

データサイエンスと 情報・システム研究機構の取り組み

情報・システム研究機構

北川 源四郎

2015年10月6日
トーゴの日



目次

1. ビッグデータとデータサイエンス

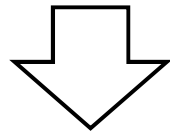
2. 情報・システム研究機構の取り組み

(1) 現在の取り組み

(2) 法人第3期の計画

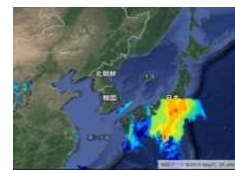
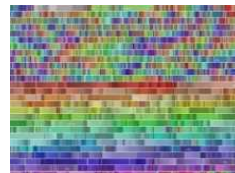
アカデミック・ビッグデータ

情報通信技術の飛躍的發展



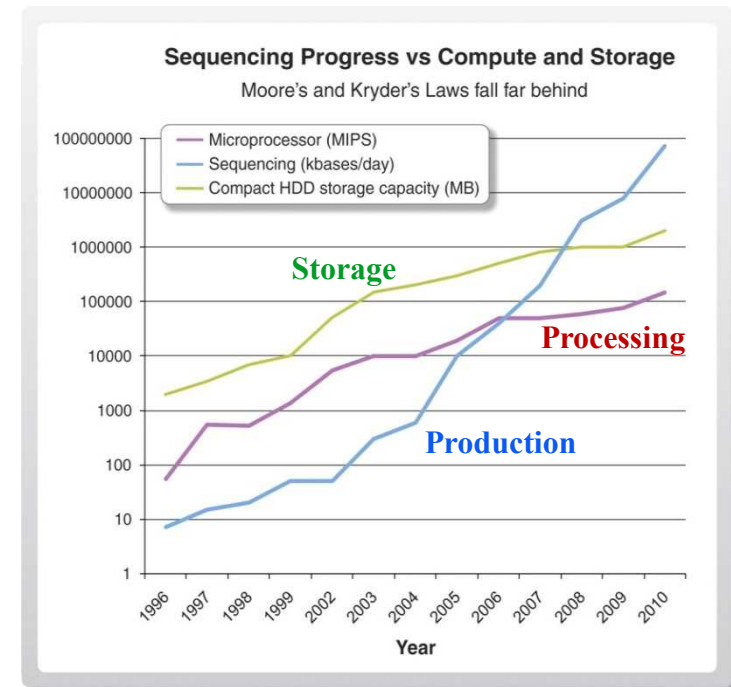
大量・大規模データの集積

- 生命科学: DNA, Micro-array data
- マーケティング: POS data
- ファイナンス: High frequency data
- 環境科学
- 防災 (地震学)
- 気象学
- 天文学 (Whole-sky CCD camera)
- 高エネルギー物理学 (LHC)
- 言語学: コーパス



情報通信技術

- 観測機器、センサー
- インターネット、通信
- データベース
- 並列計算機



(TECHTILIS Feb. 24, 2011)

ソーシャル・ビッグデータ

人間の活動を精細・網羅的に記録し、デジタル化した結果がビッグデータ

- インターネット： Web, SNS, メール, モバイルフォン
- センサーデータ： 家電, 自動車, GPS, RFID
- 取引データ： POS, 株取引, 不動産情報
- マルチメディア： 画像, 音声,
- ソフトウェア・ログ, ライフログ

活用例

- マーケティング（市場予測、顧客行動予測モデル）
- オンラインショッピングにおける推奨機能
- データ駆動型産業
- 医療・創薬・保健における個人化対応
- 社会インフラのスマート化
- センサーデータ活用（防犯, 防災, 故障検出）
- Evidence Based Policy Making, Data Journalism



