文部科学省
データ関連人材育成プログラム

～医療・創薬 データサイエンスコンソーシアム～
URL　http://md-dsc.com/

2017年12月

本コンソーシアムの目的
バイオ・医療情報の知識化の喫緊性

ゲノム・オミックス医療
臨床実装と怒涛の発展

大規模バイオバンクの
普及と集合的遺伝情報

IoTによる参加型医療
生涯型ヘルスケアの展開

生命情報/医療知識データ
爆発的拡大

ビッグデータ化による
革新的医療知識の発見の困難性

ビッグデータ、IoTから
超多次元ネットワーク

人工知能（Deep Learning）
によるデータ縮約原理

データサイエンス人材育成
研究コンソーシアム

＜ビッグデータ医療・AI創薬＞での国際競争力の飛躍的向上

我が国の将来の医療・健康（未来投資戦略2017） 国力の増大

設立趣意

近年のバイオテクノロジーの急速な発展に伴い、膨大な生命･医療情報が蓄積され、医療･創薬･ヘルスケア分野はいまやビッグデータ時代を迎えつつあります。これら生命･医療のビッグデータの解析には、ヒトの思考だけでは不十分で、人工知能（AI: artificial intelligence）を活用して「革新的(innovative)な知識」を発見することが期待されます。

我が国のこの分野での国際的競争力を高め、世界を主導することが今後、わが国の歩むべき発展の方途であり、医療・創薬・ヘルスケア分野におけるビッグデータ解析やAIを用いた問題解決能力の強化は喫緊の課題です。

本コンソーシアムでは、医療・創薬・ヘルスケア分野でのビッグデータ解析やAI応用、IoTなどの「新たなデータサイエンス」の方法論を開拓し、我が国のビッグデータ医療、AI創薬の分野での国際的競争力を向上させるための人材育成を通して、この分野を国際的に主導し、わが国の国力の増進と再興に寄与することを目的としています。

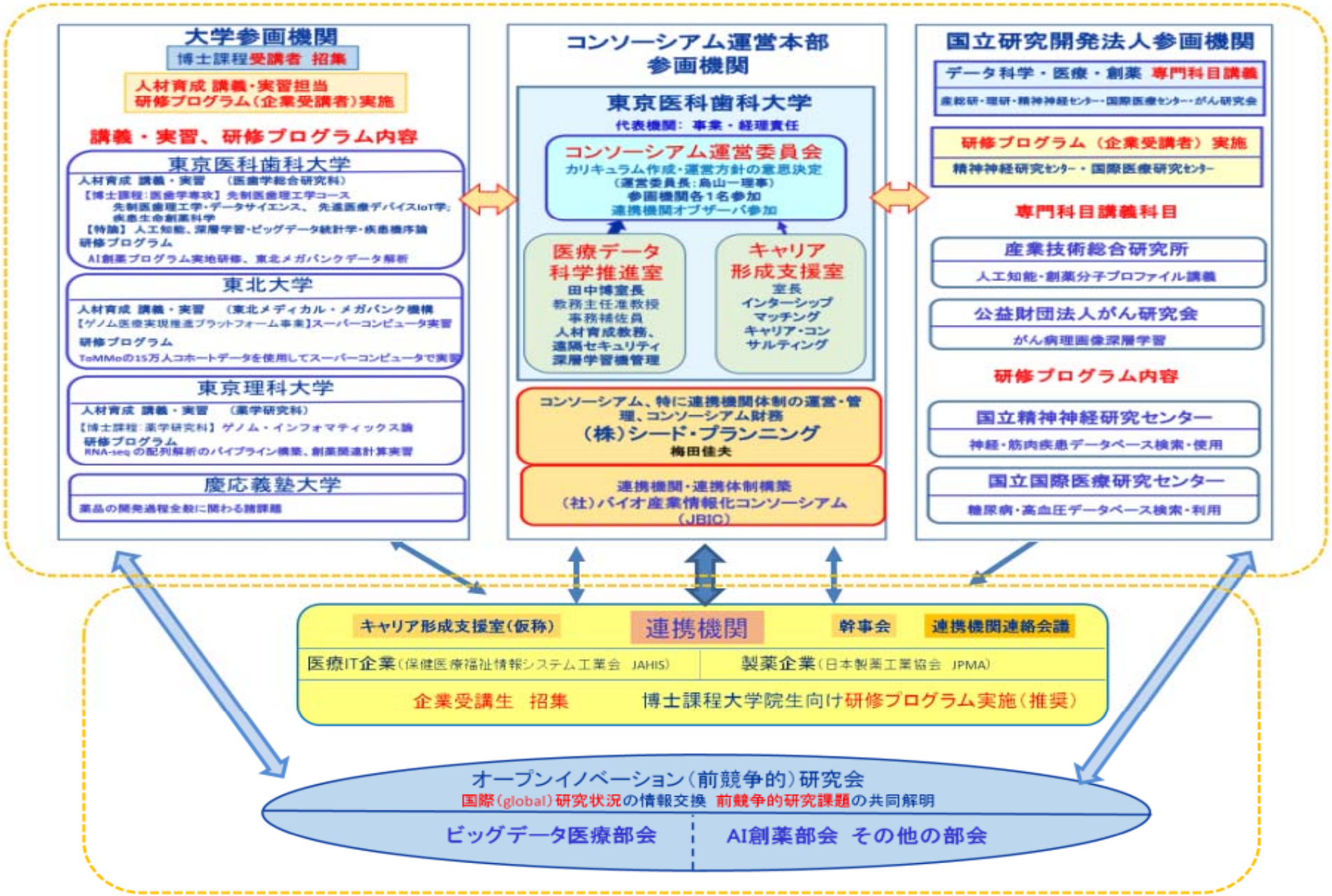
これを実行するため、①新しいデータサイエンスの教育プログラムの開発と②この分野の国際的な研究状況の情報交換を行うオープンイノベーション研究会を実施するために、本コンソーシアムを設立します。

目標

医療・創薬分野でのビッグデータ解析、人工知能（AI）応用、IoT活用において、我が国の国レベルの研究開発力を飛躍的に向上させる。

そのために、医療・創薬分野における我が国を代表する＜知のコミュニティ＞として、人材育成、Open Innovationとしての情報交換を行う研究会活動を実施するコンソーシアムを創設・運営する

