

# データアントレプレナーフェロープログラム

—2022年度講座のご紹介—

2022年2月8日

国立大学法人 電気通信大学特任教授  
データサイエンティスト協会 前事務局長  
齊藤史朗



# 1. 自己紹介

## ➤ 学問歴

- ✓ 1983年東京大学法学部を卒業：法制史のゼミにいた
- ✓ 2004年東京大学人文社会系大学院修士課程入学
- ✓ 2017年博士号取得：「日本社会学における家理論の形成と展開—その社会像と政治観」（学説史・文献学）

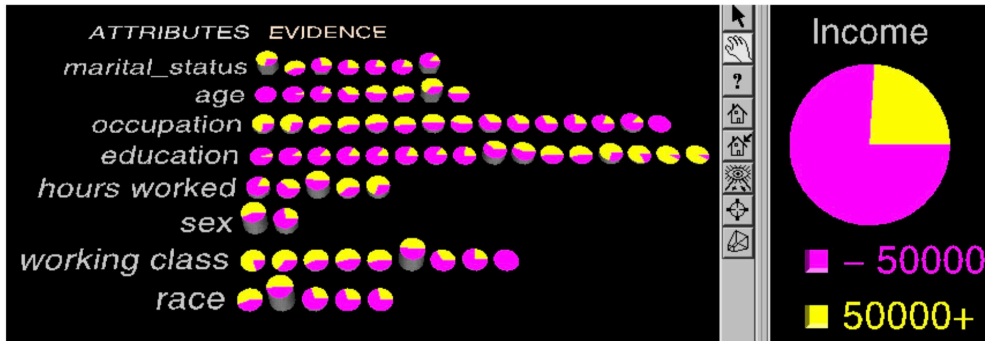
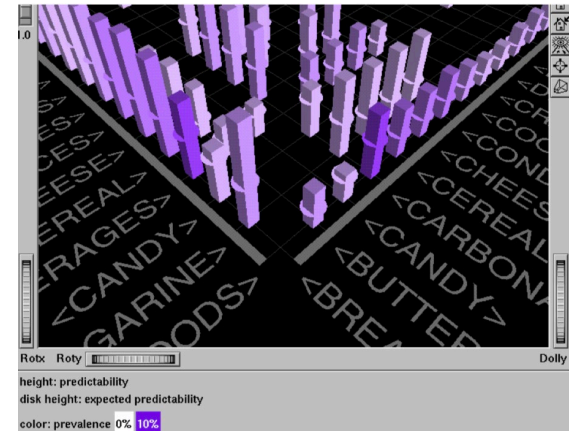
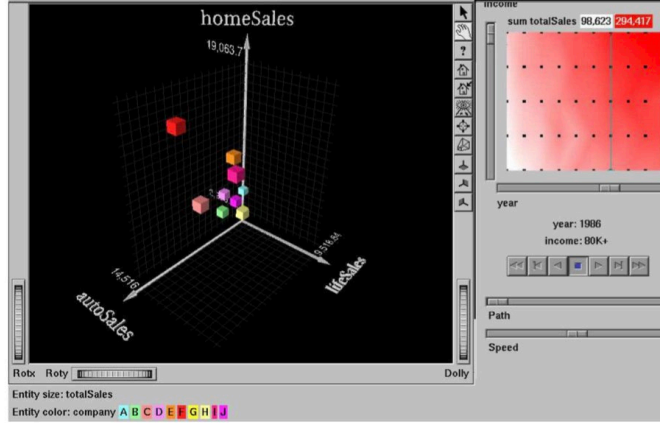
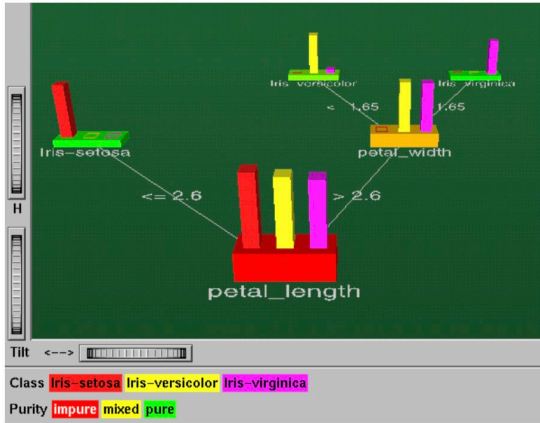
## ➤ 職歴

