第１回「統計エキスパート育成システムの構築」に向けたワークショップ
-大学統計教員育成センター設立を記念して-

本コンソーシアムが目指す 統計エキスパート育成システムの構築と 諸外国における取組例

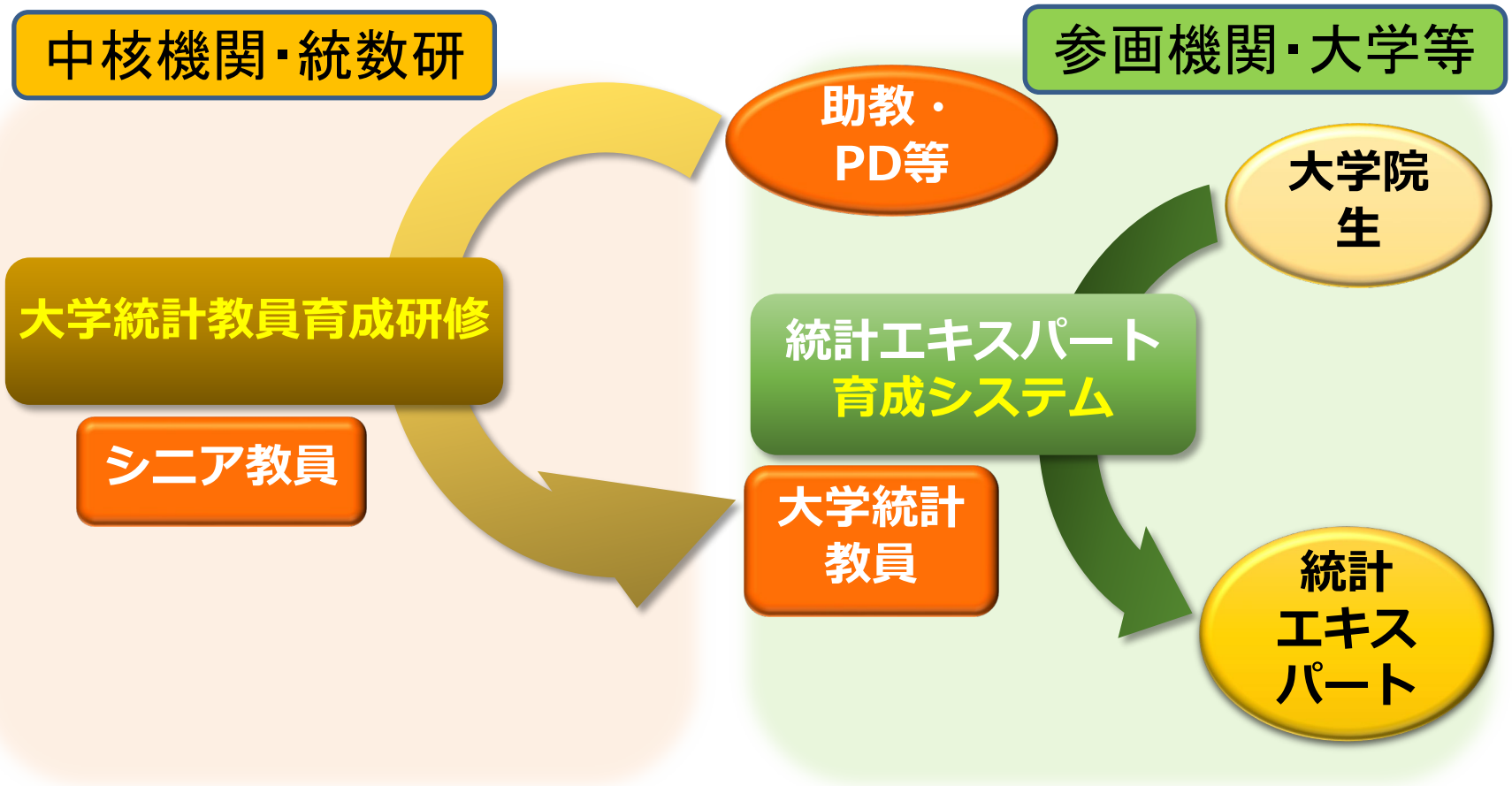
2022年 2月1日

情報・システム研究機構理事
統計数理研究所長　椿　広計

統計エキスパート人材育成プロジェクトとは

統計ｴｷｽﾊﾟｰﾄ人材育成ﾌﾟﾛｼﾞｪｸﾄとは？

プロジェクトの骨格



本プロジェクトの特徴は何か？

これまでの人材育成の取組（統数研）

○「リーディングDAT」～大学院レベルの統計教程により、
高度統計人材を育成（2017年度～）

○ **各分野の研究者育成** ～国立天文台の助教を統数研に出向させ、
天文学分野の中核的統計研究者に育成（2020年度～）

⇒ さまざまなレベルで統計エキスパートを育成

本プロジェクトでの人材育成の特徴

○ **本プロジェクト** ～統計エキスパートを育成する際に不可欠な
「大学統計教員」を育成（育成側の人材の育成）

⇒ 大学統計教員が全国の大学等で**統計エキスパートを育成**

（ ⇒ 育成された人材が指導者となり更に統計エキスパートを育成 ）

統計人材育成の好循環システムを構築！

日本の強みと弱み

George Box: 1986/06
Bell研若手研究者(V. Nair, Jeff Wuら)と共に日本(トヨタ)を視察

R.B. Reich,1984

- アメリカが10年間で50%以上シェアを落とした製造業リスト
  - Automobiles, Cameras, Stereo Components
  - Medical Equipment, Color Television Sets
  - Hand Tools, Radial Tires
  - Electric Motors, Food Processors
  - Microwave Ovens, Athletic Equipment
  - Computer Chips, Industrial Robots
  - Electron Microscopes, Machine Tools
- **高品質低コスト製品に敗退**

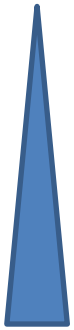
Box and Bisgaard, 1987

- The Scientific Context of Quality Improvement
- **"They Do It and We Don't"**
  - 小原、光藤、椿(2018)
    - 横幹連合コンファレンス
    - ボトムアップ小集団改善活動規模
    - 1988:製造業事業所無作為標本調査
      - 6.6万事業所
      - 75万QCサークル, **550万名**
- 当時:品質管理誌、QCC誌:
  - 問題解決ストーリ宝庫
  - 直交表実験・多変量解析の事例

