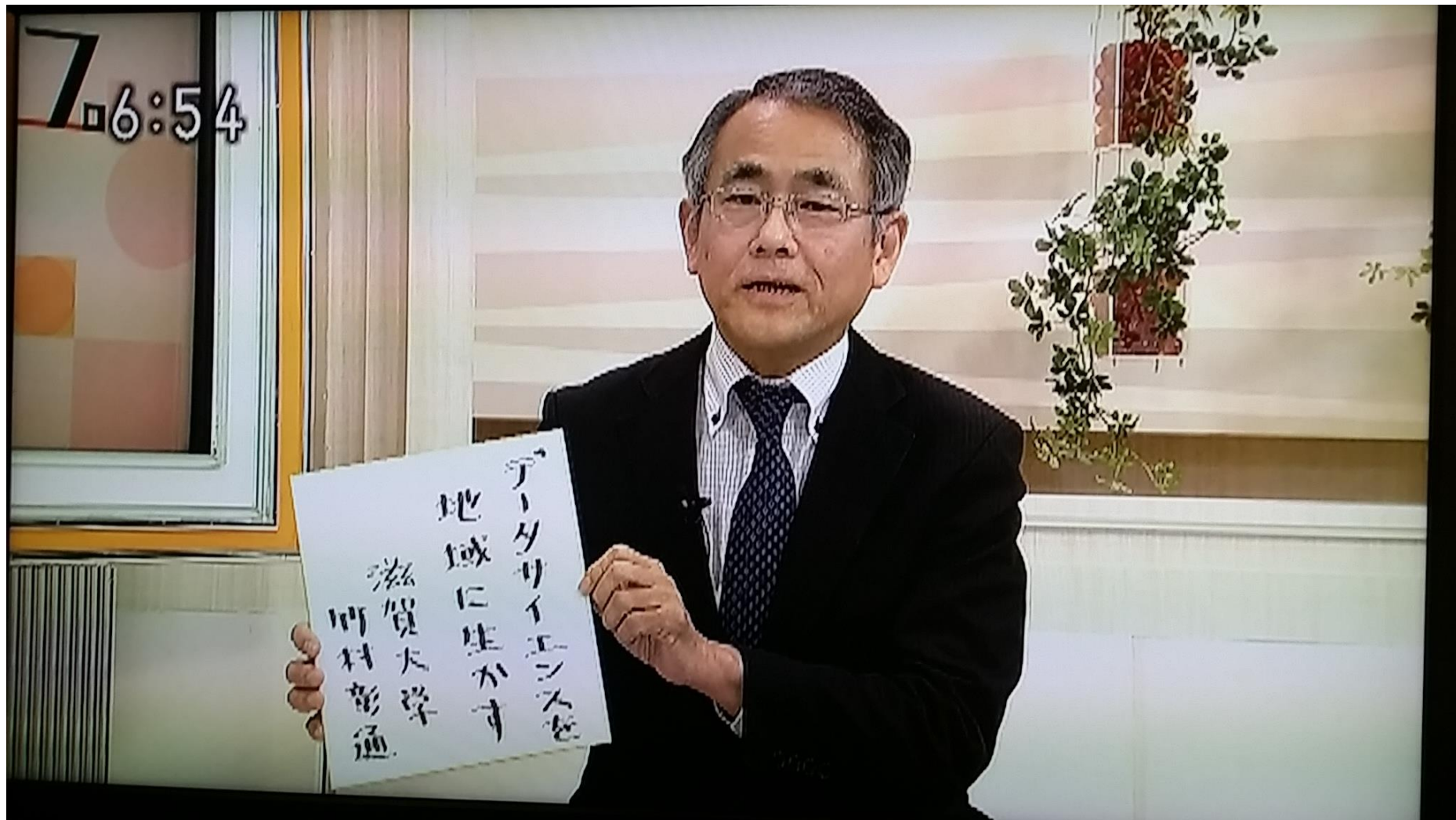
- IBMのクラウドコンピューティング環境である Bluemix のハンズオン
- ビッグデータ時代のセキュリティリスクとセキュリティ対策技術
- クラウド型コグニティブ分析 Watson Analytics によるデータ・ビジュアライゼーションのハンズオン
- お客様の音声をビジネスに生かす音声認識－音声ビッグデータの活用の広がり

など内容は多岐にわたるもの

データサイエンス実践論B (2017年秋学期)