- ✓ 様々な仕事（舞台監督なども）を経て
- ✓ 2001年12月から株式会社金融エンジニアリング・グループ
- ✓ 2018年7月から株式会社ブレインパッド
- ✓ 2019年11月から電気通信大学特任教授

## ➤ データサイエンス歴

- ✓ 1999年にSGIのMineSetというツールに出会う
- ✓ 2000年に“Mastering Data Mining”の翻訳チームに参加
- ✓ 統計やデータマイニングは独学で学ぶ
- ✓ 2014年からデータサイエンティスト協会企画委員会の委員長、2019年から2021年までは事務局長

# 参考 : SGI(Silicon Graphics Inc.)のMineSet



# こんな本出しています。



「超」文系の本です。博士論文が元になっています。漢字が多いです。数字はほとんど出てきません。数字よりもドイツ語の方が多いくらいです。



翻訳書です。共訳なので、他の人の担当のところでも、日本語が「アレ」なところも多いです。データ分析が実際にどう使われるのかを知るのにとても良い本です。



データサイエンティストの育成や管理の本です。人事関係の参考書と言っても良い本です。お腹を抱えて笑える事例が多いです。

## 2. 講座の枠組み

## ➤ 文部科学省の「科学技術人材育成費補助事業データ人材育成プログラム」(D-DRIVE)の助成を受けたプロジェクト

### 発掘

ポテンシャルのある学生を取り込む

教育機関

・学生の募集のための大学間ネットワーク

### 育成

データサイエンティストとしての素養を育成

教育機関・企業

・教育内容  
・教育コンテンツ開発のための大学間ネットワーク

### 活躍促進

活躍の場を提供：  
インターン先から就職先まで

企業・組織

・出口（インターン・就職先）としての企業・組織とのネットワーク

# データアントレプレナーフェロープログラム

(文部科学省「データ関連人材育成拠点」)



データから価値創造 \* 多様性による発展

## データアントレプレナーコンソーシアム

- 株式会社ALBERT
- 株式会社エクサエクサウィザーズ
- 株式会社キャンパスクリエイト
- 株式会社金融エンジニアリング・グループ
- コニカミノルタ株式会社
- 株式会社データフォーシーズ
- 株式会社ネットラーニング
- 株式会社野村総合研究所
- アスクール株式会社
- 株式会社GA technologies
- ネイチャーインサイト株式会社



先進的学習スペース AIA  
(Ambient Intelligence Agora)



② 育成

## Project-based Active Learning

<基礎知識習得>

### 基礎学習

- ・eラーニング基礎科目  
(確率論・統計学)  
(コンピュータサイエンス特論)  
(プログラミング言語特論)

<能動的なコミュニケーション>

### 対面学習

- ・データアントレプレナー  
実践論(企業のソリューション事例研究)
- ・データサイエンティスト  
特論(実データを使った  
課題分析、ビジネス  
戦略討議・発表)

### 実践学習

- ・データサイエンス  
共同研究講座
- ・データ解析  
インターンシップ
- ・大学指定の  
インターンシップ



③ 活躍促進

① 発掘

多様なポテンシャル  
取込み

電気通信大学  
(代表機関)

参画機関・連携機関  
(企業、団体、大学、研究機関等)



# 産学官の連携・コンソーシアム

## ➤ 産学官の連携プロジェクト

- ✓ 「データアントレプレナー・コンソーシアム」という産学の連携組織から出発した団体が運営主体となっています。
- ✓ 2020年度から神奈川県の逗子市が連携組織として参加し、自治体との連携が開始されました。

## ➤ 産業界との連携

- ✓ 講師としての参加
  - トップレベルのデータサイエンティストとして活躍する方が講師を務めています
  - KaggleGrandmasterやKaggleMasterの講義もあります
- ✓ データ・スポンサー
  - 実務に使われている実際のデータを講義用にご提供いただいています

## ➤ 自治体との連携

- ✓ 課題やデータの提供を行っていただいています

### **3. 教育の目標： どのような人材を育てたいのか**

## ➤ 文科省のDDRIVEに最初に採択された4拠点の名称

- ✓ 早稲田大グループ : Development of Data Analysis Talents
- ✓ 大阪大学グループ : 独り立ちデータサイエンティスト人材育成プログラム
- ✓ 医科歯科大学グループ : 医療・創薬データサイエンスコンソーシアム
- ✓ 電気通信大学グループ : **データアントレプレナーフェロープログラム**