椿:1989/03　MIT等に視察に行った時の認識と理想

「統計応用の標準化とネットワーク化30年間の活動を振り返る」

2011/3/6: 日本統計学会春季大会

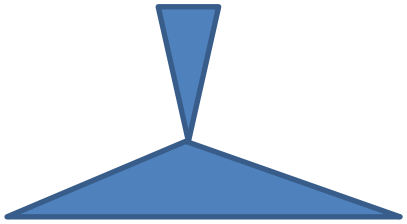


1980年後期の
アメリカ型

SCANS
Report;1992
自律的問題
解決能力は
学校教育で
育成

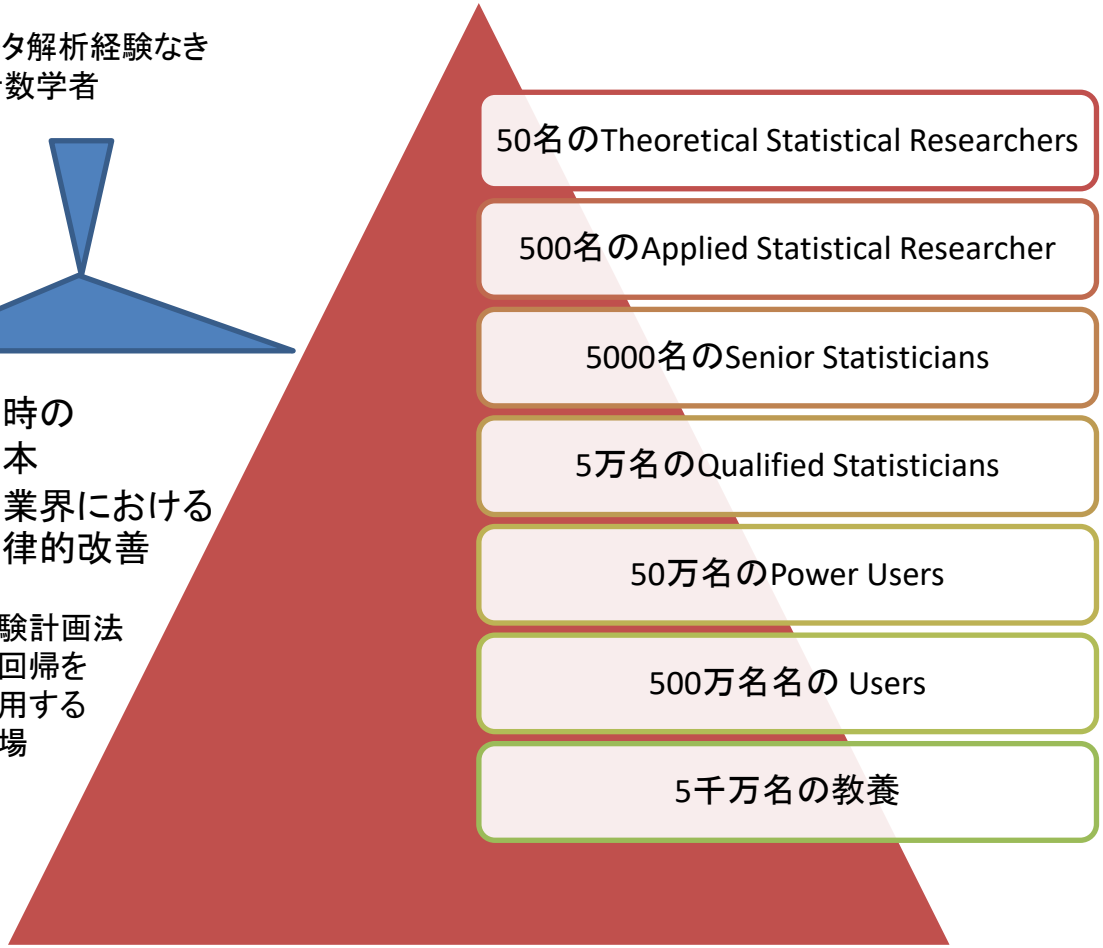
アメリカ数学会
Cobbs Report

データ解析経験なき
統計数学者



当時の
日本
産業界における
自律的改善

実験計画法
重回帰を
活用する
現場



2011概算要求時説明資料

不思議の国のビッグデータ時代

Q1:世界にあって，日本にない職種

→「**統計家**」

Q2:世界の大学にあって日本にない学科

→統計学科

ビッグデータへの対応

〇米国産業界;情報産業の高度情報コンサルテーションサービスへの転身

サービス科学の提唱

IBMによるSPSSの買収:予測型ビジネス

米国ビッグデータイニシアティブ

数理統計の延長線上としての機械学習，マイニングの基礎理論，応用の強化

〇日本の対応

情報2次産業がビッグデータ主導

大規模情報転送，圧縮

推進できる統計家の情報産業界不在

1990年代までは

日本のモノづくり産業はKAIZENで

統計最先端活用で世界から注目

〇米国労働省:**SCANSレポート(1990)**

自律的に問題解決可能な能力を

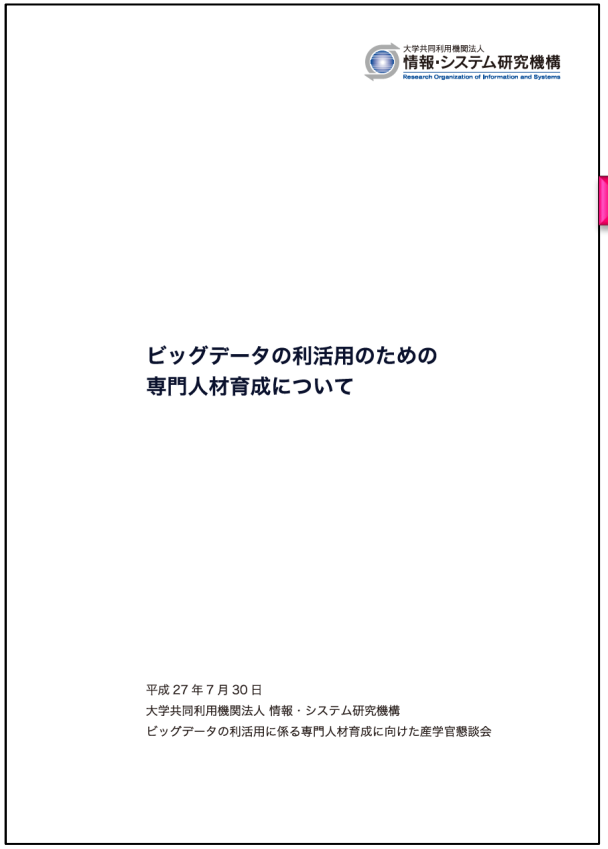
労働者に提供すべく，

学校教育の改革・統計重視

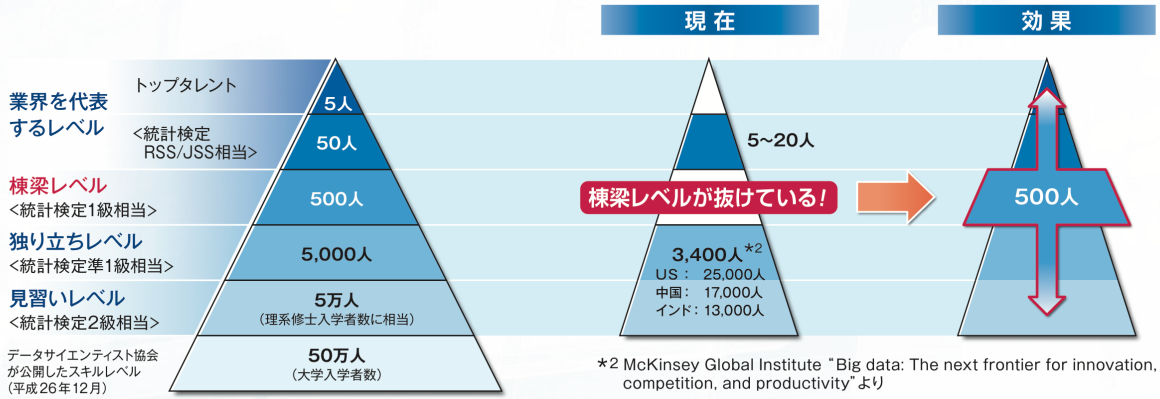
参考資料:　世界各国の統計学科ないしは統計関連組織のファカルティ数についての概要

国名	組織名	教授	准教授	助教	講師	コンサルタント	研究員	兼任	研究者計
インド	インド統計研究所(スタッフ1000名以上)						255		255
日本	**統計数理研究所**	**19**	**17**	**9**		**2**	**28**		**75**
ドイツ	ルートヴィヒ・マキシミリアン大学統計学研究所	11					55		66
イタリア	ボローニャ大学統計学科:Faculty 64名						64		64
オーストリア	ウィーン大学統計・OR研究所(Faculty22名、講師・客員35名)						22	35	57
フィンランド	*ヘルシンキ大学数学統計学科(統計学内訳不明)*	19	1		18		7	11	56
オーストラリア	*メルボルン大学数学統計学科(統計教員比不明)*	18	9		12		7		46
中国	中国人民大学統計学院	24	13	1	6	1			45
ニュージーランド	オークランド大学統計学科	12	9		15		8		44
英国	*ケンブリッジ大学純粋数学数理統計学科(統計は1/3程度)*	24			18				42
英国	オックスフォード大学統計学科	13			16		12		41
米国	ミシガン大学生物統計学科:ファカルティ数:40	40							40
米国	ワシントン大学生物統計学科	25	8	6					39
米国	UCLA統計学科	10		1	8			20	39
バングラディシュ	ダッカ大学統計生物統計情報学科	16	6	10	4				36