体制


大学参画機関
博士課程受講者 招集
人材育成 講義・実習担当
研修プログラム（企業受講者）実施
講義・実習、研修プログラム内容
東京医科歯科大学
人材育成 講義・実習 （医歯学総合研究科）
【博士課程：医歯学専攻】先制医歯理工学コース
先制医歯理工学・データサイエンス、 先進医療デバイスIoT学、
疾患生命創薬科学
【特論】 人工知能、深層学習・ビッグデータ統計学・疾患機序論
研修プログラム
AI創薬プログラム実地研修、東北メガバンクデータ解析
東北大学
人材育成 講義・実習 （東北メディカル・メガバンク機構
【ゲノム医療実現推進プラットフォーム事業】スーパーコンピュータ実習
研修プログラム
ToMMoの15万人コホートデータを使用してスーパーコンピュータで実習
東京理科大学
人材育成 講義・実習 （薬学研究科）
【博士課程：薬学研究科】ゲノム・インフォマティックス論
研修プログラム
RNA-seq の配列解析のパイプライン構築、創薬関連計算実習
慶応義塾大学
薬品の開発過程全般に関わる諸課題
コンソーシアム運営本部
参画機関
東京医科歯科大学
代表機関： 事業・経理責任
コンソーシアム運営委員会
カリキュラム作成・運営方針の意思決定
（運営委員長：鳥山一理事）
参画機関各1名参加
連携機関オブザーバ参加
医療データ
科学推進室
田中博室長
教務主任准教授
事務補佐員
人材育成教務、
遠隔セキュリティ
深層学習機管理
キャリア
形成支援室
室長
インターシップ
マッチング
キャリア・コン
サルティング
コンソーシアム、特に連携機関体制の運営・管理、コンソーシアム財務
（株）シード・プランニング
梅田佳夫
連携機関・連携体制構築
（社）バイオ産業情報化コンソーシアム
（JBIC）
国立研究開発法人参画機関
データ科学・医療・創薬 専門科目講義
産総研・理研・精神神経センター・国際医療センター・がん研究会
研修プログラム（企業受講者）実施
精神神経研究センター・国際医療研究センター
専門科目講義科目
産業技術総合研究所
人工知能・創薬分子プロファイル講義
公益財団法人がん研究会
がん病理画像深層学習
研修プログラム内容
国立精神神経研究センター
神経・筋肉疾患データベース検索・使用
国立国際医療研究センター
糖尿病・高血圧データベース検索・利用
キャリア形成支援室（仮称）
連携機関
幹事会
連携機関連絡会議
医療IT企業（保健医療福祉情報システム工業会 JAHIS）
製薬企業（日本製薬工業協会 JPMA）
企業受講生 招集
博士課程大学院生向け研修プログラム実施（推奨）
オープンイノベーション（前競争的）研究会
国際（global）研究状況の情報交換 前競争的研究課題の共同解明
ビッグデータ医療部会
AI創薬部会 その他の部会

ビッグデータ医療・AI創薬コンソーシアム活動

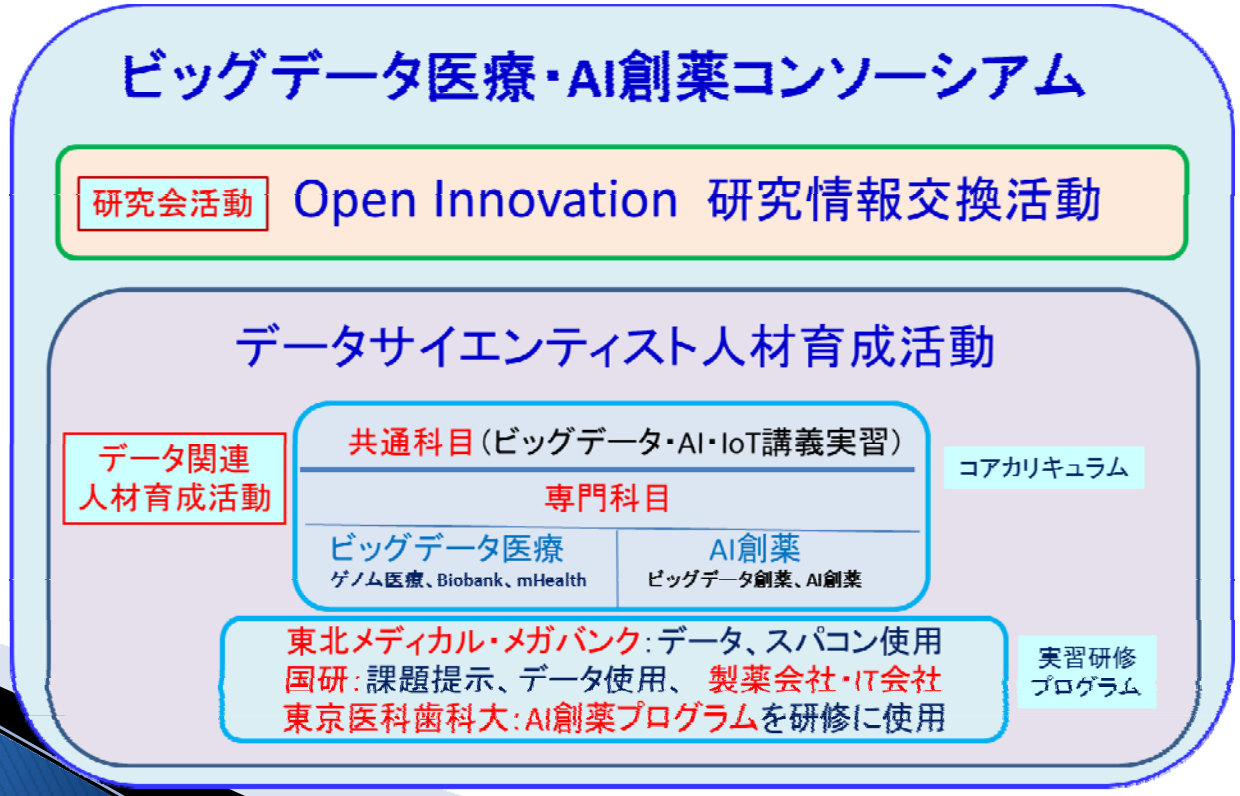
- 受講者　約60名（6000万円助成の場合）
  - A. 博士課程学生・博士学位取得者（企業もOK）**30名以上**
  - B.　企業からの受講者（学位なしOK）**30名程度**　A ≧　B
- 人材育成カリキュラムの概要

コア・カリキュラム　**講義・実習（2カ月～半年）**

- **共通科目（データサイエンス諸手法の講義・実習）**
- **専門科目（ビッグデータ医療コース、AI創薬コース）**

研修プログラム（2～3カ月）

- 東北メディカル・メガバンク機構での研修（データ、スパコン使用）
- 他の国研・がん研での研修（各研究所の医療データベース利用）
- 製薬会社・医療IT会社での研修



体制

◆コンソーシアムの体制は、東京医科歯科大学を代表機関として、この下に活動方針を決定する運営委員会を設置。

◆人材育成プログラムは、コアカリキュラムコースと研修コースから成る。

◆博士課程学生は、参画機関の大学、研究機関に限らず、広く一般の大学からも招集する。

◆参画機関の大学は講義や実習を受け持つ、医療・創薬の専門科目や現実の医療・創薬の場による研修プログラムを担当する研究所から成る。

◆連携機関としては、医療IT企業、製薬企業、医療機器企業、CRO等、医療分野におけるAIに関心がある民間企業を35社から40社程度が、本プログラムを受講すると同時に、博士課程学生のキャリア形成支援のためのインターンシップを実施する。

◆連携機関は、企業内受講希望者を招集するとともに、インターンシッププログラムとして、受講生を受け入れを検討する。連携機関連絡会議にて、プログラムのニーズやアドバイスを行う。

◆運営委員会の下、医療データ科学推進室を設置。人材育成プログラムの開発行う。同時に、キャリア形成支援室を設置。博士課程学生のキャリアコンサルティングや連携機関におけるインターンシッププログラムのマッチングを行う。

◆オープンイノベーション（前競争的）研究会ではビッグデータ医療部会、AI創薬部会、他、課題に応じて部会を開設し、産学官の情報交流や国際的な研究状況の情報交換を推進するとともに、前競争的研究分野の課題解決に向けて協議する。