データの爆発

<物質の世界>

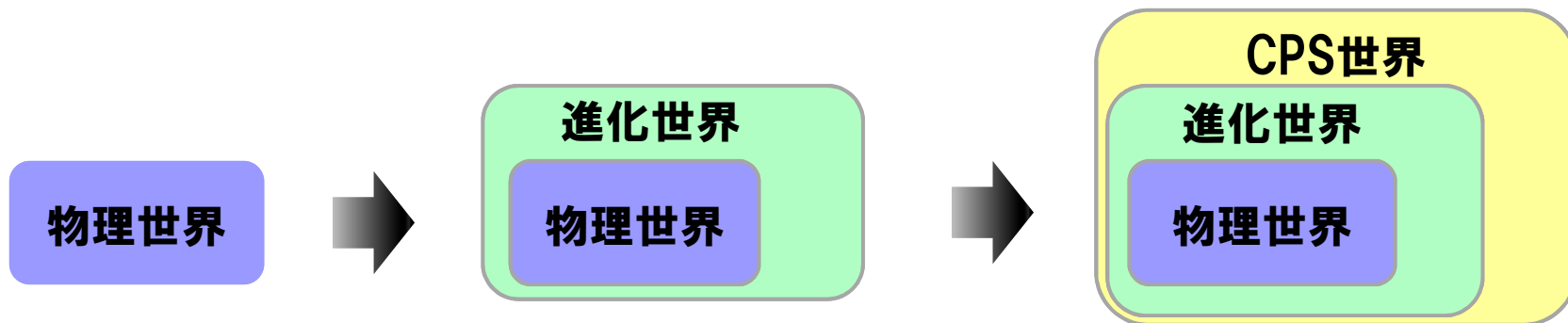
速度: 鉄道	128km/h (1932年)	→	320km/h (2005年)
旅客機	160km/h (1930年)		1000km/h (2015年)
人口: 日本	4000万人 (1900年)	→	12000万人 (2000年)
世界	10億人 (1800年)		60億人 (2000年)
GDP: 日本	70兆円 (1965年)	→	532兆円 (2015年)

<情報の世界>

ICT: ムーアの法則	2倍 (1.5年) , 100倍 (10年)		
デジタルデータ	6.2x10 ¹⁸ B (2000年)	→	4.4x10 ²¹ B (2013年)

ビッグデータの背景： 学術研究分野

- 物理モデルだけでは解決できない対象・問題が主流に
物理学 ➡ 生命科学 ➡ 人間・社会科学 ➡ CPS
- 「認識の科学」から「設計の科学」へ
 - ・ 「真理の探究」から「予測、シミュレーション、知識創造、意思決定（制御、管理）」へ
 - ・ 「物理**モデル**」から「目的達成のための**モデリング**」



ビッグデータの背景： 社会・産業分野

- 「ものづくり」から「サービス」へ
- 「大量生産・大量消費」から「個別ニーズ対応」へ
- 「(普遍的)知識の応用」から「知識と個別情報の統合」へ
- 根拠に基づく意思決定（政策決定EBP、医療 EBM)

歴史の転換点

歴史にも境界がある。

・・・数百年に一度、際立った転換が起こる。社会は数十年をかけて、次の新しい時代のために準備する。世界観を変え、価値観を変える。社会構造を変え、政治構造を変える。技術と芸術を変え、機関を変える。やがて50年後には、新しい世界が生まれる。