## ➤ 「データアントレプレナー」って何？

- ✓ わかりにくい、特に学生にはわかりにくい
- ✓ あえて「アントレプレナー」の語を冠した意味がある
  - わかりやすくするなら、こんな表現は不要

# データアントレプレナーとは？

- **データアントレプレナーとは電通大グループのDS教育の基軸**
  - ✓ 「データサイエンティストとしての素養を持ち、新たな価値を生むビジネスを創出できる人材」
  
- **データサイエンスを教えながらも、データサイエンスを自己目的としているわけではありません。**
  - ✓ このプログラムは、先端的なデータサイエンスの研究・開発をすることが目的ではありません。
  - ✓ あくまで目的は、「新たな価値」を創造する人材の輩出です。
  - ✓ いわばデータサイエンスはブレークスルーのための武器です。
  - ✓ それゆえ、データサイエンティストの実践力を鍛えることがプログラムの重要な目的となります。

# データサイエンスの「実践力」とは？

## ▶ データサイエンスとは「使う」ためのものである

- ✓ 「データ」には意味はない。データから意味を取り出して有用な「情報」にするのが「データサイエンス」
- ✓ 適用されない「データサイエンス」 = 机上の空論は無意味

## ▶ データサイエンスの「使い方」

- ✓ 単に技術を教えるのではなく、どういう場面で、どのように「使う」のかを教えてください。

## ▶ 何のためのデータサイエンス

- ✓ データサイエンスの「使い方」の先にある、「何のための」データサイエンスなのかを意識してもらうように努めています。

# DEFPの特長:90点取れるデータサイエンティスト

## ➤DEFPの特長って何？

- ✓ 実データを使ったPBL(Project Based Learning)は7年間の実績がある
  - しかし、多くの他のデータサイエンス学部等でも行われている
- ✓ 80点のデータサイエンティストではなく90点取れる人を育てる

## ➤データをよく見る

- ✓ 「データサイエンス」は「データ」の科学

## ➤アルゴリズムの性質を良く知る

- ✓ 単に手法の使い方を覚えて、形式的に精度を上げるにとどまらない
- ✓ なんのためにその手法を使うのかの吟味ができるようにする

## ➤実績の一つとして

- ✓ KaggleのJigsawコンペで6位入賞チーム4名のうち2名がDEFP生
- ✓ <https://www.kaggle.com/c/jigsaw-toxic-severity-rating/leaderboard?tab=private>

## 4. 講座の概要

## ➤フルオンライン

✓講義はZoomで行ない、専用のSlackを立てて議論・質疑を行います。

## ➤大学院生と社会人学生の共同のプログラムです

✓学生は社会人から、社会人は学生から、様々な学びがあります。

## ➤グループワークが主となります

✓データサイエンスの実業務は数名のグループで行うことが多いです。

- 一人でデータサイエンス業務を行うことにはさまざまなリスクがあります。

✓データサイエンスの技術を教えるだけでなく、仕事のやり方も学んでもらいます。

## ➤実際のデータ、実際の課題を使って学びます

✓実際のデータを分析し、モデルを作って実際の課題を解決することを常に意識してもらいます



# UECデータサイエンス・プログラム

## 基礎レベル

コンピュータサイエンス

Pythonプログラミング

ITリテラシ

Python100本ノック

EDAとモデル作成の基礎

## 応用レベル

統計検定2級自習

回帰モデル作成

判別モデル作成

特別講義

## エキスパートレベル

確率統計

DE実践論

特別講義

DS特論

深層学習の基礎

自然言語処理

レコメンド

因果推論

文科省

リテラシーレベル

応用基礎レベル

DS協会

見習いレベル

独り立ちレベル

棟梁レベル

到達目標

現在、データを活用して業務を行う上で必須とされるスキルを身につけた上で、データサイエンスのプロジェクトでリーダーに言われたことをきちんと実行できる。

DSプロジェクトでリーダーの指示の下、必要なことを自力で実行することができる。

DSプロジェクトでリーダーとしてデータ活用の目的をクライアントと相談して策定し、相応しい分析を行なった上で、結果を評価することができる。また、結果をシステム化するための要件を開発側に提示できる。