本事業のロードマップ

	1年目	2年目	3年目	4年目	5年目
受講者数（学生数）	40（20）	60（30）	60（30）	60（30）	60（30）
インターンシップ参加者数（学生数）	30（20）	50（25）	50（25）	20（10）	20（10）
就職者（学生数）	5（5）	10（10）	10（10）	10（10）	10（10）
取り組み内容	参画機関、連携機関リクルート	コンソーシアムの本格稼働	他の全国組織との連合を通したALL JAPAN体制の構築	国家レベルのビッグデータ医療、AI創薬データ科学連合	
	初期カリキュラムの設計	本格カリキュラムの設計	研究成果を取り入れたカリキュラム改変	国レベルのデータ科学者育成プログラム確立普及	
		キャリ化開発を目指した研修プログラムの検討		本人材育成プログラム修了者のキャリア形成、就職支援	

データサイエンスという概念

- **英語：Data science, data-driven science**（データ駆動型科学⇔仮説駆動）
- concept to unify statistics, data analysis and their related methods to understand and analyze actual phenomena with data
  - mathematics, statistics, information science, and computer science, in particular from the subdomains of machine learning, classification, cluster analysis, data mining, databases, and visualization.
- 言語使用の歴史
  - 1960,1974 Peter Naurの教科書、sub field of computer science
  - 1996, “Data Science, classification, and related methods” in International Federation of Classification Societies
  - 1997, Jeff Wu “Statistics = Data Science?“ Michigan 大学就任演説
  - 2001, William S. Cleveland introduced data science as an independent discipline, extending the field of statistics to incorporate "advances in computing with data" in International Statistical Review.
  - 2002, the International Council for Science (ICSU): Committee on Data for Science and Technology (CODATA) started the *Data Science Journal*
  - 2003, Columbia University began publishing *The Journal of Data Science*, which provided a platform for all data workers to present their views and exchange ideas
  - 2005, The National Science Board published "Long-lived Digital Data Collections: Enabling Research and Education in the 21st Century" defining data scientists as "the information and computer scientists, database and software and programmers, disciplinary experts, curators and expert annotators, librarians, archivists, and others, who are crucial to the successful management of a digital data collection" whose primary activity is to "conduct creative inquiry and analysis."
  - 2012 Harvard Business Review article "Data Scientist: The Sexiest Job of the 21st Century",
  - a data scientist is "a new breed", and that a "shortage of data scientists is becoming a serious constraint in some sectors“.
  - 2013, the IEEE Task Force on Data Science and Advanced Analytics, European Association for Data Science (EuADS)
  - 2014, the American Statistical Association renamed its journal to "Statistical Analysis and Data Mining: The ASA Data Science Journal" and in 2016 changed its section name to "Statistical Learning and Data Science".

データサイエンスという概念2

- 統計学との関係
  - at the Meetings of American Statistical Association, Nate Silver said, “I think data-scientist is a sexed up term for a statistician....Statistics is a branch of science. Data scientist is slightly redundant in some way and people shouldn’t berate the term statistician.”
  - data scientists as only one of the four greater job families companies require to leverage big data effectively, namely: business analysts, data scientists, big data developers and Big Data engineers.
  - He argues data science, like any other interdisciplinary field, employs methodologies and practices from across the academia and industry, but then it will morph them into a new discipline. He brings to attention the sharp criticisms computer science, now a well respected academic discipline, had to once face.
  - NYU Stern's Vasant Dhar, as do many other academic proponents of data science,argues more specifically in December 2013 that data science is different from the existing practice of data analysis across all disciplines, which focuses only on explaining data sets. Data science seeks actionable and consistent pattern for predictive uses.

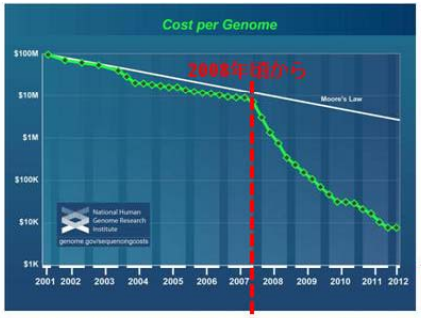
The future of data science not only exceeds the boundary of statistical theories in scale and methodology, but data science will revolutionize current academia and research paradigms. As Donoho concludes, "the scope and impact of data science will continue to expand enormously in coming decades as scientific.“

結論：データサイエンスとは、ビッグデータ時代のデータ駆動型科学であり、それはこれまでの情報科学を拡張するだけでなく、現在の科学界やサイエンスそのものの土台からの変革（パラダイム革命）をもたらすものである。

ゲノム医療の2つの流れ

- 米国の流れ
  - 次世代シーケンサの急激な発展による「シーケンス革命」からの怒涛の展開（2010から）
  - 「治療医学」レベル質的向上のためにゲノム情報を取り入れた臨床実装の推進
    - 稀少疾患の原因遺伝子変異の同定
    - がんのドライバー遺伝子変異の同定と分子標的薬の選択
    - 薬剤代謝酵素の多型性の同定と個別化投与
- 欧州の流れ
  - 社会福祉国家の理念より国民医療（医療の国民レベル）の向上
  - 「予防医学」レベル質的向上のためにゲノム情報を取り入れたバイオバンク推進
  - 大規模前向きpopulation型バイオバンク/ゲノム・コホートの確立
    - 遺伝的素因だけでなく環境要因（生活習慣）との相互作用を解明し、「ありふれた疾患」発症を予測し、これに基づいて個別化予防する。
    - 疾患を発症前に対応して発症を防ぐ「先制医療(preemptive medicine)」や「予測医療（predictive medicine）の実現を目的

米国ゲノム医療の流れ



DNA Sequencing Cost: the National Human Genome Research Institute

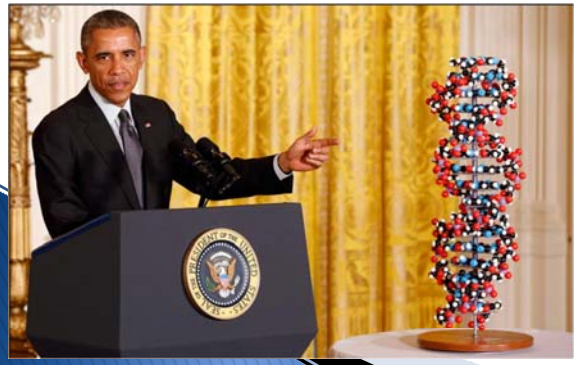
チドル
ゲノム

シーケンス革命　2007/8

2005～ NGS 454 (LS,Roche)
2007/8～454, Solexa (Ilumina),
SOLiD (LT,TF)
シーケンス革命

	HiSeq2500		Ion Proton
本体価格	約1億円		約3500万円
モード / チップ	ハイアウトプット	ラピッドラン	Ion Proton I
解析時間	11日	27時間	2時間
リード長（bp）	2 x 100	2 x 150	200
データ産出量(Gb)	約600	約120	10
試薬コスト（ヒト1人全ゲノム）	数十万円		不可 エクソームのみ