・・・この転換は2010年ないし20年まで続く。

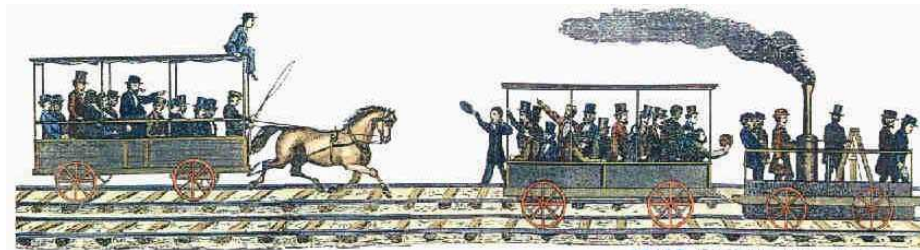
P.E. Drucker (1993) 『ポスト資本主義社会』

歴史的瞬間の再現

- 馬車 vs. 蒸気機関車

蒸気機関 => 工業化社会

1776年 ワット蒸気機関



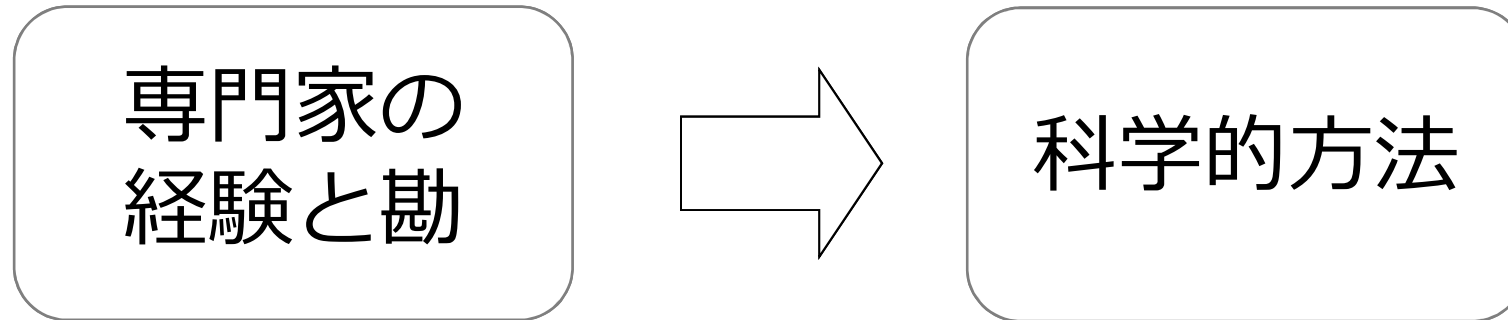
ボルチモア・オハイオ鉄道
1830年12月25日

佐野正博「技術の生存競争 ---「動力」に見る進化論」『週刊朝日百科 世界の歴史』第110号, 1991, p.695

- 棋士 vs. コンピュータ将棋

経験と勘にもとづく専門技能 vs. データ分析

社会における科学の役割



- 占星術, 航海術, 錬金術
- 工業生産過程 (ものづくり)
- 天気予報, 経済予測
- マネジメント, マーケティング
- リスク管理, ファイナンス
- 科学的発見 (発見科学)
- サービス

専門家 vs. データ分析

データ分析が専門家の経験と勘を凌駕する

- ワインのヴィンテージ評価
- 野球のスカウティング
- 人事採用
- カジノの顧客対応
- 航空会社顧客サービス
- 保険料の層別
- ネット販売の個別価格設定
- EBM 医療診断支援
- 判決予測、取引業者評価



イアン・エアーズ [著]
山形浩生 [訳]



『その数学が戦略を決める』（イアン・エアーズ著、文春文庫）

Super Crunchers, Why thinking-by-numbers is the new way to be smart

ビッグデータのインパクト

やがて、データ中心の社会が出現する。

すべての研究はデータサイエンス化する。

海外の動向（欧州）

- Datalogy : Naur (1966) Science of nature and use of data
- 探索的データ解析: Tukey (1977)
- Data Science: IFCS (国際分類学会連盟 1996)
- Data Science Journal: 国際科学会議科学技術データ委員会 (CODATA)

- e-サイエンス: Taylor (英国科学技術局長官 1999)

素粒子物理学, 地球科学, 生命科学, 社会シミュレーションなどの分野に計算機技術を適用し, 計画, 実験, データ収集, 成果の普及, 研究全過程の長期保存と活用を一体的に進めることによって, 先端科学研究を推進

- WLCG (Worldwide LHC Computing Grid)
- European Grid Infrastructure
- Open Science Grid (天文学, 素粒子物理学, 計算生物学, ゲノミクス, 分子動力学, 材料科学, 計算機科学, ナノ技術)

英国は2000年から2009年にかけて2.6億ポンド以上を投入し, Grid計算の応用, Grid Middleware開発, 14か所以上のe-Science Center支援

- FP7: Big Data Public Private Forum (2012~2014, €30億3800万)