中国(台北)	中央研究院統計科学研究所(35名)						35		35
米国	スタンフォード大学統計学科	16	2	2	0	2	5	7	34
シンガポール	シンガポール国立大学統計・応用確率学科	7	14	6	6				33
米国	ミシガン大学統計学科	16	7	3	4			1	31
米国	カリフォルニア大学バークレー校統計学科	21	2	2	1	1		2	29
米国	コロンビア大学統計学科	12	4	11	1			1	29
米国	デューク大学統計科学科	12	5	7	1		4		29
チェコ	カレル大学確率数理統計学科	10	7	9	2	1			29
中国	香港大学統計学・保険数理学科	7	7	3	6		5		28
米国	シカゴ大学統計学科	16	3	3	2			2	26
米国	ワシントン大学統計学科	16	4	4	1				25
スェーデン	ウプサラ大学統計学科	4			9		12		25
中国(台北)	国立台北大学統計学科	1	10	9			5		25
米国	カーネギーメロン大学統計学科	13	6	3			2		24
カナダ	トロント大学統計科学科	10	7	2	4				23
オランダ	ユトレヒト大学方法論・統計学科	4	4	6	8				22
中国(台北)	国立成功大学統計学科	5	8	3				6	22
南アフリカ	ステレンボッシュ大学統計・保険数理学科	6			7		8		21
中国	北京大学統計科学センター:ファカルティ18名？						18		18
フランス	ノートルダム大学応用計算数学・統計学科	3	3	10					16
イスラエル	エルサレム大学統計学科	11			5				16
スリランカ	コロンボ大学統計学科	1	1		12		2		16
米国	ハーバード大学統計学科	9	2	2	2				15
ナイジェリア	イバダン大学統計学科	0	0	0	14				14
ガーナ	ガーナ大学統計学科				7			3	10
インド	カルカッタ大学統計学科	7	2						9
韓国	仁荷大学統計学科	7							7
韓国	江陵原州大学情報統計学科	7							7
デンマーク	ウルム大学Institute of Stochasric	2	2				2		6
クウェート	クウェート大学数学統計学物理学科統計学コース	2			4				6
イラン	Yazd大学統計学科:Faculty　6名	6							6
日本	**慶應義塾大学数理科学科統計学コース**	3	1	1					5
トルコ	ヤサル大学統計学科	1	1	1			1		4

統計数理研究所が担う育成人材のレベル



ねらいは「棟梁」の育成



*2 McKinsey Global Institute "Big data: The next frontier for innovation, competition, and productivity"より

上のピラミッド型の図は、1年間に育成すべきデータサイエンティストの人数分布を表しています*1。大学では基本的なデータリテラシーを身に着け、理系修士では統計検定２級程度のスキルの習得が期待されています。その上で、「独り立ちレベル」と呼ばれる大企業で活躍するデータサイエンティスト（資本金10億円以上の企業数から推定）、さらに独り立ちレベルを指導統括するリーダーとなる「棟梁レベル」のデータサイエンティストが必要とされています。