	日程	回数	講義テーマ	会社名	講師名
1	2017/10/21(土)	1,2	・データサイエンスで実現する(したい)こと ・マーケティング戦略とデータサイエンス	株式会社野村総合研究所	福島健吾
2	2017/11/11(土)	3,4	・顧客理解や人工知能とデータサイエンス ・データサイエンティストの実際とこれから		
3	2017/11/25(土)	5	金融業界におけるデータサイエンス	株式会社金融エンジニアリング・グループ	中林三平
	2017/11/25(土)	6	不動産業界におけるデータサイエンス活用	株式会社GA technologies	小林 賢一郎 橋本 武彦
	2017/11/25(土)	7	時系列分析による需要および売上げの予測 (or チャットの会話履歴データを用いた会話ネットワークの分析)	株式会社GRI	大友
4	2017/12/02(土)	8	製造業向けIoTデータ分析	株式会社KSKアナリティクス	高木宏明
	2017/12/02(土)	9	健康経営×データ分析 ～IoT機器を用いた介入研究の事例をもとに	株式会社分析屋	青柳早苗
	2017/12/02(土)	10	銀行データを用いた顧客の行動分析	SMBC信託銀行	町田大樹
5	2018/01/20(土)	11	CRMの現場で求められる分析・仕事の進め方	ブレインパッド	兵藤 誠
	2018/01/20(土)	12	分析現場で必要となる数学力	ブレインパッド	梅田 義章
6	2018/02/03(土)	13	アドテク分野におけるデータサイエンスとエンジニアリング	株式会社 サイバーエージェント	杉尾 樹
	2018/02/03(土)	14	流通業やコンタクトセンタにおけるデータ分析技術活用事例	日本電気株式会社	三木 清一
	2018/02/03(土)	15	大企業の中でのデータサイエンティストの役割とヒエラルキーを突破する具体的方法の例示	株式会社 電通	佐伯 諭

地域への貢献 (DSに基づく地方創成)



NHK大津放送局「おうみ発630」2017年2月9日放送

滋賀大学DSセンターが提供できるノウハウ

最新の手法まで含むさまざまなデータ解析手法

- 将来データの予測のための時系列解析（ARIMAモデル、状態空間モデル）
（例）商品の今後の売り上げを予測
- 多くの要因の関係を分析する多変量解析（主成分分析、判別分析、SVM、グラフィカルモデル等）
（例）さまざまな商品の売り上げの間の関係などを分析
（例）性別や年齢などの顧客の属性による購買行動の違いを分析
- 変数間の因果関係を分析する因果分析
（例）売上げ増加のために、どのような広告手段が有効なのかを分析
- テキスト、画像、音声なども含む複雑なデータから情報を引き出し処理する機械学習
（例）カメラの動画からのデータ抽出、業務記録などのテキストの処理
- 多数の要因から重要なものを自動的に抽出するスパースモデリング
（例）生産ラインで品質に影響を与え得る多数の要因から、重要な要因を選択
- 深層学習として注目されている多層ニューラルネットワーク
（例）熟練者の技能を機械で再現するような複雑なモデルの作成

データサイエンスの手法

機械学習: コンピュータによる自動化

- 練習: 問題と答えの組をコンピュータに与えて学習させる

手書き数字(問題)

正解(答)



0



0



1

...



9



9



コンピュータ(機械)