急速な高速化と廉価化
ヒトゲノム解読計画13年,3500億円
⇒1日,10万円



先陣争いの時代

第一期

ゲノム多型性の認識
Hapmap 計画(2002)
GWAS研究など

薬剤代謝酵素の多型性の判別・電子カルテで警告・Preemptive PGx
Vanderbilt大病院

2005～ NGS 登場
(454,Solexa,SOLID)
2007/8～
シーケンス革命

Undiagnosed Genetic Diseaseの原因遺伝子POC同定
MCW小児病院

国際がんコンソーシアム開始
ICCG (2008年)
2011頃からがん変異成果報告

Cancer Driver Geneの同定と抗がん剤治験
Dana Faber CC

国家政策の時代

第二期

ゲノム・オミックス医療の臨床実装の普及
ゲノム・オミックス情報のビッグデータの出現

ゲノム医療の国家的取組み
NIH "BD2K "計画 ・各種ゲノムコンソーシアム開始

オバマ大統領　年頭教書
Precision Medicine initiative 政策の発表

精密医療普及期

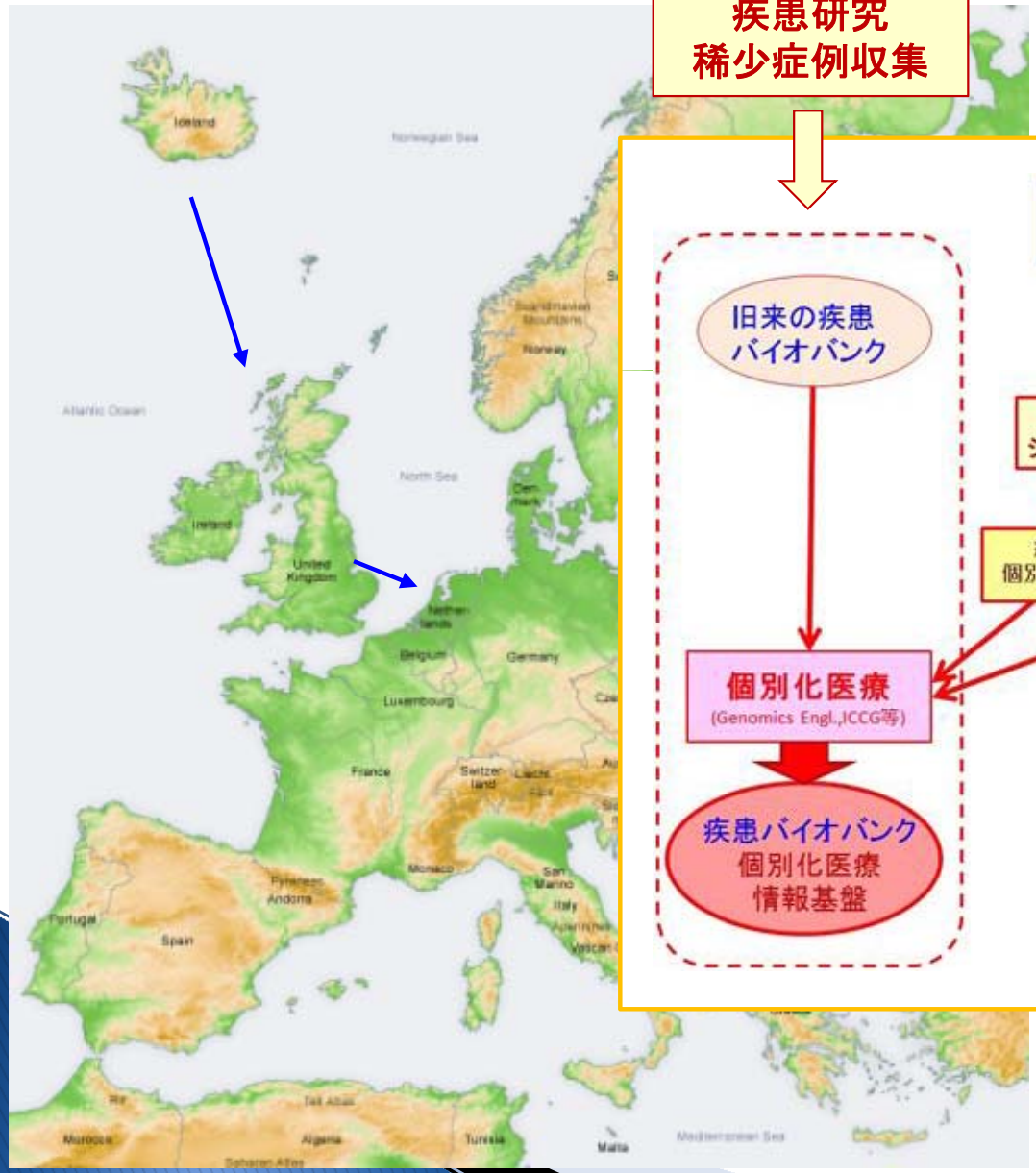
第三期

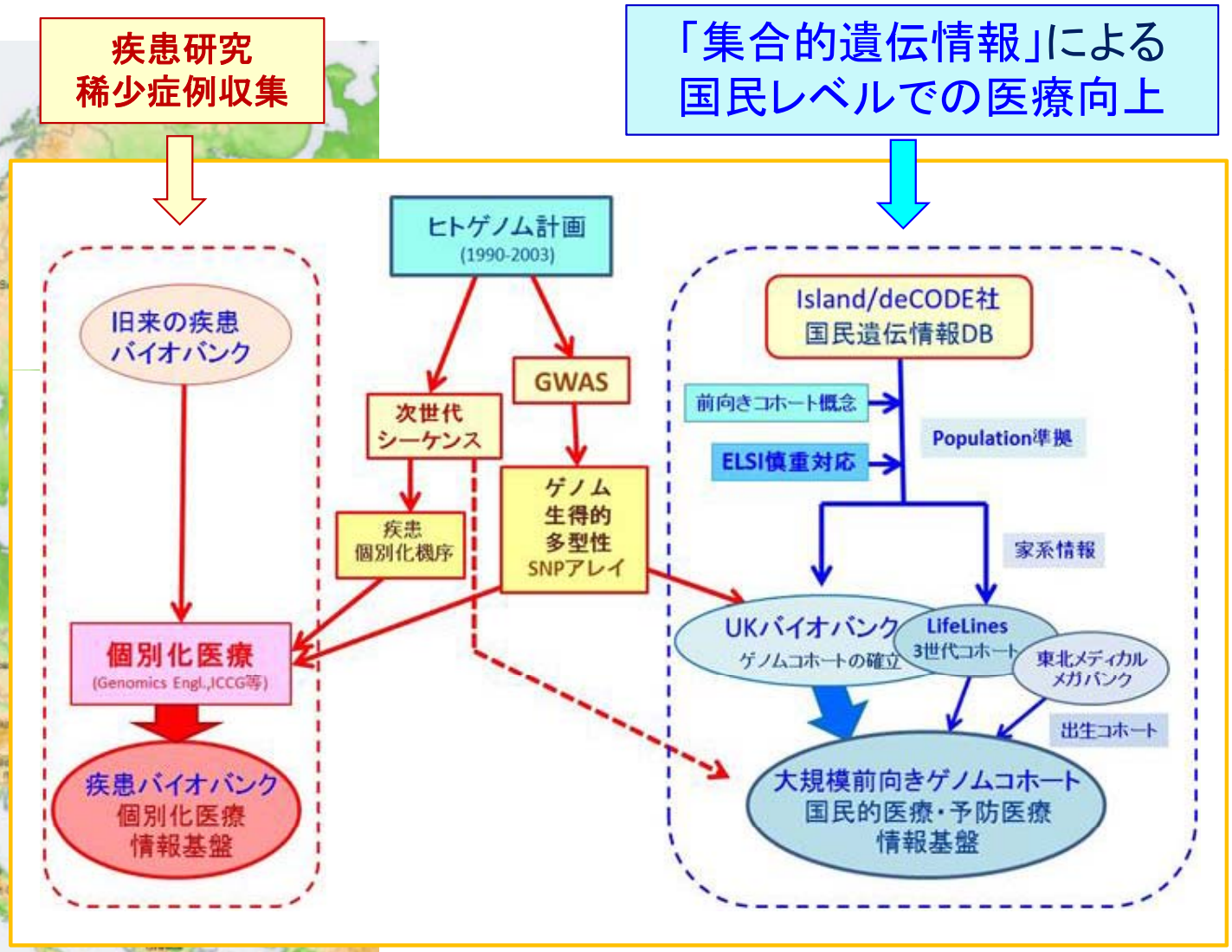
100万人コホート:バンダービルト大学設開始(2016-2020)
NCI"National Cancer MoonShot" 10年計画開始
各州でのプレシジョン医療計画開始(カリフォルニア、ペンシルバニア)

