

Open for Data Summit Autumn



データサイエンティストの育成のための大学教育 ー 滋賀大学データサイエンス学部モデル

竹村 彰通

滋賀大学データサイエンス教育研究センター長

《本日の内容》

- ・ データサイエンスとは？
- ・ 海外の統計学の動向
- ・ 統計学からデータサイエンスへ
- ・ 滋賀大学の改革構想
- ・ 育成する人材像
- ・ カリキュラムの概要
- ・ 企業との連携 – 実際のデータにもとづく演習

データサイエンス－価値創造のための新たな科学

客観的な存在としての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科学

データサイエンス（滋賀大モデル）

データアナリシス
大規模データを分析・解析
するための知識とスキル
(統計学)



データエンジニアリング
大規模データを加工・処理
するための知識とスキル
(コンピュータ科学)

新たな知見



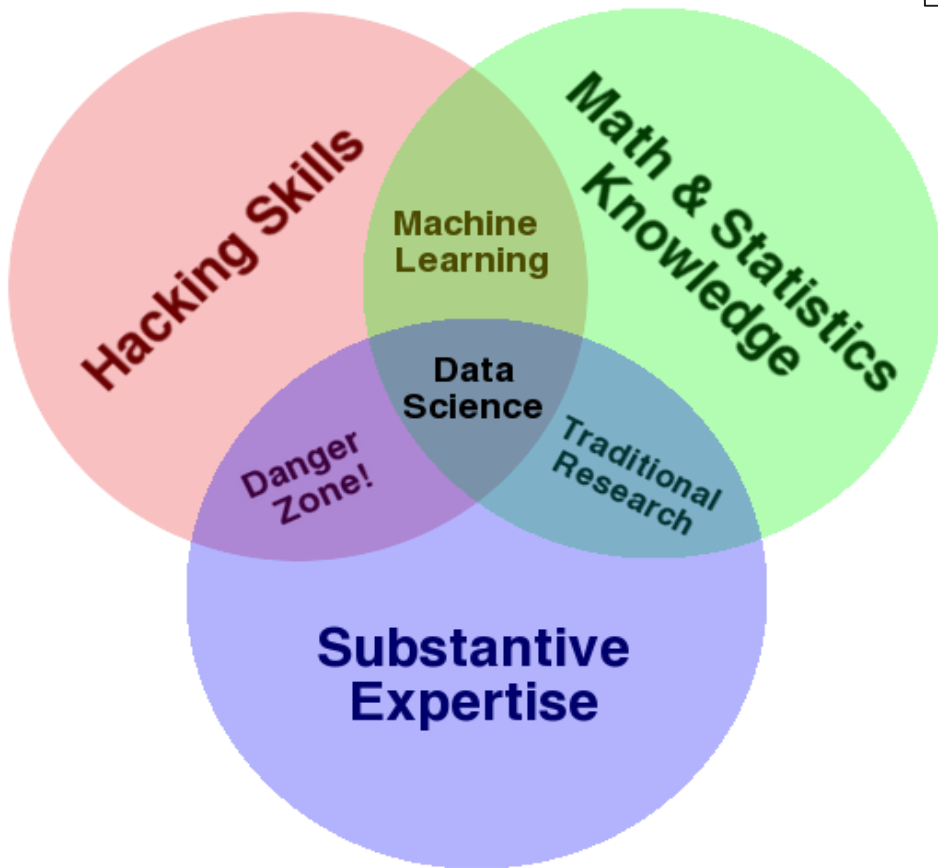
価値創造
ビジネスや政策などの領域で課題を読み取り、データエンジニアリングとデータアナリシスにより得られた知見を現場の意思決定に生かす
(領域分野での成功体験＝PBLの繰り返し)



データサイエンティスト

Data Science Venn Diagram

<http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>



DS = 統計 + 計算機 + 領域
分野の考え方は一般的

©Drew Conway Data
Consulting

データサイエンス分野の重要性

- 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」（「科学技術基本計画」（第5期））
 - 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、**統計学部がなかった**
- 「分野点在型」**

諸外国での統計学部の数

- アメリカでは100程度。
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250以上といわれており、さらに増えている
- これに対して日本はこれまでゼロ

米国における統計学関係の博士取得者数の推移 (2002-2012)