- 欧州 : [Horizon2020](#) (FP7の後継)

2014-2020年にビッグデータ関連事業に9億ユーロ以上を投入

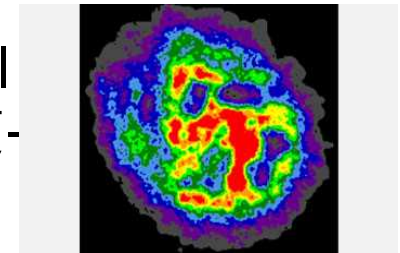
海外の動向（米国）

NSF（数学）

- 数理学の重要課題「mathematical and statistical challenges posed by large data sets」（2004年度-

情報学

- CDI (Cyber-enabled Discovery and Innovation, 2007-2011年度)
- **CPS** (Cyber-Physical Systems, 2009年度-)
- Materials Genome Initiative (2011年度-)
計算ツール, 実験ツール, 数値データを材料イノベーション基盤とし, 実験データやノウハウを蓄積したビッグデータを活用するプロジェクト
- **Big Data Research and Development Initiative**
 - NSF, NIH, DOD, DARPA, DOE, USGSを通して2億ドルの財政支援により, **ビッグデータの最新技術を構築**.
 - **スパコンとインターネット**が過去の連邦政府の投資によって飛躍的に発展したと同様に, 科学的発見, 環境・生命医学研究, 教育および国家安全保障における**ビッグデータ活用が飛躍的に進む**と指摘



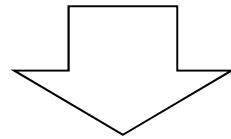
NSF

国内の動き

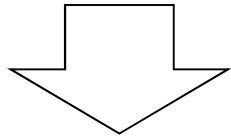
- 情報・システム研究機構設立 (2004)
- 情報爆発プロジェクト
情報爆発時代に向けた新しいIT基盤技術の研究
2006 – 2010 科研費特定領域研究
- 情報大航海プロジェクト
2007-2009 経済産業省
- さきがけ「知の創生と情報社会」
2008 –
- 平成25年度戦略目標
「分野を超えたビッグデータ利活用により新たな知識や洞察を得るための革新的な情報技術及びそれらを支える数理的手法の創出・高度化・体系化」
- AIP : Advanced Integrated Intelligence Platform Project
2016 – (予定) 人工知能/ビッグデータ/IoT/サイバーセキュリティ統合プロジェクト

帰納的推論(統計学)のトレンド

1 9 世紀：大量データからの因果律の発見



2 0 世紀：小標本に基づく精密推論



2 1 世紀：ビッグデータに基づく意思決定
(大標本への回帰? . . . しかし)

データは多ければ良い訳では…

アメリカ大統領選挙予測 (1936年大恐慌後)

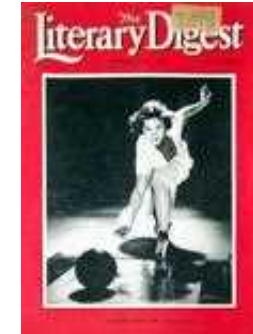
- The Literary Digest:
230万人調査 Landon 57% (370人:161人)
- The American Inst. of Public Opinion (Gallupの前身)
サンプル数 3000 Roosevelt 54%



F. Roosevelt 対



A. Landon



George H Gallup

結果：Roosevelt 60% (46/48州, 8人: 523人)

原因： 調査の偏り

- Literary Digest 読者 + 自動車保有者 + 電話利用者 1000万人
- Gallup 層化抽出 (割り当て法)

ビッグデータの活用場面

- ◆ ビッグデータを推定精度向上に使うのはCPが悪い
 - 複雑な現象の変数間の関係性の検出
 - 個性を持った対象に対する個別化対応の実現
 - レアイベントの発見