ところが現在、我が国ではデータサイエンティストのピラミッドがきわめて細いこと、特に「棟梁レベル」のデータサイエンティストが不在であることが深刻な問題として指摘されています*1。「棟梁レベル」人材が不在では、「独り立ちレベル」人材の活躍の場も制限されてしまいます。そこで本事業では、「棟梁レベル」人材を戦略的に育成します。「棟梁レベル」の人材育成の効果は、同レベルの充実にとどまりません。「棟梁レベル」の充実は、分野をリードする世界的なトップタレントが生まれる土壌となることが期待できます。同時に、「棟梁レベル」の人材の存在は、大学や企業でのデータサイエンティスト育成の促進につながります。

*1 情報・システム研究機構ビッグデータ利活用に係る専門人材育成に向けた産官学懇談会「ビッグデータ利活用のための専門人材育成について」より

科学技術 総合戦略

科学技術イノベーション総合戦略 （2015年6月 閣議決定）

○ 科学技術基本計画の中長期方針の下、各年度の重点項目を明確化

○ 成長戦略の一環として、毎年策定、閣議決定

○ 第5期科学技術基本計画の始動に向けた3つの政策分野

⇒ **大変革時代の挑戦**、地方創生、2020東京ｵﾘﾝﾋﾟｯｸ･ﾊﾟﾗﾘﾝﾋﾟｯｸ

重点的取組

・**「超スマート社会」の実現に向けた共通基盤技術や人材の強化**

（IoT、**ﾋﾞｯｸﾞﾃﾞｰﾀ解析、数理科学、AI、**ｻｲﾊﾞｰｾｷｭﾘﾃｨ、ｾﾝｻ、ﾛﾎﾞｯﾄ、素材、ﾅﾉﾃｸﾉﾛｼﾞｰ等）

「重点的課題」に、**強烈な問題意識**が示された

第1章 大変革時代における挑戦 2. 重点的課題

「我が国では欧米等と比較し、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く、・・・危機的な状況にある。」

統計エキスパート
人材育成プロジェクト創成の経緯

プロジェクト創設の経緯 ～ ＡＩ戦略 2019

ＡＩ戦略 2019 （2019年6月 統合イノベーション戦略推進会議決定）

○ **3つの理念**（「人間尊重」、「多様性」、「持続可能」）を掲げ、
Society 5.0 を実現し、SDGs に貢献

○ **4つの戦略目標**を設定（人材、産業競争力、技術体系、国際）
～ **AI時代に対応した人材を育成、**持続的に実現する仕組を構築

小中学校	高校	大学	社会人
基礎的学力・情報活用 【100万人卒/年】	文理問わず数理・データ関連教育 【100万人卒/年】	AI・数理・データサイエンス教育/エキスパート教育 【50万人卒/年】	リカレント教育/待遇 【多くの社会人に教育機会を提供】
	大学入試 **応用基礎を重視する入試に採用する大学への重点支援**	**応用基礎** **大学・高専生が自らの専門分野へのDS・AIの応用力を習得（25万人規模/年）** ➢ **AI×専門のﾀﾞﾌﾞﾙﾒｼﾞｬｰ**を可能とする環境 ➢ 専門教育レベルの**コース認定の導入**	**エキスパート** **年間2000人、トップ100人育成** ➢ PBL中心の**AI実践スクール制度** ➢ 若手の**海外挑戦機会の拡充**

大学を始め各段階で人材の育成を進めてきたが・・・

○ ここで顕在化した問題が、**大学の統計教員の不足**

大学共同利用機関法人 情報・システム研究機構 統計数理研究所

国民各層への関連施策との連携

関連施策との関係と各ターゲット

AI戦略2019における目標

- トップレベル 100
- エキスパート 2,000
- 応用基礎 250,000（理系学生＋文系学生の1/3）
- リテラシー 500,000（大学・高専全員）

データ関連人材育成プログラム

- ○ データサイエンス等のスキルを習得させる研修プログラムを開発・実施し、キャリア開発を支援
- ○ アカデミアにおける多様な分野の研究者や、コンソーシアムに参加する企業等で活躍する高度人材を育成
- ○ H29年に４コンソーシアム、その後１コンソーシアムを採択

数理・データサイエンス・AI教育の全国展開

- ○ 全国的なモデルカリキュラム・教材の開発、教育に活用可能な社会の実課題・実データの収集・整備等
- ○ 6大学を拠点校として設置。更に30大学を協力校・特定分野協力校、61大学を連携校として設置（令和２年８月31日現在）
- ○ 全国の国公私立大学等への普及・展開
- ○ 私立大学の取組への支援強化

課題：統計の教育・研究の中核人材の不足

統計エキスパート人材育成プロジェクト（新規）

統計学の教育・研究のエキスパート人材を育成

- ○ 人材育成プログラムと共同研究により統計学のエキスパートに育成
- ○ 育成された人材が大学等で核となり、統計学や融合領域に係る教育・普及を行う好循環体制を構築
- ○ 統計学を用いた他分野との融合領域の研究振興を図る
- ○大学・大学共同利用機関法人等のコンソーシアムを公募により支援

出典：文部科学省 公募説明会 資料

文部科学省による公募 ～ 背景と目標

背　景

○ データ駆動型研究の進展に伴い、統計的素養を必要とする
リアルタイム・**ビッグデータ解析等へのニーズが増大**

○ しかし、多数の統計学部から多くの専門人材を輩出する諸外国に比べ、
我が国では、**高度な統計学のスキルを有する人材は極めて少ない**

目　標

○ 大学院生等に講義・指導などができる**大学統計教員**を育成し、
これらの教員が中核となって更に**統計エキスパート**を育成する
体制を構築（統計エキスパート人材育成エコシステム）