## ➤ DS基礎講座（総講義時間：80時間）

✓対象:社会人

✓開講期間

- ・ 2022年4月から6月(応募者多数の場合には10月から12月に再度実施します。)

## ➤ 独り立ちデータサイエンティスト養成講座（総講義時間：134時間）

✓対象:社会人

✓開講期間

- ・ 2022年4月から9月

## ➤ データアントレプレナーフェロープログラム（総講義時間：144時間）

✓対象:大学院生および社会人

✓開講期間

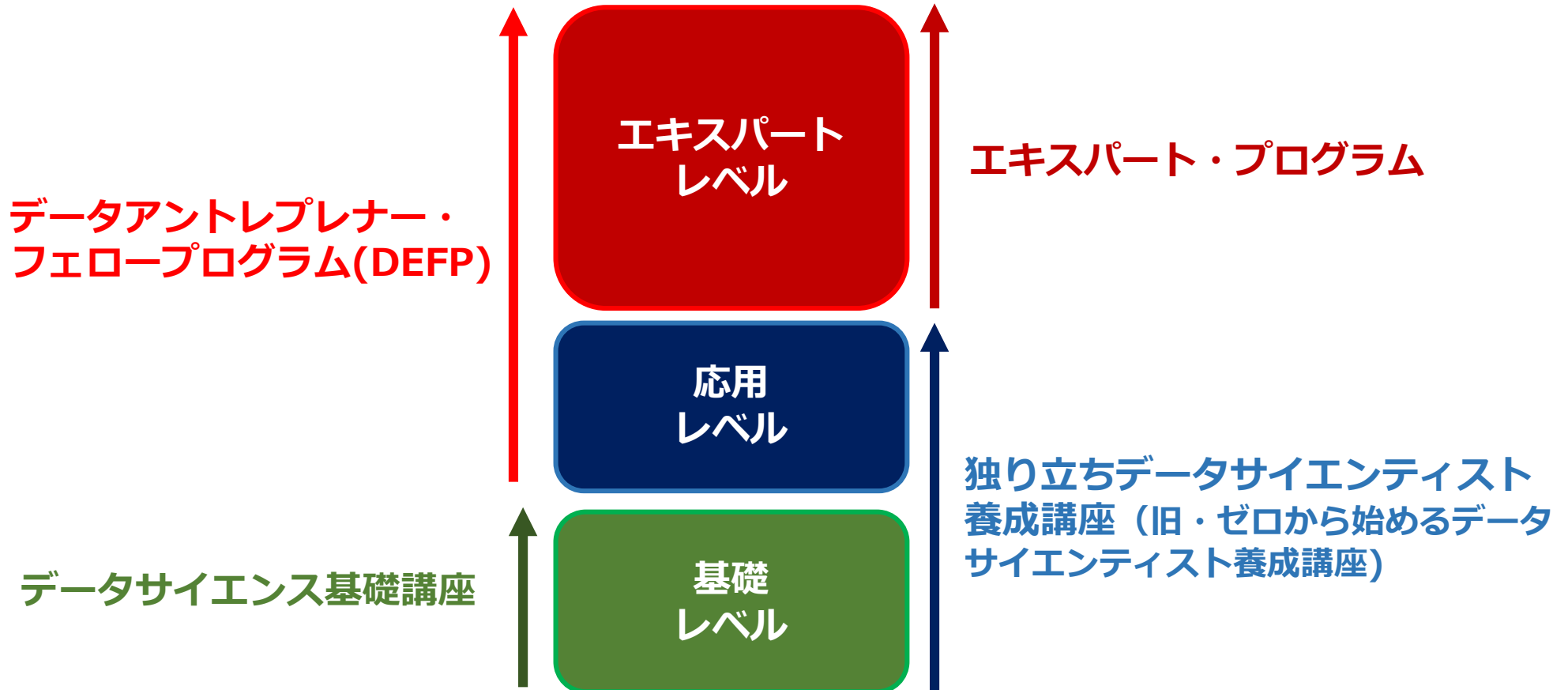
- ・ 2022年4月から2023年3月

## ➤ エキスパートプログラム（総講義時間：90時間）

✓「独り立ちデータサイエンティスト養成講座」修了生で、特に許可を得たものは追加の費用をもって、エキスパートプログラムを受講することができます。