2007年
2009年
2010年
2011年
2012年
2013年
2014年
2015年
2016年
2017年

オバマ大統領　Precision Medicine Initiativeを開始
2015年1月　大統領一般年頭教書演説

第2の流れ　欧州のバイオバンクの普及





医療ビッグデータ時代

(1) 次世代シーケンサ（Clinical Sequencing)による
「ゲノム/オミックス医療」における網羅的分子情報収集/蓄積
(2) Biobank/ゲノムコホート普及による分子・環境情報の蓄積
(3) モバイルヘルス(mHealth)によるWearable センサの連続　計測による生理データの蓄積（unobstructed monitoring)

急激な大量データの出現
コストレス化かつ高精度化

!!

ゲノム：13年→1日(1/5000) 3500億→10万円(1/350万)

個別化医療・医療の国民レベルの向上
医療/ヘルスケアの適確性の飛躍的な増大

NIH BD2K計画の実施

- 医療におけるデータ科学の全米COE創成
  - Center of Excellence (COE) in Data Science
    - Uni. Pitts: Center for causal modeling and discovery of biomedical knowledge from big data
    - UCSC: Center for big data in translational genomics
    - Harvard: Patient-centered information commons
    - **その他、コロンビア大学、イリノイ大学など11施設　32M$**
- Data Scientist 人材養成
- データ発見索引 DDI (Data Discovery Index) Consortium
  - Data discovery index coordination consortium (DDICC)
  - データベースカタログの発展・PubMEDのDB版
  - UCSD: BioCADDIEを中心にDDI研究開発

NIH「ビッグデータから知識へ」計画 *"Big Data to Knowledge" (BD2K) initiative*

- **以前の同様な国家プロジェクト** 1999
  - "生命医療情報科学技術計画(BISTI)"
  - "Biomedical Information Science and Technology Initiative"
- BD2K: "Big Data to Knowledge" Initiative 開始 2013
  - 「データ・情報学に関するNIH長官諮問委員会」WG
    - WG on Data and Informatics for Advisory Committee to the Director (ACD) of NIH
  - 焦点化したワークショップ, 知識発見指標や人材養成のためのCOEに対する提案の募集
  - 研究費の配分　2014年から
  - **データ科学のための副長官**（Associate Director of Data Sciences）を任命 **Boune, PhD.**
- Francis Collins長官談:「NIH全規模での優先計画」
  - 生命医療研究に喫緊の重要性を持つ、指数的に増大する生命医療データを活用する
  - 「ビッグデータの時代は到来した」
  - NIHが、この革命を作り上げる、**様々な異なったデータ種類に対するアクセスの統合・分析に主導的な役割を果たす。**
- http://bd2k.nih.gov

医療データサイエンスの定義

- ビッグデータ医療時代のデータ駆動型科学であり、
- 「医療ビッグデータから知識を取り出す」
  方法論の科学である。
- 医療ビッグデータとは
  - ゲノム・オミックス医療
  - 大規模バイオバンク
  - 大規模生命情報データベース
  - IoT、モバイルヘルス
  - 医用機能画像、医用分子機能画像
  - その他

データサイエンスの領域

- 統計学(statistics)：
  - X古典的統計学　単変量検定推定
  - 探索的データ統計学（explanatory)
  - 多変量解析・データマイニング手法
- 機械学習・知識発見(machine learning, KDD):
  - 探索的統計学と違う「立て付け」で非線形問題へ
  - ほぼ同じ課題を数理的手法によって解決
- 人工知能（AI：Artificial Inteligence)
  - 機械学習と被る。ニューロコンピュータは別領域として認識。

共通の認識
ビッグデータから知識を取り出す

医療の「ビッグデータ革命」
～何が新しいのか～

1）臨床診療情報
- 従来型の医療情報
  - 臨床検査、医用画像、処方、レセプトなど

2）社会医学情報
- 従来型の社会医学情報
  - 疫学情報・集団単位での疾患罹患情報

旧来のタイプの
医療データの
大容量化

3）新しい種類の医療ビッグデータ
- 網羅的分子情報・個別化医療
  - ゲノム・オミックス医療
  - システム分子医学・Precision Medicine
- 生涯型モバイル健康管理（mHealth）
  - ウェアラブル・生体センシング

新しいタイプの
医療ビッグデータ

医療の「ビッグデータ革命」
～何が新しいのか～

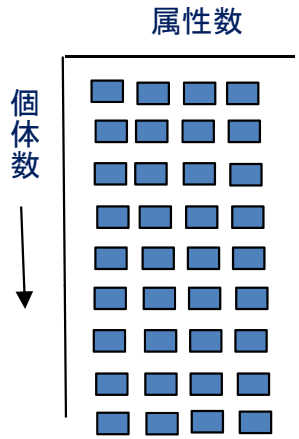
＜目的もデータ特性も従来型と違う＞

従来の生命医療情報の「ビッグデータ」
Big “Small Data” (n≫p) n-bigdata

医療情報・疫学調査 属性数：10項目程度

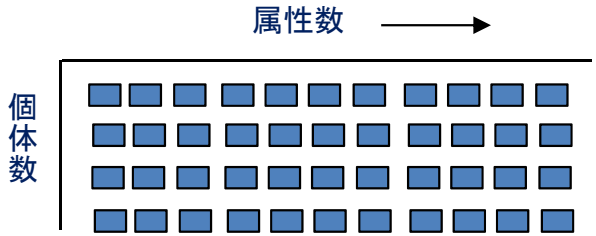
– 目的：Population MedicineのBig Data

⇒個別を集めて「集合的法則」を見る



網羅的分子情報,、mHealth、大規模生体ビッグデータ
Small “Big Data”(p≫n) p-bigdata

１個体に関するデータ属性種類数が膨大
属性に比べて個体数 少数:従来の統計学が無効
とくに多変量解析:GWASで単変量解析の羅列



– 目的：例えば医療の場合Personalized Medicine

⇒大量データを集めて「個別化パターン」の多様性を摘出する

新しいデータ科学の必要性

ビッグデータ時代の統計手法の限界

- 重回帰モデル

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \varepsilon\,.$$

- 観測

$$Y = \mathbf{X}\beta + \varepsilon.$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