① 事業期間5年間で、少なくとも**約30名の大学統計教員**を育成
② ①の教員が更に人材を育成し、10年間で
約500名の統計エキスパートを育成

大学統計教員：　専門の学術領域で**博士の学位**を有し、**大学院修士水準の統計学の講義**や統計活用研究の指導を行うことのできる人材

統計エキスパート：**統計を駆使**して学術研究や産業界等に貢献できる人材。
このプロジェクトで育成するのは、**大学院修士水準**の統計エキスパート。

文部科学省による公募 ～ コンソーシアム設立

コンソーシアムの設立

事業の目標を達成するため、**大学と大学共同利用機関により**
コンソーシアムを設立。

コンソーシアムの構成機関

「中核機関」: 中心的な役割を果たす機関
「参画機関」: 中核機関と協働して事業を実施する機関

※ 事業終了後も、人材育成エコシステムを自立的に継続・発展させる
※ 他の関連施策（「数理・ﾃﾞｰﾀｻｲｴﾝｽ・AI教育の全国展開」
「ﾃﾞｰﾀ関連人材育成ﾌﾟﾛｸﾞﾗﾑ」）と連携しながら取り組む

対象とする取組

○ **人材育成**に関する取組（人材育成プログラムの開発、共同研究 等）
○ **コンソーシアム構築**に関する取組（コンソーシアムの設立・運営・活動）

対象とする経費

（令和3年度総額 3億1,200万円）

・シニア教員・業務担当者等の雇用、人材育成プログラムの開発、施設・設備の購入・賃借、共同研究・学会発表、育成対象者の業務代替、コンソーシアム運営
※ 育成対象者自身の人件費は、補助対象外

大学共同利用機関法人 情報・システム研究機構 統計数理研究所

統計エキスパート育成プロジェクトとコンソーシアムの形成

大学共同利用機関法人 情報・システム研究機構
統計数理研究所

文部科学省による公募 ～ 採択の際の留意事項

応募申請の採択

○ これらの審査を経て、大学共同利用機関法人 情報・システム研究機構 2021/06**統計数理研究所** による応募申請が、**文部科学省により採択**

○ 委員会の面接審査での指摘事項を中心に、**次の留意事項**が提示

採択に際しての留意事項

- 参画機関の積極的な協力の下、コンソーシアム全体として**密接な連携体制**を構築
- **シニア教員の質**を適切に管理し、研修担当教員との役割分担を明確に
- 自身の研究業績につなげるなど、**育成対象者の今後のキャリア形成**に配慮
- **国際的に活躍できる人材**の育成につながる機会の創出等に向け、検討
- **世界水準に見合う教科教育科目**を研修カリキュラムに設定
- 人材育成が計画どおりに進まなかった場合の対応策等の**リスク管理**を検討
- 人材層を厚くするため、**統計学以外の分野からの育成対象者**も積極的に選抜
- 事業終了後の発展に向け、**博士レベルの統計エキスパート**の育成方策を検討

プロジェクトの概要 ～ 具体的な目標

統計学の大学院課程の現状

- **総合研究大学院大学** 複合科学研究科 統計科学専攻（博士課程、学年定員5名、統計数理研究所が基盤機関）が、これまで唯一
- **滋賀大学** 大学院データサイエンス研究科が、2019年に新設

具体的目標：統数研

- 統数研は、2年間の**「大学統計教員育成研修」**を、3期にわたり実施
- **各期10名以上**の若手研究者（助教、ポスドク等）を、大学等から受入れ
- 若手研究者2名につき、おおむね1名のシニア大学統計教員を、メンターとして割り当て（第1期は、**6名のシニア教員**を雇用）

具体的目標：大学等

- 大学等は、事業期間中に、**1名以上の若手研究者**を研修に派遣
- 大学内に、**統計エキスパートを育成するシステムを構築**
- 育成された大学統計教員1名につき、原則として**毎年3名以上の大学院修士水準の統計エキスパート**を育成

大学共同利用機関法人 情報・システム研究機構 統計数理研究所

プロジェクトの概要 ～ コンソーシアムの運営

統計エキスパート人材育成コンソーシアム

構成員

中核機関 ～ 統計数理研究所（中心的役割を果たす機関）
参画機関 ～ 全国21の大学等（協働して事業を実施する機関）
協力機関 ～ 全国 5の大学等 （事業に協力する機関）

※ 参画機関には、事業担当教員と研修担当教員を配置
※ 協力機関は、教材・カリキュラム開発、知識・力量測定などに協力する機関、事業期間中に参画機関となる見込みの機関など

○ **協力機関** ～ コンソーシアムの活動がより効果的に進み、参画機関の拡充がより円滑に進むよう、構成機関に追加。

運営組織

・**総　会** ～ 中核機関・参画機関・協力機関の全会員
・**運営委員会** ～ 中核機関・統数研5名、参画機関・大学等5名
・**事務局** ～ 統計数理研究所 大学統計教員育成センター