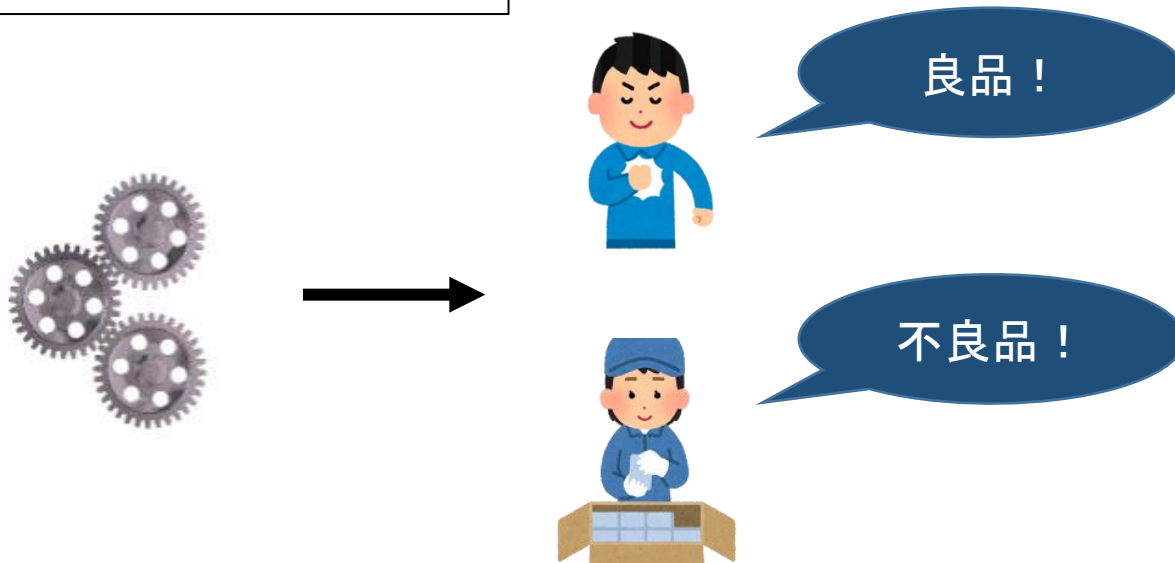


どんな手書きが
どの数字かを学習

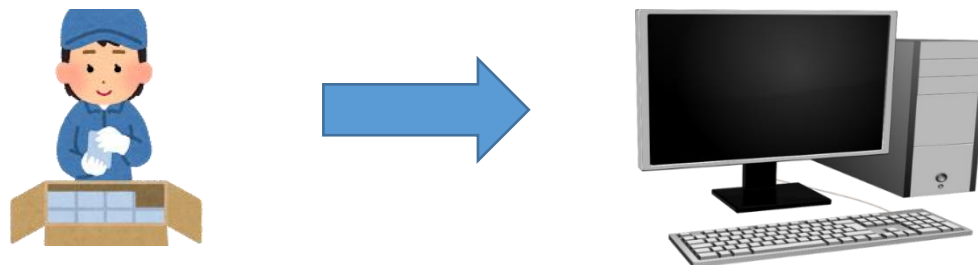
Alpaydin and Kaynak (1998)

良質の問題と答えの組(データ)を 大量に用意することが鍵

“良質のビッグデータ”



熟練検査員の力をコンピュータへ



複雑な関数が利用可能に 深層学習（ディープラーニング）

- 大量のデータ
- 高速なコンピュータ
- 本番に強くする技術（過学習回避）
- データがあればあるほど性能が良くなる
 - 巨大企業の隆盛
 - Google, Apple, Facebook, Amazonなど
- 良質なデータが重要
 - 練習問題と正答の組
 - 熟練の技をコンピュータに

回答と特徴量の関係

$$\text{回答} = f(\text{特徴量})$$



人工知能(AI): 我々の生活を格段に便利にする技術

- 記者「A Iが**人間の脅威**になると心配する人もいる。A I研究は今後、どこに向かっていくのでしょうか。」
- 杉山「A Iが感情を持って人を攻撃するのではないかと誤解されているのは、『人間っぽい』というイメージを抱かせる『人工知能』という言葉の悪い影響だろうが、**現状の技術はとてもそんなレベルではない**。しかし**我々の生活を格段に便利にする技術**なのは確かで、A Iの開発競争は今後いっそう熾烈になるだろう。」



杉山将: 東大教授・理研AIPセンター長
YOMIURI ONLINE 2017年2月13日

開発競争のためにも、
組織的な人材育成が鍵!

A I による画像認識が誤る例

Facebook によってアダルト画像と判断されて拒否された
クリスマスカード用の絵



<http://www.mirror.co.uk/news/uk-news/facebook-mocked-after-slapping-adult-11511010>

NHKスペシャル 「AIに聞いてみた どうすんのよ!?ニッポン」 より

(2017/7/22放送)



AIの分析結果から読み解いた提言



01.

健康になりたければ病院を減らせ



- 相関関係と因果関係を混同
- AIの予測性能は非常に向上しているが、
- ブラックボックス的で「説明」は不得意

まとめ

- ビッグデータの時代はさらに進んで行く
- データが新たな経済的資源
- 分析手法(“AI”)も進化
- 滋賀大学のデータサイエンス学部は先進的な取組みとして評価されている
- 多くの企業・団体と連携検討中
- 今後も社会連携を進めていく