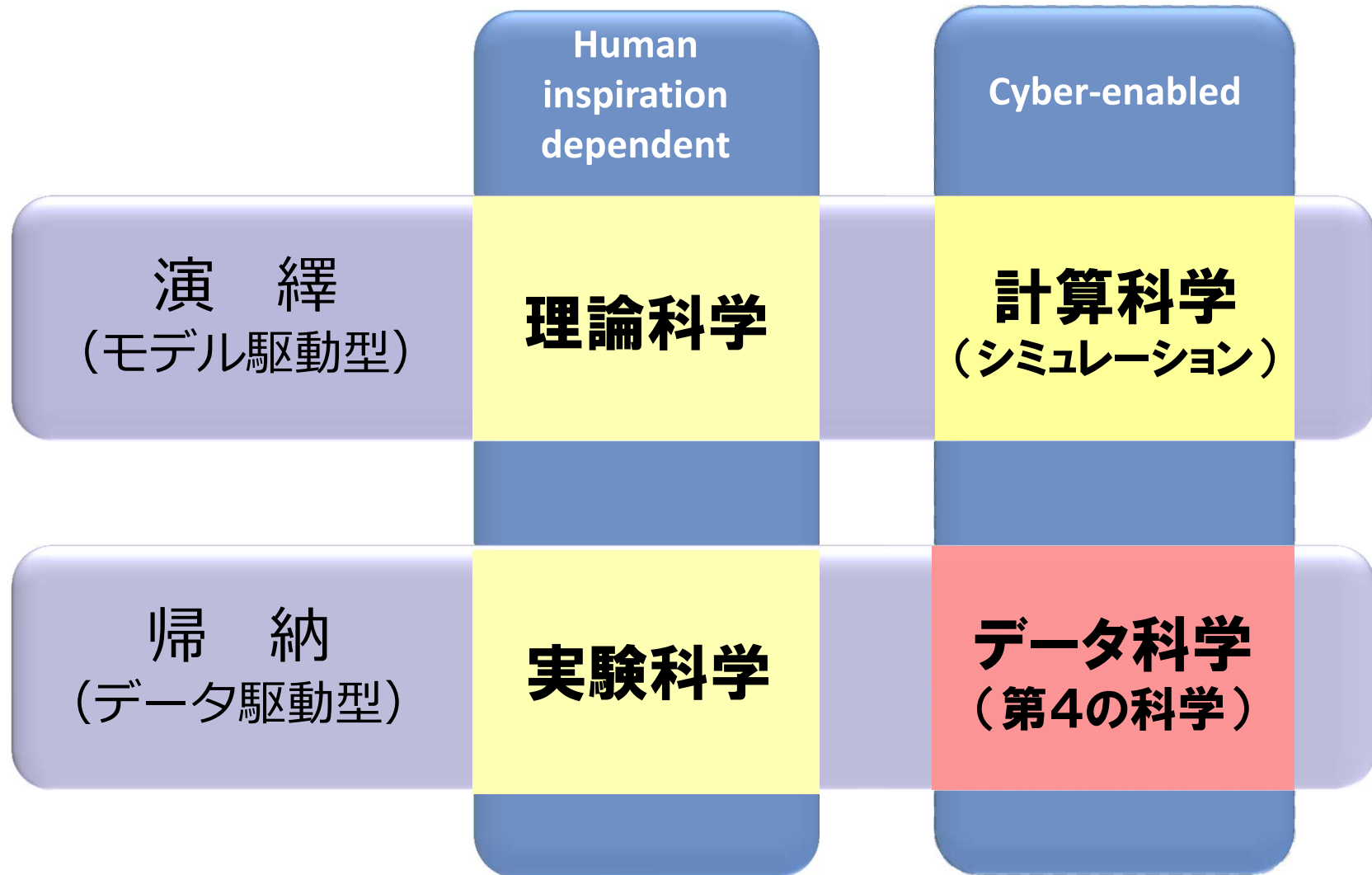
ビッグデータ活用における課題

- ビッグデータには膨大な知識や価値が埋もれている。
- 現在の方法・技術では有効活用は困難。
 - 多くは構造化されていない
 - 価値密度が低い
 - 不均一（形式，精度，観測頻度，非定常性），スパース
- ビッグデータを効果的・効率的に集約し，**革新的な科学的方法**により知識発見や価値創造を行うことが重要。