大学共同利用機関法人 情報・システム研究機構 統計数理研究所

参　考

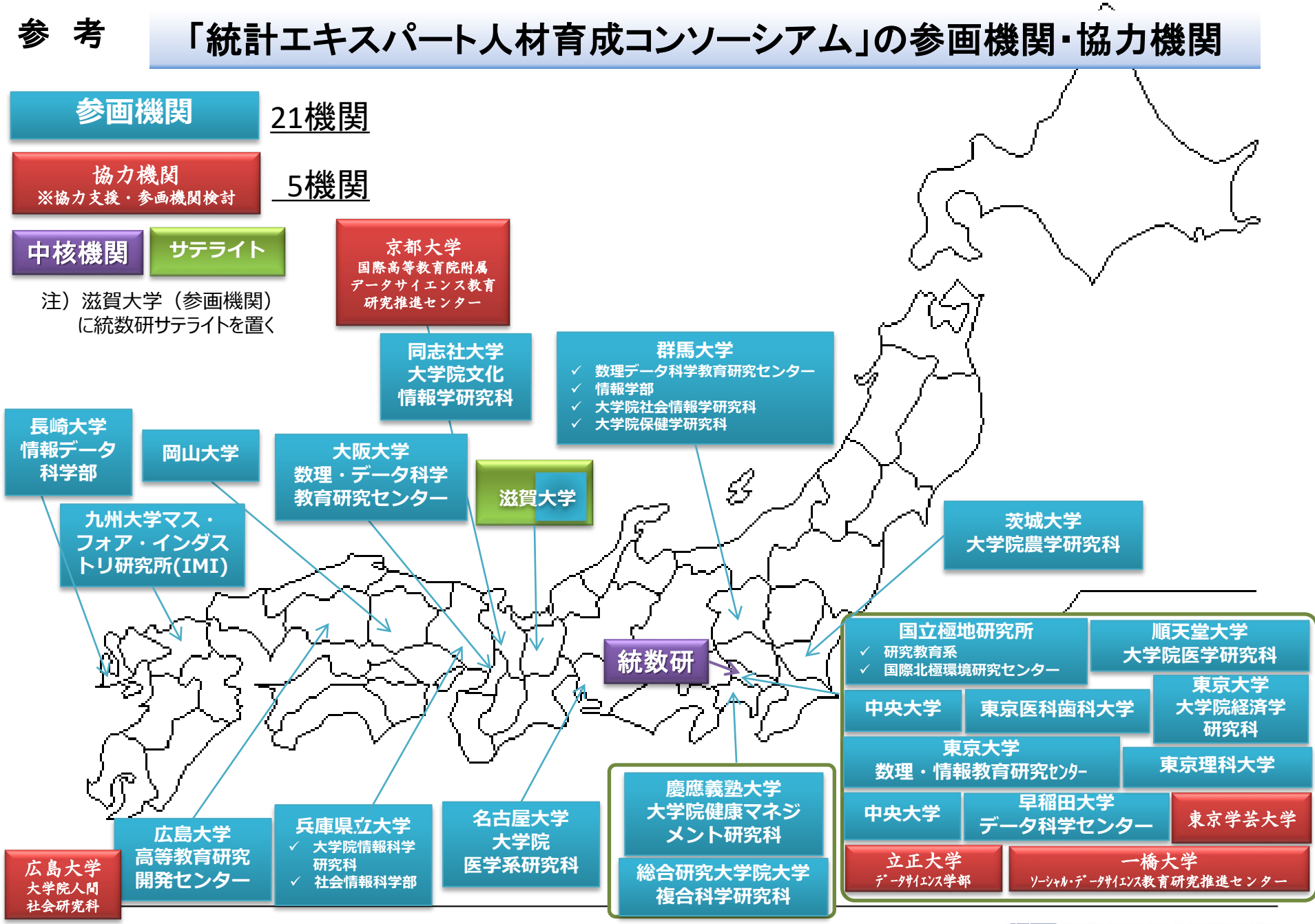
「統計エキスパート人材育成コンソーシアム」の参画機関・協力機関

参画機関 21機関

協力機関
※協力支援・参画機関検討 5機関

中核機関　**サテライト**

注）滋賀大学（参画機関）
に統数研サテライトを置く



大学共同利用機関法人 情報・システム研究機構
統計数理研究所

コンソーシアム運営・組織の概要

運営組織:2021/08始動

総会
・中核機関、参画機関、協力機関
・基本的事項の審議

運営委員会
・中核機関、選出された参画機関
・事業実施に関する重要事項の審議

意見の反映
運営方針

（今後も拡充）
参画機関
統計エキスパート育成システムの開発と実装

事業・研修担当教員、育成対象者、アシスタント、シニア教員（助言）

事務局:2022/01/01設立
統計数理研究所　大学統計教員育成センター

統括部
コンソーシアムの運営

研修部

研修主幹
研修の実施
（シニア教員、育成対象者、アシスタント）

教育システム開発主幹
教材・カリキュラム開発

統計教員育成要請
質保証された統計教員育成
教材・カリキュラムの提供
参画機関に移行

統計数理研究所
データサイエンス共同研究基盤施設
・多様な共同研究
・多様なデータの研修事業への提供

連携

協力

協力機関
・教育システム開発
・人材質保証等
・参画検討機関

統計数理研究所 サテライト：
滋賀大学（研修の西の拠点）

プロジェクトの概要 ～ 研修の基本的考え方

大学統計教員 育成研修の基本的考え方

1. 主たる目的は、大学統計教員の育成

～統計学者の育成ではなく、育成対象者の専門分野を尊重

2. 修士課程学生に対する 4科目の講義を可能にする

～「基本的な統計学の講義」・「発展的な統計学の講義」・「専門分野と統計学が融合した講義」（2科目）

3. 研修の基本は、講義体験の繰り返しと達成度の確認

～ 達成度WG（育成対象者・研修担当教員・メンター・副メンター等）により、各育成対象者の研修進捗状況・達成度を確認

4. 育成対象者の研修計画・研究計画を、半年ごとに策定

～シニア教員は、全員で互いに協力して研修をサポート

5. 教員育成の方法を開発する

～ 日本には、大学教員を育成するという発想がない

大学共同利用機関法人 情報・システム研究機構 統計数理研究所

プロジェクトの概要 ～ 研修科目・カリキュラム

育成対象者全員に必要なスキル

基本科目

（基本1）「統計ベースライン特習」
分野共通の統計学的知識について、
不足している知識体系を補充

（基本2）「統計教育力育成演習」
模擬講義や作問練習などによって、
目指す分野における教育力を育成

（基本3）「統計研究力強化演習」
論文執筆技術、着実に研究業績を
挙げる能力。准教授以上に要する、
学生に研究させる能力＝研究指導力
（参画機関でのFD活動）

各自の研究ニーズに応じたスキル

個別科目

（個別１）先端的データサイエンス特論
（例） Ⅰ 因果推論、Ⅱ 先端的機械学習、
Ⅲ スパースモデリング
先端的データサイエンスの理論を習得。
対象者に要する技術によって科目を選択

（個別2）先端的データ分析演習
実データの性質にあった分析方法を選択。
最先端の統計理論、機械学習技術を実際に使用

（個別3）コンサルテーション演習
実問題に対するコンサルテーション能力を育成

（個別4）教材開発演習
対象者の進路にあった教材開発技術を習得。
教科書の執筆、教育用データの作成のスキルも学ぶ

（個別5）共同研究演習
企業や公的機関からの研究課題に対し、
ステークホルダーや教員とともに共同研究

プロジェクトの概要 ～ 育成対象者とシニア教員

第1期研修 育成対象者とメンターの配置 （A方式11名）

参画大学	学　位	メンター	副メンター
東京医科歯科大学	博士（工学）	岩崎	折笠
兵庫県立大学	博士（情報学）	中西	岩崎
同志社大学	博士（文化情報学）	中西	岩崎
群馬大学	博士（保健学）	岩崎	折笠
岡山大学	博士（経済学）	国友	中西
東京理科大学	博士（薬学）	水田	中西
東京理科大学	博士（情報科学）	国友	岩崎
滋賀大学	博士（工学）	神保	中西
滋賀大学	博士（工学）	神保	中西
長崎大学	博士（理学）	岩崎	折笠
国立極地研究所	博士（情報）	水田	岩崎