- パラメータ推定

$$\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 \to \min$$

$$\widehat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\vec{y}$$

pービッグデータではデータ行列(相関行列)は次元落ちになり求まらない

医療の「ビッグデータ」革命は どんな既存のパラダイムに挑戦しているか

- Population medicineのパラダイム転換
  - <One size fits for all>のPopulation医療はもはや成り立たない
  - Personalized (Precision) medicine
    - 個別化医療を実現するために＜個別化・層別化パターン＞を網羅的に調べる:どこまでの粒度で個別化・層別化すればよいか
- Clinical research（臨床研究）のパラダイム転換
  - 臨床研究を科学にする従来の範型RCTは個別化に破綻した
  - <statistical evidence based>呪縛からの解放
  - 「標本」統計・「推測」統計学に限定されない臨床研究
  - Real World Data:ビッグデータ知識生成（BD2K)
- 創薬の戦略パラダイムの転換
  - ビッグデータ創薬の可能性
  - 創薬・育薬のReal World Dataの利用
  - Transdisease Omics, Drug networkのDual Network Topologyによる創薬

カリキュラム（2017）

今年度はトライアルにつき、自由選択

20171211時点

1.データサイエンス関係（共通科目）

コース名	講座名
システム医学概論 （システム医学）	システム医学概論
	脳機能画像による脳内・認知機能の評価とシステム医学
	システム医学概論
	ファンクショナルゲノミックスによる創薬ターゲットの探索と再生医療ツールの同定
	疾患の機序解明に向けたシステム医学的アプローチ
システム医学概論（システム薬理学）	医学・生物学研究におけるデータ駆動型アプローチ
	分子ネットワークにおける数理モデルの構築と応用
	遺伝子間制御ネットワークの推定アルゴリズムとそのシステム医学への応用
人工知能・機械学習概論（理論）	教師付き学習
	教師なし学習
	強化学習
人工知能・機械学習概論（医学的応用）	概論、機械学習による画像診断支援
	オミクス解析における機械学習の利用
	（特別講義）調整中
	機械学習を用いた音響／センサデータからの異常検知
	介護予防、医療活用のためのウェアラブルの信号処理と知識化技術
ゲノム統計学・ビッグデータ統計学	ゲノム・ビッグデータ統計解析学

2.データサイエンス関係（共通科目　実習）

コース名	講座名
	Rプログラミング演習（R基礎演習）
	Rプログラミング演習（Rによる機械学習）
	パイソンプログラミング演習
	深層学習総合演習

カリキュラム（2017）

今年度はトライアルにつき、自由選択

20171211時点

4. AI創薬関係（専門科目）

コース名	講座名
ビッグデータ創薬・AI創薬特論	医療・創薬におけるビッグデータAIの可能性
	機械学習とデータ駆動型ドラッグリポジショニングによる創薬
	ビッグデータ創薬・AI創薬特論
	たんぱく質相互作用ネットワーク解析とDeep Learninnguを利用した、新規薬剤候補の探索について
分子創薬プロファイリング総論	溶液NMRを中心とした創薬研究
	オミックスデータの創薬活用
	創薬支援ソフトウェアmyPrestoでの薬物探索・ドッキングの計算理論
	創薬支援ソフトウェアmyPrestoでの薬物探索・ドッキングの計算演習
	創薬支援に向けた情報統合
インシリコ創薬化学	化合物情報に基づく創薬インフォマティクス タンパク質構造情報に基づく創薬インフォマティクス
	生体分子動力学シミュレーションの基礎 分子動力学シミュレーションの創薬応用
	生体高分子の量子化学計算 フラグメント分子軌道法の創薬への応用
	インフォマティクスとシミュレーションを融合したインシリコスクリーニング 創薬におけるインシリコ薬物動態、毒性予測
現代薬理統計学	医療系データベースの利活用と社会福祉、臨床研究

カリキュラム（2017）

今年度はトライアルにつき、自由選択

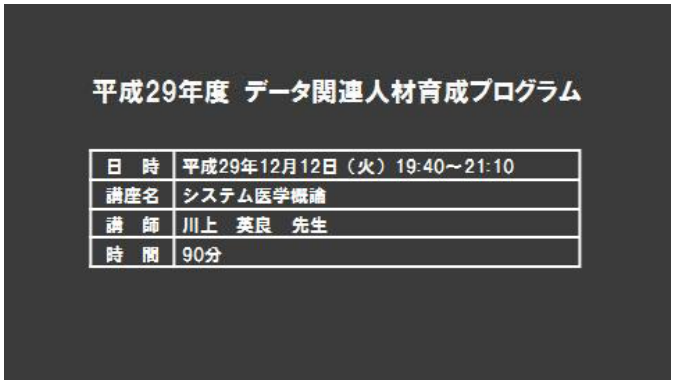
20171211時点

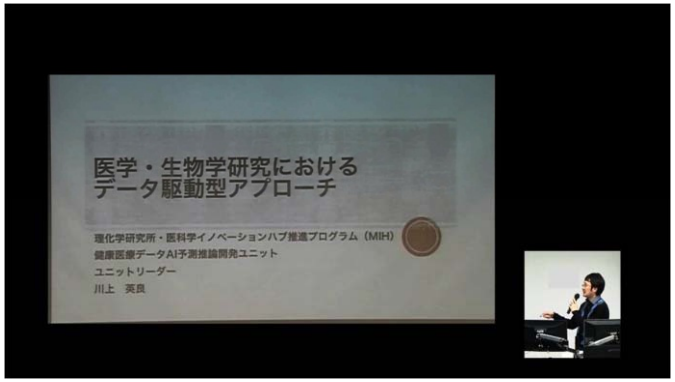
3. ビッグデータ医療関係（専門科目）

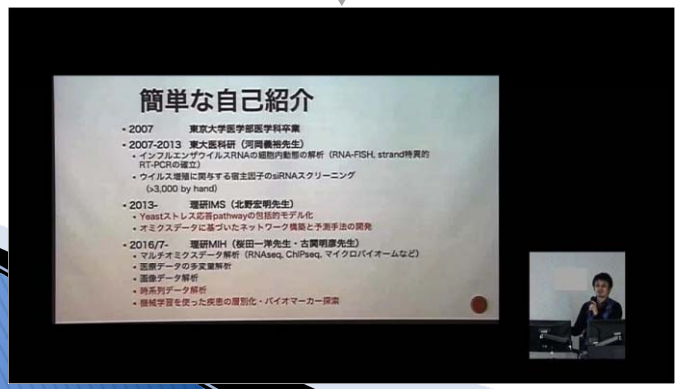
コース名	講座名
Precision Medicine概論	Precision Medicine概論
Biobank/Genome Cohort概論	Biobank/Genome Cohort概論
	東北メディカル・メガバンクにおけるコホート調査データ管理
	調整中
	ヨーロッパを中心としたバイオバンク
	（特別講義）
生命医学ビッグデータベース論	国外の疾患ゲノムバリエーションデータベース
	がんのデータベース
	臨床ゲノム情報統合データベース事業の取り組み
	調整中
疾患データ科学	疾患ゲノムコホート・バイオバンク：NCBN
	心血管代謝性疾患のデータ科学
	がん関連データベース概論 がんのゲノム研究におけるがん関連データベースの利活用の実際 クリニカルシーケンシングにおけるがん関連データベースの利活用の実際
	神経疾患（総論）、遺伝性小脳失調症
	多因子疾患のデータ科学
	未診断疾患、IRUD　beyond ,神経疾患（各論）
	稀少疾患、国際協調、筋ジストロフィー