National Science Foundation, Survey of Earned Doctorates

Subfield of study	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Statistics(mathematics)	166	191	226	267	301	360	303	353	327	332	365
Mathematics/statistics, general	133	150	81	104	107	145	140	173	163	143	195
Mathematics/statistics, other	93	88	73	60	61	15	63	68	82	72	97
Biometrics and biostatistics	81	84	100	129	107	120	122	115	127	137	174
Statistics (social sciences)	54	48	31	22	22	29	20	17	21	22	18
Econometrics	14	23	18	30	33	30	36	31	34	28	40
Management information systems /business statistics	91	86	94	95	137	142	156	136	108	100	102
TOTAL	632	670	623	707	768	841	840	893	862	834	991

日本における統計学の博士取得者数の推移 (2002-2012)

- 総合研究大学院大学複合科学研究科統計科学専攻
(学術・統計科学)

2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
5	6	7	4	6	6	6	4	7	4	5

日本でなぜ統計学部ができなかったのか

- 日本の学部は産業分野にならう縦割り型。特に工学部
- 明治時代に大学が作られた時からの伝統
- 統計学や情報学のような「横串」の方法論は、「必要に応じて後から勉強すればよい」という傾向
- 統計学者の中にも、統計学は各分野で応用されてはじめて意味があるとする者も多かった
- 日本は少子化の影響で大学の学部再編は最近非常に困難

最近横串分野がイノベーションに貢献
日本は非常に遅れている

統計学からデータサイエンスへの海外の動向

- 海外の大学には、もともと統計学の独立した学部・学科が存在
- 独立した統計学のセクションがあると、情報工学、コンピュータ科学、ときには数学も加えて、**データサイエンス・プログラムが柔軟に生成されやすい**
- 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**されている
- **統計学専攻は近年の理系学位（STEM）の一番人気で、卒業後の給与も一番高い。**

アメリカの学部・大学院教育の動向

アメリカ統計学会ニュースレターから
(amstat news, February & April, 2015)

統計学・学士号	取得者比率 2011-13/ 2003-05	データサイエンスを意識したカリキュラム
Purdue University	875%	Big Data course programming language (C, Python, Java, etc.) プログラミング言語
University of California, Berkeley	224%	the upper division electives are almost all centered on data analysis (statistical learning theory) 上級選択科目としての統計的学習理論
University of Illinois, Urbana-Champaign	452%	new analytics courses emphasizing data management and statistical analysis of databases, Big Data methods ビッグデータ手法に重点化したコース a new statistical programming course 統計的プログラミング
Carnegie Mellon University	191%	experiential learning through the use of real data sets 実際のデータ利用 programming and software engineering in R, and databases and data management visualization, data mining プログラミング、R、可視化、マイニング

データサイエンス分野の人材不足

「第4次産業革命に向けた人材育成総合イニシアチブ」 ～未来社会を創造する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～

別紙 4

- 「第5期科学技術基本計画（平成28年1月閣議決定）」において謳われている「超スマート社会」の実現、及び「理工系人材育成に関する産学官円卓会議における行動計画」等を踏まえ、関連施策の一體的な推進が求められている
- 生産性革命や第4次産業革命による成長の実現に向けて、**情報活用能力を備えた創造性に富んだ人材の育成が急務**
- 日本が第4次産業革命を勝ち抜き、未来社会を創造するために、特に喫緊の課題であるAI、IoT、ビッグデータ、セキュリティ及びその基盤となるデータサイエンス等の人材育成・確保に資する施策を、初中教育、高等教育から研究者レベルでの包括的な人材育成総合プログラムとして体系的に実施**

参考：必要とされるデータサイエンス人材数(※)