⇒ **データ中心の新たな方法論：第4の科学が必要**

- 情報統合技術
- 新NP問題

第4の科学(データ中心科学)

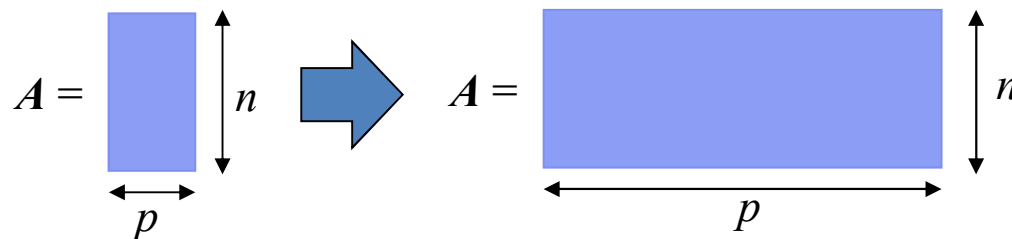


ビッグデータと次元の呪い

● 回帰モデル

$$y = Ax + \varepsilon$$

n データ数
 p 説明変数の数



マイクロアレイ、マーケティング
 $p \gg n$ 新 NP問題

● 離散的最適化

$$\max_{\theta} f(\theta) \quad \theta = (\theta_1, \dots, \theta_p)'$$

$$\theta_j \in \{1, \dots, k\} \Rightarrow k^p$$

ルービックキューブ	10^{20}
将棋	10^{150}
囲碁	10^{365}



ベイズモデリングによる異種情報の統合

1. 普遍的知識と個別的情報の統合
個別化 (**Personalization**)
2. シミュレーションモデルとデータの統合
データ同化 (**Data assimilation**)
3. 知識発展のスパイラル
能動的モデリング (**Active modeling**)
4. 事前情報とデータ情報の統合
ベイズ推論
5. 時間発展とデータ更新
フィルタリング (**Sequential filtering**)

鍵はBayes modeling

個別化：統計的ものの見方の変化

帰納：個々の具体的事実から一般的知識へ

- 集団として捉える  本質を捉える

大量生産・大量消費を目指した20世紀

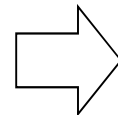
21世紀：個人に焦点をあてた科学技術へ

背景

資源の有効利用のための選択と集中

価値観の多様化（個人，個性，個別，固有の尊重）

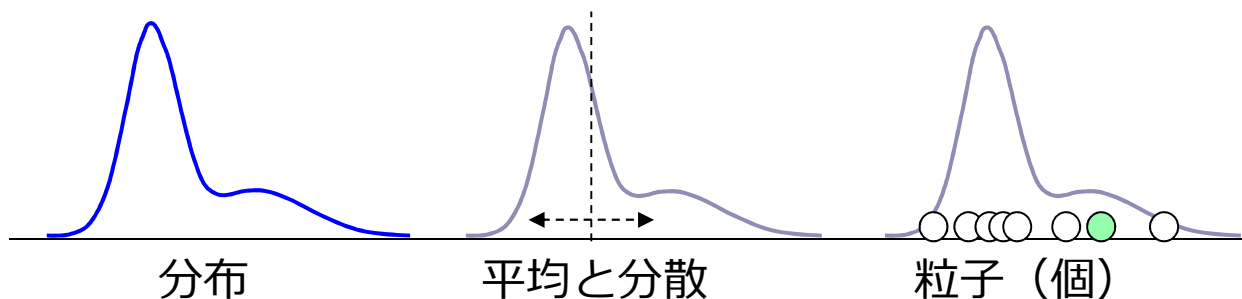
平均を見る



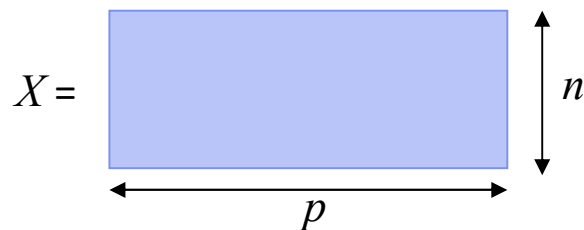
個を捉える

個別化(Personalization, 平均から個へ)