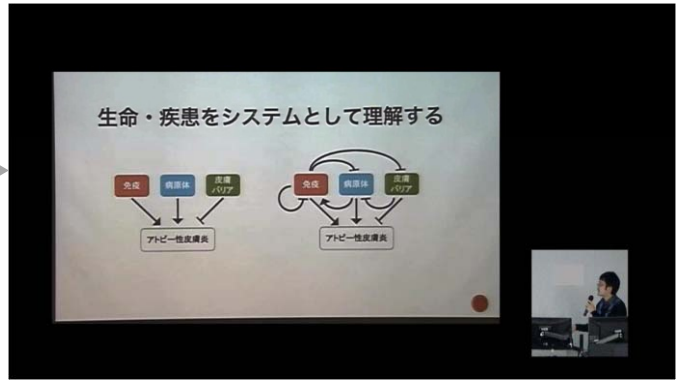
e-Learning（2017）

配信準備中（12月講義は1月配信予定）









本編：
90分前後

本編（つづき）：
90分前後

エンドロール：
10秒

人材育成　主な講師陣

データサイエンス関係　6教科

東京医科歯科大学　名誉教授
東北大学　東北メディカルメガバンク機構
田中　博　機構長特別補佐

東北大学
ゲノム遺伝統計学分野
理化学研究所
革新知能統合研究センター兼任
田宮　元　教授

理化学研究所
革新知能統合研究センター
杉山　将 センター長

産業技術総合研究所
人工知能センター
辻井 潤一　センター長

【2018年度から】
慶応義塾大学
先端生命科学研究所 所長
冨田　勝　教授

東京医科歯科大学　医療データ科学推進室　長谷武史　特任准教授

順天堂大学　茂櫛薫　講師

東京大学先端科学技術研究センター　辻真吾　特任助教

人材育成　主な講師陣

ビッグデータ医療　4教科

東北大学
東北メディカル・メガバンク機構
山本雅之　機構長

国立国際医療研究センター 研究所
遺伝子診断治療開発研究部
メディカルゲノムセンター
加藤　規弘　センター長

国立精神・神経医療研究センター
水澤 英洋　理事長

公益財団法人がん研究会
がん研究所
野田哲生　所長

公益財団法人がん研究会
がんプレシジョン医療研究センター
次世代がん研究シーズ育成プロジェクト
森　誠一　プロジェクトリーダー

AI創薬　5教科

京都大学
ビッグデータ医科学分野
奥野　恭史　教授

京都大学
薬剤疫学分野
川上浩二　教授

産業技術総合研究所
創薬分子プロファイリング
研究センター
堀本 勝久　副センター長

理化学研究所　本間光貴　氏
産業技術総合研究所　広川　貴次　氏

研修プログラムの具体構成

- 研修機関・期間
  - 研修プログラムを実施可能な参画機関・連携機関
  - 期間は２カ月から３カ月、現場ＯＪＴ研修
- 研修課題　（企業受講者研修ＯＫ）

実施機関	研修名	研修期間（日時	研修場所	受け入れ人数	主な内容
東京医科歯科大学	Deep Learning	2018年2月～3月の間の2週間を予定（1週間に2日程度、3～4時間程度を予定）	東京医科歯科大学 医療データ科学推進室	2週間で2名、2～3月で合計8名を	Deep Le るAI創薬プログラムを使用した計算創薬演習
精神・神経研究センター	（調整中）	2018年1月～3月の間の1ヶ月間を予定（1週間に2日、朝～夕方）	精神・神経 セン		
国際医療研究センター	J-DREAMS（診療録直結型全国をはじめとした、電子カルテ情報を活用した医療用ビッグデータの構築・管理・運用の実際	1月～3月の間 から2ヶ月を 間に2日、朝		2名	J-DREAMSをはじめとした当センターの医療情報管理部門が扱う、様々な医療用データベースの仕組みについて学び、実際の運用・管理の一部を行う。医療情報管理部門の定例の会議に出席し、電子カルテの運用管理および、そこで扱われる診療情報の特性について学ぶ。J-DREAMSのデータを用いた簡単な解析も可能（ただし、解析結果は、J-DREAMSに帰属する）
がん研究会	（調整中）	2～3月 で、1から2ヶ月 予定	がん研究会	1名	（調整中）
慶應義塾大学先端生命科学研究所	（来年度　予定）	（来年度　予定）	（来年度　予定）	（来年度予定）	（来年度　予定）
東京理科大学	（来年度　予定）	（来年度　予定）	（来年度　予定）	（来年度予定）	（来年度　予定）
東北大学メディカルメガバンク	（来年度　予定）	（来年度　予定）	（来年度　予定）	（来年度予定）	（来年度　予定）

2017年度の研修実施につ いては調整中

キャリア形成支援

- 「キャリア形成支援室（仮称）」の創設
  - 「キャリア形成支援室（仮称）」が博士課程学生の自己理解を支援し、適切なキャリアルートを提示し、マッチングにおいて適切なアドバイスを行う
- 連携機関と博士課程学生のマッチング
  - 連携機関として参加した企業は、育成過程を修了した大学院学生を研修プログラム（インターンシップ）に採用できる。
  - 医療データサイエンティストとしてのキャリア形成に適したインターンプログラムとのマッチングを、学生、連携機関、「キャリア形成支援室（仮称）」、博士課程の指導教員で検討を行う。

製薬会社・ＩＴ企業、ベンチャー（インターンシップの実施）

- 博士課程学生がコアカリキュラムで学習した方法を用いて与えられた課題に取り組む
- 連携機関（製薬・医療ＩＴ企業）は、研修プログラム（インターンシップ）を開催・実施することは推奨されるが義務ではない

オープンイノベーション（前競争的）研究会

産官一体となってビッグデータ医療やAI創薬等、課題に応じた研究部会を設置。

アカデミア、製薬企業、IT企業が交流をもつ場を構築することによって、わが国の医療・創薬分野での国際競争力の向上と研究の世界的先端性の達成を目標とするAll Japanのコンソーシアムへと発展させる。

－全体会；1回、各研究部会は１～２か月／回（平成２９年度は各部会２回開催予定）

－課題毎に適宜部会を形成し、意見交換を実施する

部会の例）

ビッグデータ医療部会

- 医療ビッグデータ解析、電子カルテの臨床情報とゲノム情報の統合、Real World Data準拠型臨床研究、バイオバンク構築システム論、疾患データベース構成論などの研究会活動を行う
- 国際的なビッグデータ医療についての研究状況の共有とAIを用いた応用研究についてのopen innovation として情報交換を行う