- 世界トップレベルの育成（5人/年）
- 業界代表レベルの育成（50人/年）
- 橋梁レベルの育成（500人/年）

- 独り立ちレベルの育成（5千人/年）
- 見習いレベルの育成（5万人/年）

現状（MGLU-8-1）

日本：3,485人

US：25千人

EU：17千人

滋賀大最先

- リテラシーの醸成（50万人/年）

大学入学数/年：約60万人

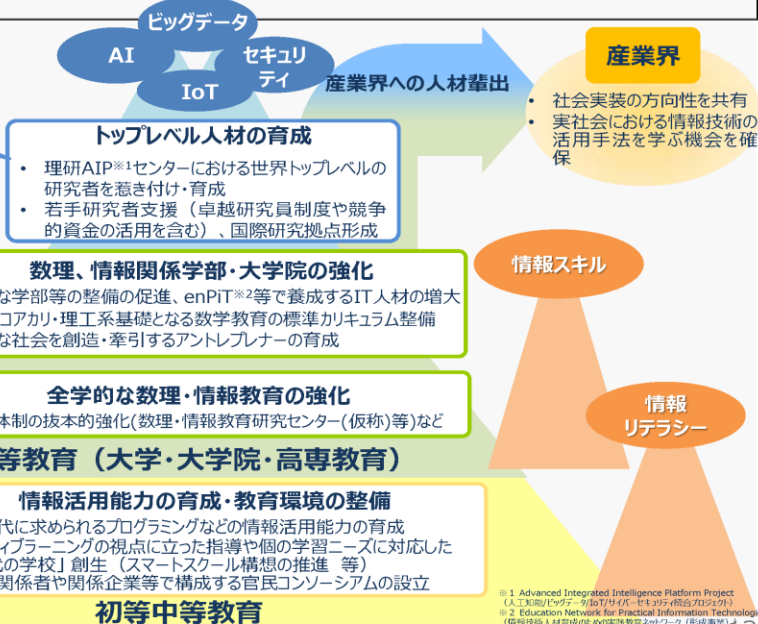
- 小学校における体験的に学習する機会の確保、中学校におけるコンテンツに関するプログラミング学習、高等学校における情報科の共通必修履科目化といった、発達段階に即したプログラミング教育の必修化

- 全ての教科の課題発見・解決等のプロセスにおいて、各教科の特性に応じてICTを効果的に活用
- 文科省、経産省、総務省の連携により設立する官民コンソーシアムにおいて、優れた教育コンテンツの開発・共有等の取組を開始

高等学校：約337万人（3学年）

中学校：約350万人（3学年）

小学校：約660万人（6学年）



※注：左吹き出しの人数は「ビッグデータの活用のための専門人材育成について」（大学共同利用機関法人情報・システム研究機構、平成27年7月）から引用

「科学技術イノベーションによる未来社会創造プラン」（p.13）
2016年4月文部科学省

平成28年4月19日 産業競争力会議文部科学大臣提出資料（関連資料）

「第4次産業革命に向けた人材育成総合イニシアティブ」 ～②高等教育段階における取組～

全学的な数理・情報教育の強化

～非数理・情報学部を含め、数理・情報の学修を強化～

・ 全学教育研究組織（数理・情報教育研究センター（仮称））の整備

※教育体制の抜本的強化、他分野と数理・情報を融合した教育研究の実施等

（例：大阪大学が数理・データ科学教育研究センターを2015年10月に設立）

・ 理工系の基礎となる数学教育の標準カリキュラムを整備

数理・情報の専門人材の育成強化

・ 数理、情報関係学部・大学院の整備・拡充の促進

※数理、情報関係学部・大学院の新設、定員増等

（H29年度新設を構想中 **滋賀大学（データサイエンス学部）**、名古屋大学（情報学部））

・ 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（enPiT※拡充、数理を産業に活かす実践教育の推進）

✓ 数理・情報教育プログラムの開発（大学における医療・農業・経営・公共政策等他分野と情報・数理を掛け合わせるプログラム開発、高等専門学校における情報教育パイロットプログラム開発）

✓ コアカリキュラムの策定（大学の情報教育コアカリキュラムの策定、高等専門学校のコアカリキュラムの導入促進）

など

※ Education Network for Practical Information Technologies（情報技術人材育成のための実践教育ネットワーク（形成事業））2

滋賀大学の改革構想 — 文理融合大学への変革

改革方針

人文社会系大学から文理融合型大学への 大転換

- 教育・経済2学部 & 分離キャンパスの大学からの脱却
- 社会的要請の強い新学部創設
- 専門職業人育成機能の強化

滋賀大学の強み・特色

- 経済学部のビジネススクール型教育
- 教育学部における情報教育課程



日本初のデータサイエンス学部創設
データ駆動型価値創造人材の育成
先行事例のない教育プログラムの確立

滋賀大モデル

データアナリシス

大規模データを分析・解析
するための知識とスキル
(統計学)

データエンジニアリング

大規模データを加工・処理する
ための知識とスキル
(コンピュータ科学)