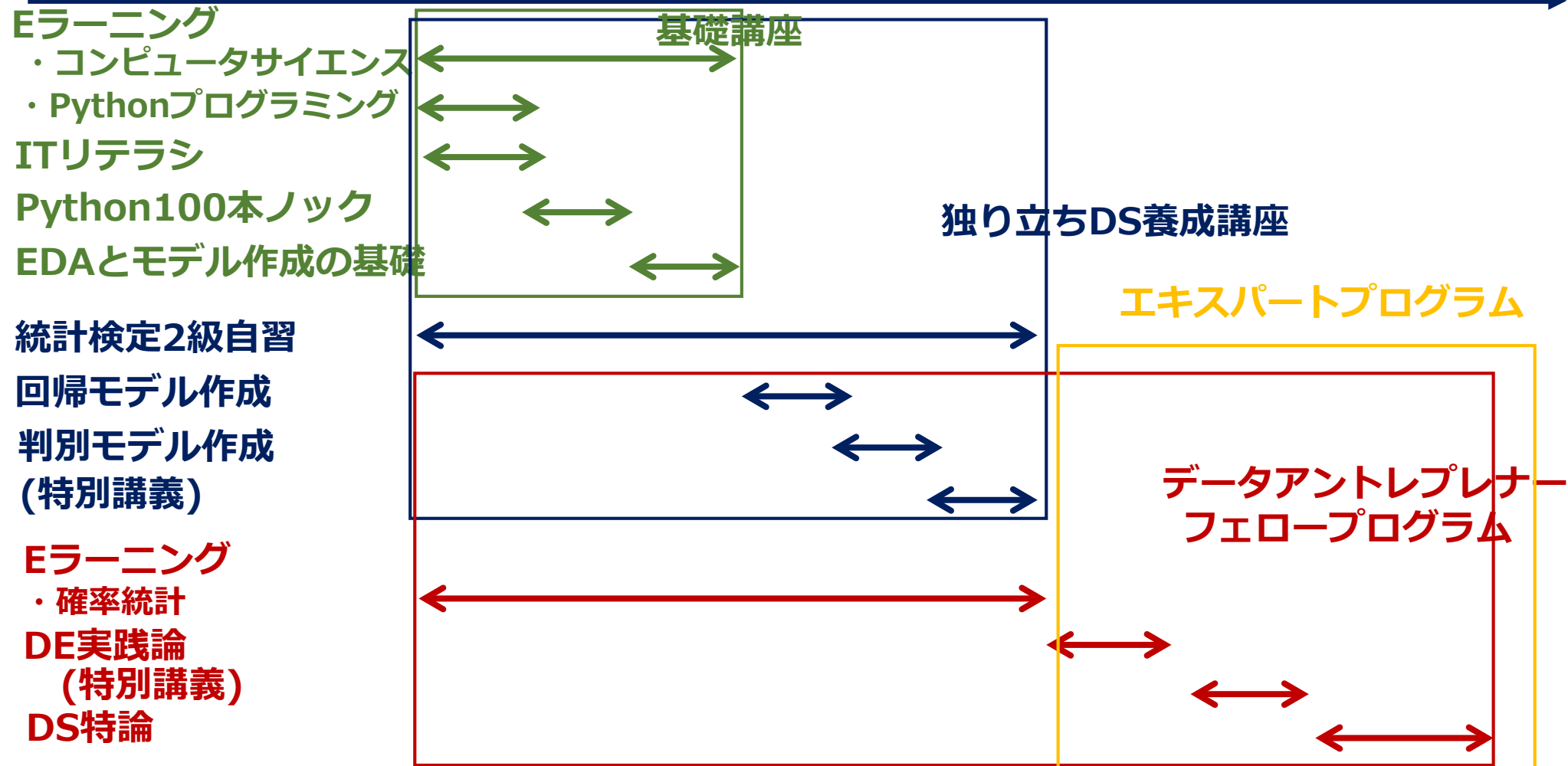
※毎週土曜日 13時から18時に講義を行います。

※講義時間の2倍から3倍の自宅学習時間が必要です。



# UEC DSプログラム2022・開講スケジュール

4月 5月 6月 7月 8月 9月 10月 11月 12月 1月 2月 3月



## ➤データアントレプレナー実践論の最終課題

- ✓ 5日目の最終日に、参加者各人が「データサイエンスを使ったサービスを考えてプレゼンテーションを行う」ピッチコンテストを実施

## ➤準備としての講義

- ✓ 5日の講義のうち最初の3日は座学（講師の話聞き、質問する）
- ✓ 座学の講師はコンソーシアム企業で、実際にデータサイエンスを活用している方々
  - 1日目：石川氏（ナイトレイ）、阪田氏（Panasonic）、赤間氏（フリー）