●平均から個性へ



●本質は究極の条件付け



新NP問題 ($n \ll p$)

- マイクロアレイ
- マーケティング

サービスの個別化

- 個別化医療
- オーダーメイド創薬
- マイクロ・マーケティング
- テーラーメイド教育
- ロジスティクス
- 情報提供



データ同化：モデルとデータの統合

- 気象学・海洋学の分野で開発
- 数値シミュレーションと観測データを統合

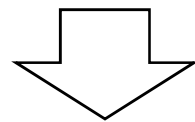
– シミュレーションモデル

複雑な現象を完全には表現できない

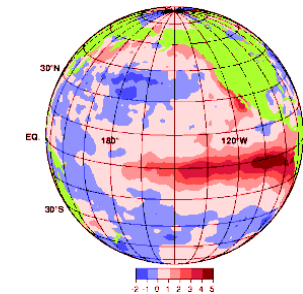
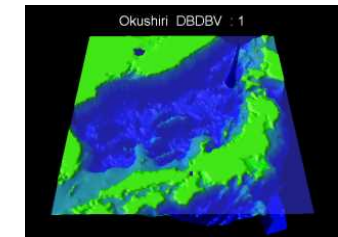
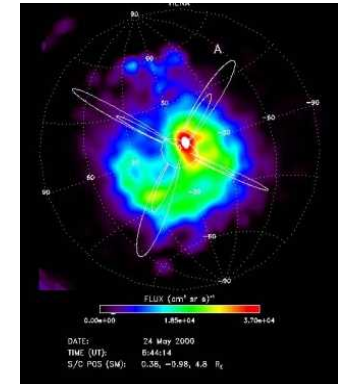
- 厳密な初期条件
- モデル不確実性（境界条件, パラメータ, 構造）

– 観測データ は有限 (スパース)