新たな知見

価値創造

ビジネスや政策などの領域で課題を読み取り、データ
エンジニアリングとデータアナリシスにより得られた
知見を現場の意思決定に生かす
(領域分野での成功体験 = PBLの繰り返し)

- 価値創造を重視
- 職場で、一人でもデータ分析を担当できる人材
- コミュニケーション力を重視

アメリカでは
統計 + 計算機
だけで就職できる

データサイエンス学部における育成人材像

データ駆動型価値創造 逆T（またはΠ）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

逆T型

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

逆Π型

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

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)

データサイエンスの
専門知識とスキル

データサイエンス学部のディプロマポリシー

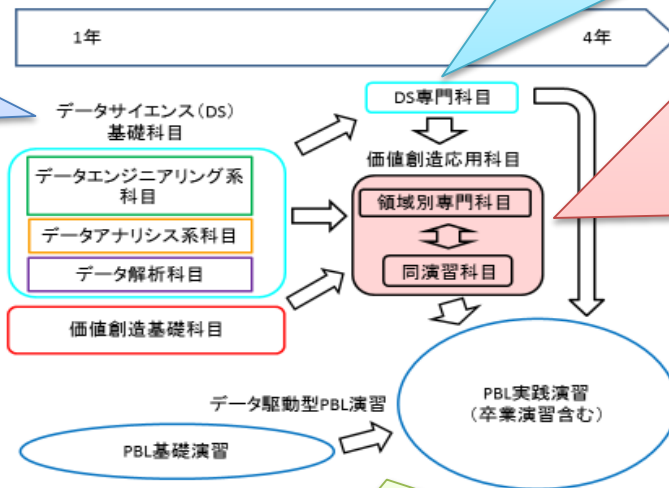
- データエンジニアリングとデータアナリシスの専門知識とスキルを修得し、データサイエンスの基礎的力量を備えている。
- データサイエンスの基礎を応用して、多様な領域でのデータ駆動型価値創造を導くための実装力を備えている。
- 多様なコミュニケーションの力量を備え、データ利活用の現場で相互補完的な専門性を有する仲間と協力して、組織目標を追求できる。
- データ駆動型価値創造社会の哲学・倫理・政治等について、バランスのとれた見識を有している。
- 上記のようなデータサイエンティストの専門的力量とイノベーティブな心の習慣を背景に、卒業後の現場での課題に対応して、自律的な学習を進めることができ、多様な領域における価値創造のための創造的イノベーションにも貢献できる。

履修モデル、想定される就職先、キャリアプラン

3つの履修モデル&人材像	想定される就職先
データエンジニア	• IT系企業、シンクタンク、製薬企業や医療機関等 • 企業のデータ分析部門 • 大学院
データアナリスト	• 企業のデータ分析部門 • 経営コンサルタント、シンクタンク、製薬企業や医療機関等 • 企業の財務経理、経営企画、マーケティング、生産管理等の諸部門 • 大学院
データコンサルタント	• 企業の財務経理、経営企画、マーケティング、生産管理等の諸部門 • 国または地方自治体 • 企業のデータ分析部門

カリキュラムマップ

- 統計学、コンピュータ科学・情報工学から構成
- 相対的に統計分析によるアナリシスを重視
- 汎用解析ソフトの利用・訓練



ビッグデータ解析のための発展的科目
データマイニング、テキストマイニング、
機械学習、ベイズ理論、最適化など

- 文理融合の実践を重視
- データサイエンス基礎を
応用する多様な領域を準備
- ・ マーケティング
- ・ ファイナンス
- ・ 会計
- ・ 医療・健康・福祉
- ・ ビジネスエコノミクス
- ・ 環境
- ・ 教育
- ・ 保険・リスク
- ・ 公的統計
- ・ 心理
- ・ 地域文化情報

- 本学DSプログラムの真骨頂ー現場のデータを利用した価値創造PBL演習での成功体験
- ナレッジマネジメント理論に基づくPBL演習の設計
 - DS教育研究センターにおける価値創造プロジェクトが企業、自治体、非営利団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