  - 2日目：西牧氏（IBM）、和田氏（データフォーシーズ）、島田氏（Sony）
  - 3日目：岩倉氏（富士通）、守屋氏（日立製作所）、羽間氏（エクサウィザーズ）
- ✓ 4日目は、グループワークによって、チャットボットのサービスを考えます。
  - デザイン思考によるサービス・イメージの策定
  - グループごとのプレゼンテーションを実施

## ➤ 実社会の実データを素材としたデータ分析演習

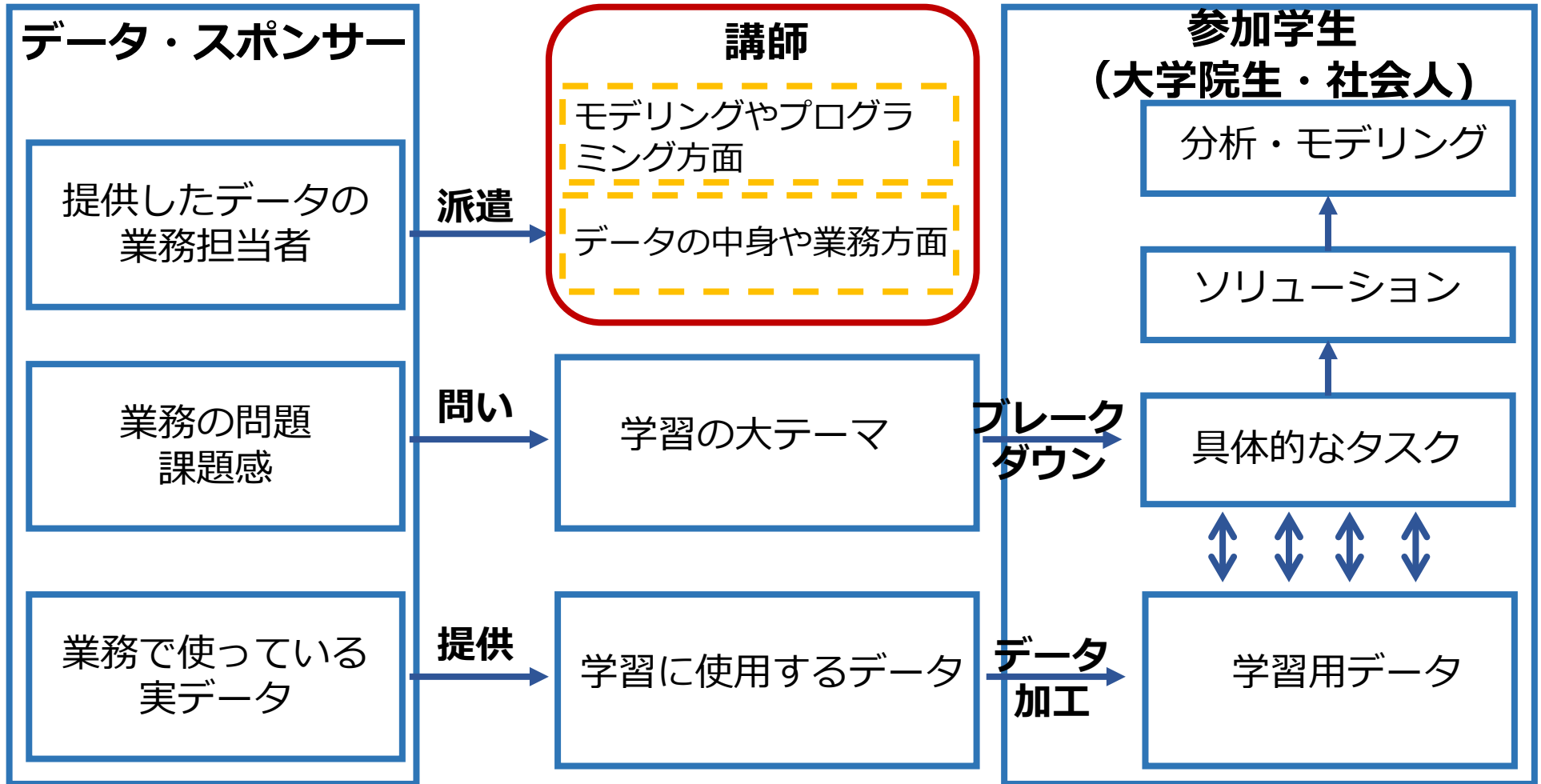
- ✓ データスポンサー様からデータとともに課題をいただく
  - 課題は少し粒度の荒いものをいただく
- ✓ 課題をブレークダウンした分析テーマを受講生（チーム）が決定
  - テーマ自体が設定されている「コンペ」とは異なる
  - データに基づかない「アイデアソン」とも異なる