【 メンター （シニア教員）】

岩崎　学、折笠秀樹、国友直人、神保雅一、中西寛子、水田正弘

大学共同利用機関法人 情報・システム研究機構 統計数理研究所

プロジェクトの概要 ～ 第1期研修の開始！

第1期研修 開講式

【日　時】 2021年10月4日（月）

【場　所】 オンライン開催（ZOOM）

【参加者】 育成対象者、研修担当教員（所属大学等）、
シニア教員（統数研）、文部科学省など、50名超

祝辞、シニア教員紹介、育成対象者自己紹介

第1回 全体研修

【日　時】 2021年10月18日（月）

【場　所】 オンライン開催（ZOOM）

【参加者】 全ての育成対象者・シニア教員

【内　容】 **育成対象者（教師役）による講義体験**

各自の研究分野について、講義40分、質疑20分

大気リモートセンシング、「線虫」の健康寿命、多目的最適化問題・・・

大学共同利用機関法人 情報・システム研究機構 統計数理研究所

プロジェクトの概要 ～ 研修のスケジュール

大学統計教員育成研修　スケジュール

第1期：　2021年10月　～2023年 9月
第2期：　2023年 4月　～2025年 3月
第3期：　2024年 4月　～2026年 3月

各期4クール（半年ごと）で、基礎から発展、個別科目へ

育成研修	2021年度	2022年度	2023年度	2024年度	2025年度
第1期研修	←→	←→	←→		
自己点検		←→			
第2期研修			←→	←→	
中間評価			←→		
第3期研修				←→	←→

評価・改善の取組

2022年10月（第1期開始1年後）～**自己点検**（統数研）⇒第2期の改善へ
2023年10月（事業開始3年目）　～**中間評価**　（文科省）⇒第3期の改善へ

大学共同利用機関法人 情報・システム研究機構 統計数理研究所

アメリカから今学べること

アメリカ合衆国
労働統計局雇用統計2019/05

https://www.bls.gov/ooh/math/mathematicians-and-statisticians.htmなどより作成

- 米国政府職業小分類867職種中：
  「統計家」第5位、「情報セキュリティアナリスト」第10位の成長率予測
  - 「第1位は風力タービンサービス技術者（7000名：62%成長）
  - 統計家も2018年から職業分類に追加された
    データサイエンティストも数理科学職（主として産官で活躍）

アメリカ合衆国標準職業分類	分類コード	雇用者数	10年成長率	トップ産業雇用者数
統計家	15-2041	42,700	34.6%	科学研究開発,6190
データサイエンティスト/その他	15-2051/99	33,200	30.9%	計算機システム設計,9100
オペレーションズリサーチアナリスト	15-2031	105,100	24.8%	企業マネジメント,9930
アクチュアリー	15-2011	27,700	17.6%	保険会社,11,140
数理科学（中分類）	15-20XX	211,700	26.5%	
情報セキュリティアナリスト	15-1212	131,000	31.2%	計算機システム設計,36,280
計算機科学（中分類）	15-12XX	4,633,400	11.5%	

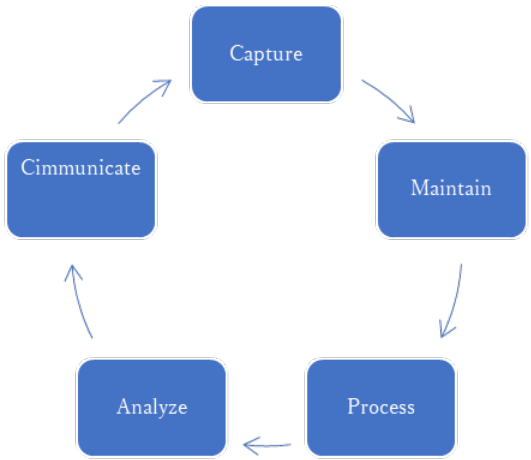
米国労働統計局が示す社会における「統計家(15-2041)」の役割

- 特定の質問や問題に答えるために必要なデータを決定
- ビジネス、工学、科学、その他の分野における実用的な問題を解決するために、数学理論と数理技術を適用
- データ収集のための調査・アンケート（設問の設計、ターゲットとなる集団からの適切な標本の決定。サンプリング、意思決定に必要なサンプルの大きさの決定）、実験を計画
- データ分析のための数理モデル・統計モデルの開発
- データを解釈し、専門家および専門家ではない者に分析をレポート
- 統計ソフトウェアを利用したデータを分析し、トレンドや関係性を明らかにし、データの妥当性や限界も検討
- レポート（ビジュアライゼーション、集計表）を作成して、ビジネス等の意思決定を支援
  - **42,700名中：全米146統計学専攻の修士統計等取得者が65%、**
  - **博士取得者が20%、学部卒業が15%**

データ・サイエンティスト(15-2051)のミッション

- データ・オリエンテッド・プログラミング言語と視覚化ソフトウェアを使用して、原データを意味のある情報に変換するための一連の手法または分析アプリケーションを開発および実装
- データマイニング、データモデリング、自然言語処理、機械学習を適用して、大規模な構造化データセットと非構造化データセットから情報を抽出して分析
- データの調査結果を視覚化、解釈、および報告
- 動的データレポートを作成
  - ただし、『統計家（15-2041）』、『地図製作者および写真製作者（17-1021）』および『健康情報技術者および医療登録者（29-9021）』を除外
  - 現時点では学部卒業生が主体
- データサイエンティスト協会のスキルセットよりは狭い
  - データサイエンス力×データエンジニアリング力×ビジネス力
  - https://www.datascientist.or.jp/common/docs/skillcheck_ver3.00.pdf

米国データサイエンス大学院ランキング1位：UC Berkley:　オンライン大学院



データサイエンスのライフサイクル
これができるエキスパートを育成

- Capture
  - データ取得、データ入力、信号入力とデータの抽出
- Maintain
  - データウェアハウスの作成、分析可能なデータの編成（クレンジング、ステージング）、データ処理、データアーキテクチャー
- Process
  - データマイニング、クラスタリング／分類、データモデリング、データの要約
- Analyze
  - 探索的分析と検証的分析、予測分析、回帰分析、テキストマイニング、質的分析
- Communicate
  - レポーティング、可視化、ビジネスインテリジェンス、意思決定