開講科目の詳細

共通科目
を重視

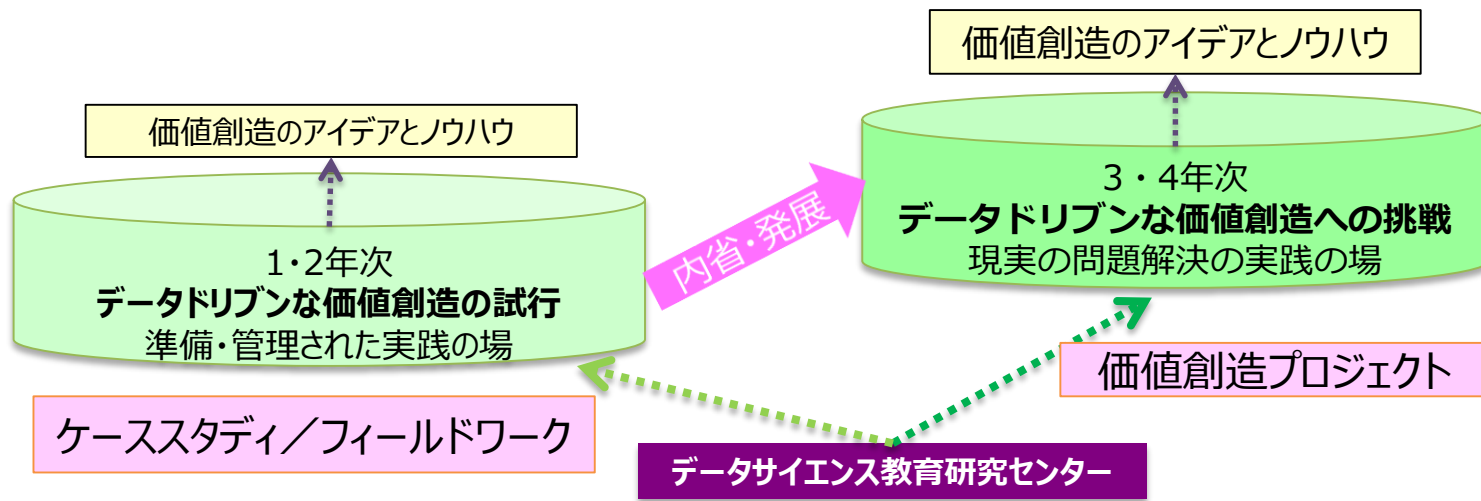
1 年	2 年	3 年	4 年
学部共通の修得科目			
〔教養教育科目群〕 数学基礎・導入 データサイエンス入門、線形代数への招待、解析学への招待、確率への招待 〔専門教育科目群〕 〔データサイエンス基礎科目〕 〈データエンジニアリング系科目〉 計算機利用基礎 データ構造とアルゴリズム データベース 情報科学概論 プログラミング I 〈データアナリシス系科目〉 解析学 統計数学 回帰分析 線形代数 統計学要論 多変量解析入門 〈データ解析科目〉 基礎情報活用演習 A, B 〔価値創造基礎科目〕 プレゼンテーション論 データサイエンス実践論 価値創造方法論 情報倫理 実践データ概論 価値創造実践論等		〔卒業要件〕 〔教養教育科目群〕 29単位以上 大学入門科目 2単位 全学共通教養科目 20単位 外国語科目 5単位 体育科目 2単位 〔専門教育科目群〕 99単位以上 データサイエンス基礎科目 〈データエンジニアリング系科目〉 11単位 〈データアナリシス系科目〉 16単位 〈データ解析科目〉 4単位 データサイエンス専門科目 自由選択 価値創造基礎科目 8単位以上 価値創造応用科目 9単位以上 データ駆動型PBL演習科目 12単位 合計 128単位以上 〔価値創造応用科目〕 各種領域分野における 講義及び演習セット	