## ➤ 実務視点からのフィードバック

- ✓ 参加者がテーマを設定する段階で、講師の他、データスポンサーの業務ご担当者が指導（過去の経験や、ビジネス的な筋の良し悪しなど）を行います

## ➤ データサイエンス・スキルを総合評価

- ✓ 大きな課題をデータに基づいてブレークダウンできているか
- ✓ タスクについてデータを使ったソリューションが提案できているか
- ✓ データ分析の出来の良し悪し



- **2016年**：全日食様のPOSデータ
  - ✓ 初回と最終回に分析部署担当の方が出席してご指導下さる。
- **2017年**：アスクル様の売り上げデータ（中分類まで）
  - ✓ よりLTVの高いお客様（アスクル様社内の基準による）となっていたいただくための施策を考える。
  - ✓ 毎回、事業部長様が出席してご指導下さる。
- **2018年**：ワシントンD.C.の不動産データ分析のデータ
  - ✓ 不動産価格の予測を行う
  - ✓ 毎回不動産AIの先端企業であるG.A.テクノロジーズ様が指導担当。
- **2019年**：DeNA様の携帯ゲームデータ
  - ✓ ユーザーエクスペリエンス向上のための施策を考える
  - ✓ 当該ゲームのマネージャー様が毎回出席してご指導下さる。
- **2020年**：小田急電鉄様の乗降客数データ、小田急カードのデータ
  - ✓ コロナ後の鉄道事業の戦略を考える
  - ✓ 小田急カードのご担当者が毎回出席してご指導下さる。



## 5. 受け入れ条件、受講条件や修了条件など

## ➤アドミッション・ポリシー（受け入れ条件）

- ✓統計検定2級（大学教養課程の「確率・統計」履修）レベル
- ✓なんらかのプログラミング言語を触ったことがある
- ✓なんらかのデータ分析を行ったことがある

## ➤カリキュラム・ポリシー（受講条件）

- ✓リアルタイムの講義には出席すること(7月から2月初めまで毎週土曜日午後)
- ✓予習は不要ですが、復習に講義時間の2倍から3倍の時間が必要です
- ✓4月からオンデマンド講義は受講可能なので、自信のない科目は他講座のコンテンツを使って学んでおいて下さい。
  - 確率・統計のeラーニングのレベルは統計検定準1級レベルなので早めの着手をおすすめします。

## ➤ディプロマ・ポリシー（修了条件）

- ✓9月までに確率・統計のeラーニングを修了していること
- ✓リアルタイムの講義に出席してチームに貢献していること

## 6. 費用

## ➤ DS基礎講座

- ✓ 対象:社会人
- ✓ 費用:30万円+税

## ➤ 独り立ちデータサイエンティスト養成講座

- ✓ 対象:社会人
- ✓ 費用:50万円+税

## ➤ データアントレプレナーフェロープログラム

- ✓ 対象:大学院生および社会人
- ✓ 費用:50万円+税 (社会人博士は25万円+税)
  - 学生は無料

## ➤ エキスパートプログラム

- ✓ 対象:「独り立ちデータサイエンティスト養成講座」修了生で、特に許可を得た者
- ✓ 費用:30万円+税