キーワード
ゲノム医療、バイオバンク、モバイルヘルス

AI創薬部会

- ヒトの網羅的分子情報や化合物との相互作用情報など「生体ビッグデータ」のもつ「創薬力」をAIにより抽出し活用する。
- 多くの生体ビッグデータベースを使って生命系全体としての反応を予測する「分子プロファイル計算創薬・ＤＲ」の研究の現状を総覧。
- AI創薬の国際的な研究情報交換やOpen Innovation としてのAI創薬の共同研究。
- Deep Learningを用いて疾患の薬剤標的を探索するシステム（東京医科歯科大学）を軸に応用性の高い「AI創薬」方法論を追跡し、個々の製薬企業の「創薬力」の向上に寄与する

キーワード
計算創薬：ビッグデータ創薬/DR、AI創薬/DR

ヘルスケア、介護など課題毎に部会形成を検討

連携機関活動促進

- 連携機関の対象
  - 製薬企業
  - 医療情報系企業
  - その他、ヘルスケアのビッグデータ・AIに関心のある企業
  - ベンチャー企業（従業員50人未満、設立5年以内）
- 連携機関のメリット
  - 大学・国研の専門家より、データサイエンス人材育成プログラムの講義を受講できる
  - 東北メディカル・メガバンク機構のデータやスパコンを使用して多因子疾患の発症解析など研修プログラムを受講できる
  - 東北メディカル・メガバンク機構の場合は試料・情報分譲申請の上、入手した情報で分析した成果は、解析研究を行った企業の知的成果となる
  - 国研やがん研のデータを用いる研修プログラムを受講できる
  - 日本オミックス医療学会の「データサイエンティスト育成課程修了認定証」を受理できる
  - 育成過程を修了した大学院学生を研修プログラム（インターンシップ）に採用できる
  - ビッグデータやAIの医療・創薬応用のオープンイノベーション情報交流活動（ビッグデータ医療部会、AI創薬部会）で情報交換や研究協力を実施できる

連携機関・連絡会議

連携機関連絡会議では、受講カリキュラムやオープンイノベーション研究会プログラムのニーズやアドバイスを行うため、連携機関の代表者が、集まって意見交換を行う。

- **目的**
  - インターンシッププログラム募集内容の検討
  - 教務内容に関するニーズ収集・大学への意見のとりまとめ
  - オープンイノベーション研究会（全体、ビッグ医療部会、AI創薬部会）に関するニーズ収集（プレコンペティテブな領域における情報交換）
- **運営方針（案）**
  - 2か月毎に1回程度開催（2017年；1回、2018年；6回）
  - 受講者の有無にかかわらず、連携機関の代表者１名は連絡会議に出席していただく。
  - 連絡会議で取りまとめた意見を運営委員会において提言する。

スタートアップスケジュール

2017年12月11日（月）　受講開始

2017年12月25日（月）　キックオフシンポジウム

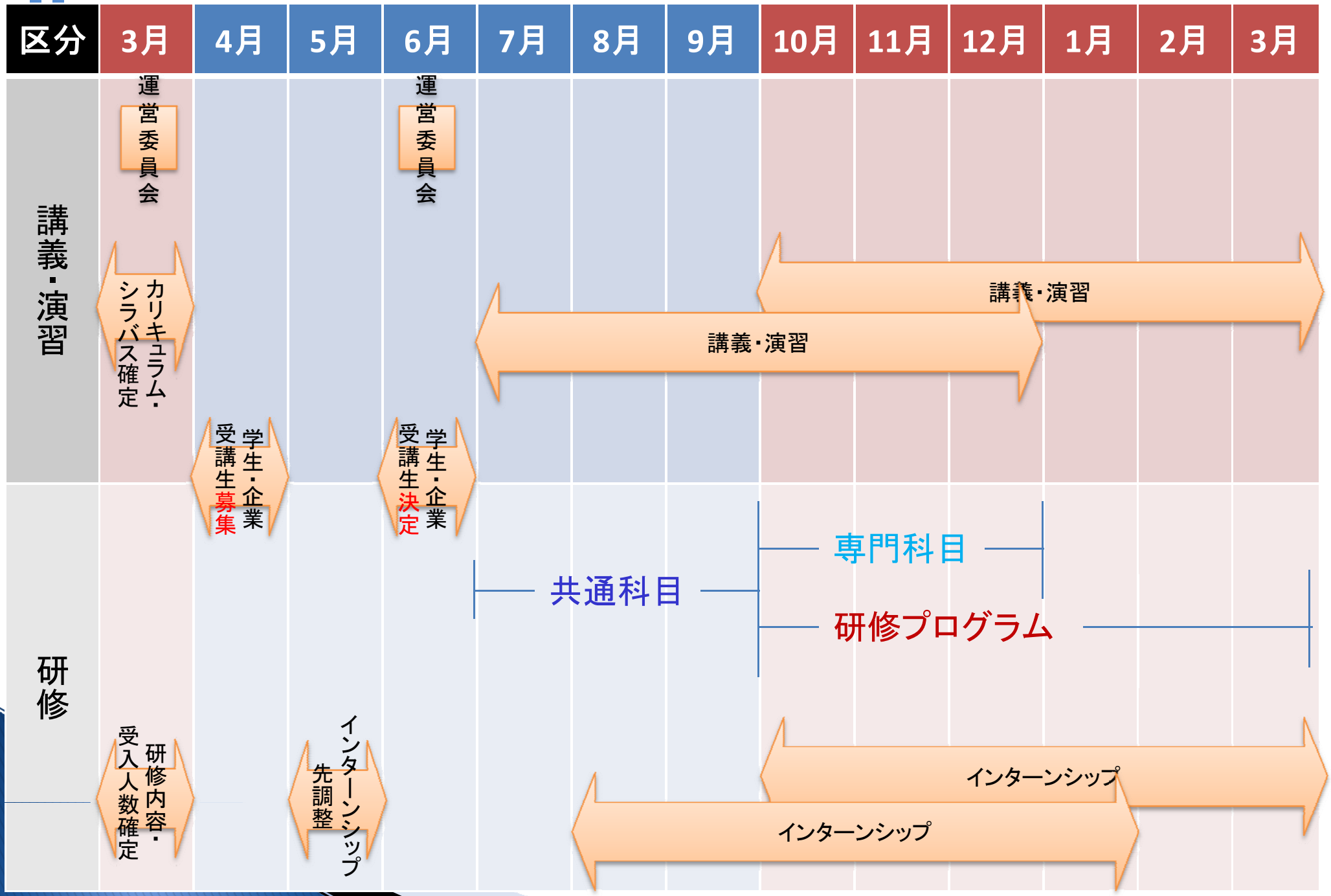
2018年1月下旬　連携機関・連絡会議

注）スケジュールが変更になる場合もありますので、最新状況は事務局にお問い合わせください。

ご清聴ありがとうございます

来年度（30年度）の計画

2018データ関連人材育成プログラム年間スケジュール（予定）



次年度の予定（会費）

企業会員（予定：2018年度以降）

	対象	会費対象	税別
A会員	大企業	受講費(1名分）+連絡 交換コンソ参加費	200万円／年
B会員	ベンチャー企業	受講費 情報交換コンソ参加	100万円／年
C会員	大企業、ベンチャー、他	連絡 交換コンソ参加費	50万円／年

調整中

※2017年度モニター機関参加者は優先的にカリキュラム、研修を選択可
※1社当たり2名目以降は会費割引を検討

注）スケジュールが変更になる場合もありますので、最新状況は事務局にお問い合わせください。

ご清聴ありがとうございます