履修モデル
に対応した
専門科目群

1 年	2 年	3 年	4 年
型別の修得科目			
◆ データエンジニア型 進路：IT系企業、シンクタンク、製薬企業、医療機関、企業のデータ分析部門、大学院等	【データサイエンス専門科目】 プログラミングⅡ、Ⅲ ビジュアルプログラミング 応用数学 プログラミング設計	情報セキュリティ 情報ネットワーク 情報理論 機械学習	最適化理論 データマイニング総論 情報学特論（人工知能） 情報活用演習等
	【データ駆動型PBL演習科目】 DS入門演習 DSフィールドワーク演習	DS実践価値創造演習	情報科学系教員ゼミ DS上級実践価値創造卒業演習
◆ データアナリスト型 進路：経営コンサルタント、シンクタンク、製薬企業、医療機関、企業の財務経理/マーケティング/生産管理、企業のデータ分析部門、大学院等	【データサイエンス専門科目】 多変量解析 テキストマイニング 質的データ解析 機械学習	時系列解析 生存時間解析 ハイズ理論 空間統計	統計学特論 情報活用演習等
	【データ駆動型PBL演習科目】 DS入門演習 DSフィールドワーク演習	DS実践価値創造演習	統計科学系教員ゼミ DS上級実践価値創造卒業演習
◆ データコンサルタント型 進路：企業の財務経理/経営企画/マーケティング/生産管理、国、地方自治体等	【データサイエンス専門科目】 社会調査法 標本調査法	品質管理 社会調査実践演習	テキストマイニング データマイニング総論
	【データ駆動型PBL演習科目】 DS入門演習 DSフィールドワーク演習	【価値創造基礎科目】 ミクロ経済学、マクロ経済学、経営学、簿記会計、財務会計、管理会計、証券分析とポートフォリオマネジメント	企業・官公庁等実務経験教員ゼミ DS実践価値創造演習 DS上級実践価値創造卒業演習

データ駆動型価値創造PBL演習

- 演習では、実際のビジネスの現場で得られるデータを扱うことにより、**実際のデータ**から価値創造を経験
- さまざまな**企業**や地方公共団体との**連携**を依頼中



IBMとの連携

- IBMによるビッグデータ解析に関する事例紹介の講義
- IBMの Bluemix 環境を用いた演習
- Watson の機能などを含めたアプリ開発を学生時代から経験

その他の企業等との連携 (協定締結済)

- 株式会社**オプトホールディング**との協定: データ分析コンテストサービスの教育への利用
- **京都銀行**と滋賀大学の地方創生に関する包括協定: ビッグデータを利用した地方創生施策の研究
- 総務省**統計研修所**との協定: 公務員等を対象とした統計教育についての連携

引き続き多くの企業と交渉中

受験：こういう高校生に受けてほしい

アドミッションポリシー

新たな文理融合の姿

- 高校の様々な教科・科目の学習を通して、バランスよく、文・理の基礎的知識を身に付けてきた、潜在豊かな人
- コミュニケーション力を有し、多様な人々と協働して、理想の未来に向けた価値創造に貢献したい人
- 物事を筋道立てて考えることができ、人間社会や自然の現象を数理的に分析することに関心のある人
- 情報ネットワーク、プログラミング、コンピュータグラフィックス（視覚化）などに関心がある人

- 数学IIBまで受けられるので、経済学部志望の学生でも受けられます
- 統計とコンピュータを社会的な課題に応用したい文系志向の人材を求めています

入学者選抜の方法

アドミッション・ポリシー

- 《本学部の教育プログラムの特徴》
- 多様な領域での価値創造を実践するデータサイエンティストの育成
 - 統計分析、コンピュータ科学、情報学から成るデータサイエンスの基礎を学び、データを活用した価値創造の理論と実践
 - マーケティング、ファイナンス、保険、会計、歴史文化、交通、防災、医療・健康、バイオ等の領域における文理融合の実践的訓練



このため、次のような人材を求めています。

- ◆ 高校の様々な教科・科目の学習を通して、バランスよく、文・理の基礎的知識を身に付けてきた、潜在性豊かな人
- ◆ コミュニケーション力を有し、多様な人々と協働して、理想の未来に向けた価値創造に貢献したい人
- ◆ 物事を筋道立てて考えることができ、人間社会や自然の現象を数理的に分析することに関心のある人
- ◆ 情報ネットワーク、プログラミング、コンピュータグラフィックス（視覚化）などに関心がある人

定員 100名

一般・前期	一般・後期	AO入試
50名	20名	30名

※初年度のAO入試は定員20名で一般型のみ
一般・前期の定員は60名

統計教育の普及を目指して高校教員も対象としたプロジェクトとして実施

一般選抜・前期

**基礎的・基本的な知識・技能を重視
《一般教科・科目型》**

科目	英語(200点)	数学(200点)
数学 出題範囲	【共通】 数学Ⅰ・数学A、数学Ⅱ・数学B(数列、ベクトル) 【選択】 ①上記共通範囲、②数学B(確率分布と統計的な推測)、③数学Ⅲのいずれかから選択	