専門職修士課程（産業界への人材輩出）そのカリキュラム

DS修士課程基礎コース
2.5（統計学）： 2（計算機科学）： 0.5（応用AI）

科目名	キーワード
DSプログラミング入門	オブジェクト指向、Pythonによるデータ分析、モジュール等の開発、Jupyter Notebookによるプレゼン、GitとGithub
研究計画とデータや分析への適用	研究計画、研究の問いかけの定式化、データと意思決定、認知バイアスの理解、説得とアクションのためのデータ、データと固有知識の統合、データで物を語る
DSのための統計学	研究計画、様々な量的研究方法と統計的分析技法、記述統計、推測統計、サンプリング、実験計画、仮説検定、最小二乗法と回帰分析、ロジスティック回帰、Rによる実社会データ分析実践
データ工学の基礎	データセットの保存、管理、処理システムの基礎、分析によるソリューションのアーキテクチャー、分散データ処理、リレーショナルデータベース、グラフデータベース、ストリーミング分析、クラウドコンピューティング
応用機械学習	実験計画、機械学習アルゴリズム、特徴量エンジニアリング、予測か説明か、ネットワーク解析、協調フィルタリング、

DS修士課程発展コース
2: 1: 2: 2(コミュニケーション・倫理・マネジメント)

科目名	キーワード
実験と因果推論	実験計画に基づく因果関係の推論。無作為化、統計分析と結果の伝達、データのクリーニング、データマイニングと探索
データの背景：人間と価値	倫理と法的枠組み、政策分析、DSライフサイクル全体での倫理問題、データの収集やタイプなどへの倫理・政策的制約、ケーススタディ
クラウドとエッジでの深層学習	画像やビデオなど大規模データの深層学習などの処理環境の実践、クラウド、分散ストレージ、イーサリアムブロックチェイン、Keras, IBM Watson, 利用可能なApacheのソフトウェア
質的回答・時系列・パネルデータのための統計的方法	断面的データと時系列データの可視化、確率と数理統計の重要概念、古典的回帰分析、変数変換、モデルの規定、因果推論、操作変数法、時系列モデル（ARIMA、GARCH、VARモデル）、統計的予測、時系列データを伴う回帰分析
大規模機械学習	単一あるいは複数計算機上での機械学習アルゴリズムの実装、AWS, テラバイト級データの作業における問題点、ペタバイト級データの機械学習パイプライン、アルゴリズムの設計（決定樹、グラフ処理、最急降下法）、並列処理
深層学習を用いた自然言語処理	言語現象の紹介と機械学習による分析、情報抽出、機械翻訳、センチメント分析、要約
データの可視化	探索的データ分析、効果的な文書コミュニケーション、効果的なデータの可視化プレゼンテーション、人間の知覚のデザイン

Berkleyの企業幹部向けDSリテラシーコース
10週間70時間オンライン2850$:

- 確率的意思決定
  - データ分析の前に離散・連続データの比較等DSの背後にある基礎概念等
- 標本データの生成
  - 用語、抽出方法の種類、標本データの知る、第1種の過誤と第2種の過誤等
- 仮説検定
  - 仮説検定、信頼区間、実験の基本原理、4M（動機、方法、メカニズム、メッセージ）モデルによる問題の定式化
- 標本データからの情報推論
  - 直線パターンや曲線パターンの探索とデータに線形モデルを当てはめる様々な方法の理解と産業界での応用
- 基本的回帰モデル
  - より精緻なビジネス意思決定の理論の機関となる単回帰分析の使い方、
  - その利用上の注意とビジネスの意思決定がどう改善するのか
- 発展的回帰モデル:重回帰分析の様々な利用方法
- 予測と機械学習
  - 機械学習の基礎原理と様々な応用を分かりやすく解説、教師付き学習と教師無学習、時系列回帰予測等
- **A/BテストとのDSチームの効果的構築**
  - 組織のデータ主導型文化構築、データサイエンティストとの効果的連携戦略と陥りやすい問題

Berkley 統計学科(J. Neyman設立)
学部レベル

- 確率・統計・数理：確率・統計入門
  ゲーム理論、確率過程
- 統計的推論：データ、推測、意思決定
- 統計モデル：線形モデル、時系列入門
- データ収集技術：標本調査、実験計画法
- 統計計算・データ解析：データに関する計算概念
  Rによるプログラム入門・発展
- DS準備：DSのための線形代数、DSのための確率
  DSのための統計的方法
- DS：DSの基礎、現代統計的予測と機械学習

大学院(オンラインではない)

- 確率論
  - 確率論、確率と確率過程の発展、確率の応用、数理金融への確率解析の応用
- 統計学基礎　理論統計
  - 理論統計のトピックス、**確率と統計のトピックス**
- 統計モデル
  - 統計モデルの理論と応用、時系列分析
- 統計的推論
  - 因果推論、ノンパラメトリック技法とロバスト技法
- 統計計算
  - 統計計算入門
- DS
  - DSの原理と技術、再現・検証可能な統計的DS
- **統計専門職実務**
  - **統計コンサルテーション、確率統計の教育**
  - **研究　　統計研究セミナー**

弟子→孫弟子→ひ孫弟子

- UC Berkley統計学科:Neymanの36年間
- Neyman　1945~1981 35名Ph.D授与
  - 内6名が　Berkley 教員(Lehman, Le Cam, Hodges)
    - 79名のPh.D授与
    - 内2名がLehmanの弟子でUCB教員(Bickel, Doksum)
      - » Lehmanは1981年時点で34名の弟子
      - – Bickel&Doksumで1981年まででで31名Ph.D授与
        - » Bickelの弟子 : Jeff Wu
        - » Doksumの弟子、三浦良造先生、V. Nair
  - Neyman, Lehman, Bickel and Doksumの共通点
    - 古典的統計学の教科書・教程の創始者

おわりに

周回遅れの日本
本コンソーシアムが目指すべきもの

- 分野別基幹統計教育システムを
  研修生とメンターとで共創
  - **教育システム＝教材＋教育プロセス＋教育人財**
- コンソーシアムが協力してCAPDoサイクルで
  基幹統計教育システムを改善
- 基幹統計教育システムを各大学院に実装と
  その質保証制度の確立
- 基本的な力量を有する有為なエキスパート輩出
- その上で、各大学が独自の領域で競争