一般選抜・後期

**課題解決に向けた思考力・表現力を重視
《総合・合科目型》**

科目	英語(300点)	総合科目(300点)
総合科目 趣旨		➢ 数学、統計、情報の内容を組み合わせた合科目の総合問題 ➢ データの解釈をもとに、問題解決的な思考力・判断力を問う ➢ 一定量の記述表現を要求する

大学入試センター試験 ⇒ 基礎的・基本的な知識・技能に関する総合評価

国語	外国語	数学	地歴	公民	理科	計
200点	200点 筆記(200点)、リスニング(50点)を200点に換算	200点	300点(100点×3科目) 〈地歴・公民から2、理科1〉又は〈理科2、地歴・公民から1〉			900点

AO入試

**課題発見・解決力、その基礎となる思考力・判断力・表現力、および学ぶ意欲を重視
《実践的な学力を総合的に評価》
(基礎的・基本的な知識・技能の一定水準の習得を前提: センター試験利用)**

評価項目	高大連携・接続型	データコンペ活用型	一般型
意欲・展望・学習計画	志願理由書	志願理由書	志願理由書
データを活用した問題解決における思考力・判断力・表現力、コミュニケーション力	本学でのデータサイエンス講座講義と演習 プレゼン、レポート	データ活用コンペ等の参加経験者 実績報告レポート	①事前エッセイ提出 ・MOOCによる課題 ・面接で口頭試問 ②小論文試験 データ分析を含む
人物	面接	面接	面接

※統計検定3級以上の合格者、独立行政法人情報処理推進機構主催基本情報技術者試験以上の合格者、財団法人全国商業高等学校協会主催情報処理検定試験各部門第1級合格者については、資格取得実績を加点する。

まとめ

- 滋賀大学のデータサイエンス学部構想は先進的な大学改革として評価されている
- 政府や企業からの期待も高い
- 多くの企業・団体と連携検討中
 - 実務家による講義
 - 学生のインターンシップ、実習
 - 共同研究（価値創造プロジェクト）
- 積極的な支援・協力をお願いしたい

ワークショップ、セッション、および資料は、IBMまたはセッション発表者によって準備され、それぞれ独自の見解を反映したものです。それらは情報提供の目的のみで提供されており、いかなる参加者に対しても法律的またはその他の指導や助言を意図したものではなく、またそのような結果を生むものでもありません。本講演資料に含まれている情報については、完全性と正確性を期するよう努力しましたが、「現状のまま」提供され、明示または暗示にかかわらずいかなる保証も伴わないものとします。本講演資料またはその他の資料の使用によって、あるいはその他の関連によって、いかなる損害が生じた場合も、IBMは責任を負わないものとします。本講演資料に含まれている内容は、IBMまたはそのサプライヤーやライセンス交付者からいかなる保証または表明を引きだすことを意図したものでも、IBMソフトウェアの使用を規定する適用ライセンス契約の条項を変更することを意図したものでもなく、またそのような結果を生むものでもありません。

本講演資料でIBM製品、プログラム、またはサービスに言及していても、IBMが営業活動を行っているすべての国でそれらが使用可能であることを暗示するものではありません。本講演資料で言及している製品リリース日付や製品機能は、市場機会またはその他の要因に基づいてIBM独自の決定権をもっていつでも変更できるものとし、いかなる方法においても将来の製品または機能が使用可能になると確約することを意図したものではありません。本講演資料に含まれている内容は、参加者が開始する活動によって特定の販売、売上高の向上、またはその他の結果が生じると述べる、または暗示することを意図したものでも、またそのような結果を生むものでもありません。パフォーマンスは、管理された環境において標準的なIBMベンチマークを使用した測定と予測に基づいています。ユーザーが経験する実際のスループットやパフォーマンスは、ユーザーのジョブ・ストリームにおけるマルチプログラミングの量、入出力構成、ストレージ構成、および処理されるワークロードなどの考慮事項を含む、数多くの要因に応じて変化します。したがって、個々のユーザーがここで述べられているものと同様の結果を得られると確約するものではありません。

記述されているすべてのお客さま事例は、それらのお客さまがどのようにIBM製品を使用したか、またそれらのお客さまが達成した結果の実例として示されたものです。実際の環境コストおよびパフォーマンス特性は、お客さまごとに異なる場合があります。