- 物理的または予算上の条件



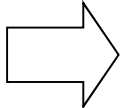
データを使ったシミュレーションモデルの改良
= データ同化

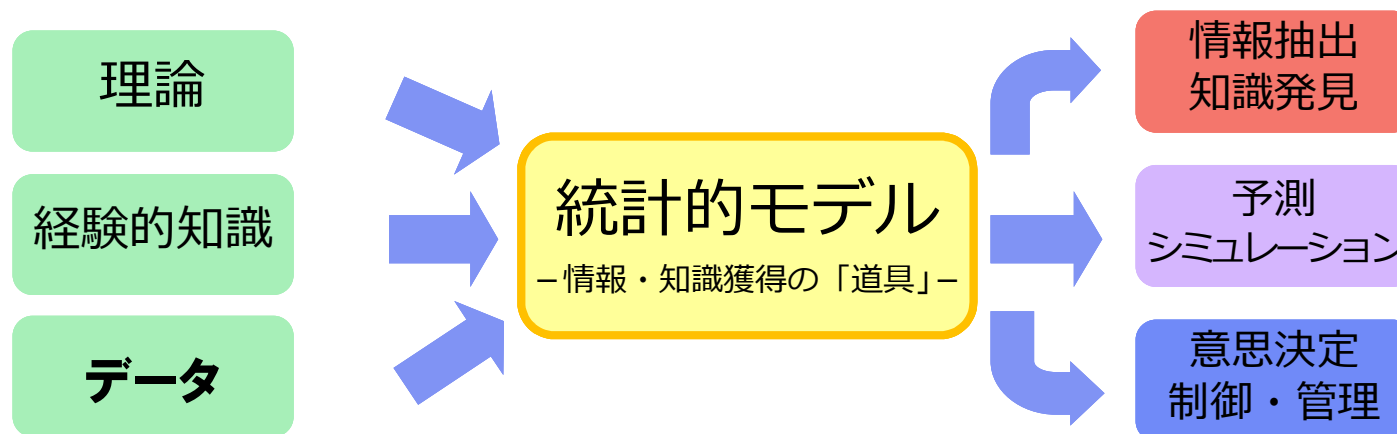


統計数理研究所
データ同化研究開発センター

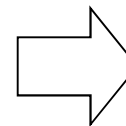
モデリングと知識発展

対象に関するあらゆる取得可能な情報 (対象に関する理論, 経験的知識, 観測データ) およびモデリングの目的

事前情報とデータの持つ情報の統合  ベイズモデリング



モデルを通して知識が得られる.
知識はモデルを改良する

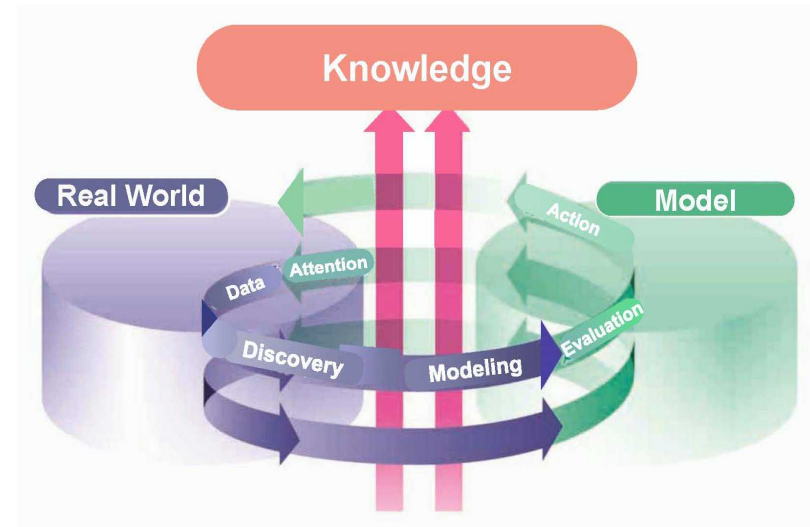


知識発展の
スパイラル

ベイズの定理と知識発展

ベイズの定理

$$p(x|y) = \frac{p(y|x)p(x)}{\int p(y|x)p(x)dx}$$



逐次フィルタリング

$$p(x_n | Y_{n-1}) = \int p(x_n | x_{n-1}) p(x_{n-1} | Y_{n-1}) dx_{n-1}$$

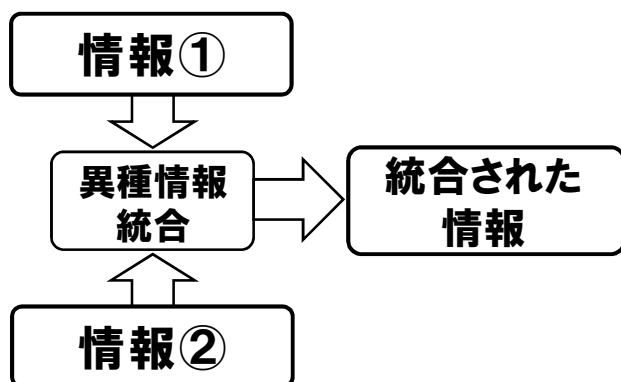
Prediction

データ

$$p(x_n | Y_n) = \frac{p(y_n | x_n) p(x_n | Y_{n-1})}{\int p(y_n | x_n) p(x_n | Y_{n-1}) dx_n}$$

Filtering

異種情報統合の諸様相



	ベイズモデル	粒子フィルタ	個別化技術	データ同化	正則化	外挿・補間
情報①	事前分布	予測分布	一般的知識	シミュレーションモデル	観測モデル	空間モデル
情報②	観測データ	観測データ	個人情報	観測データ	正則化項	近接データ
統合情報	事後分布	フィルタ分布 平滑化分布	個人化知識	データ同化シミュレーション	スパース回帰	平滑化分布

情報統合の切り札：ベイズモデリングの実用化



Thomas Bayes
(1702-1761)

ベイズの定理 18世紀

事前分布の問題
哲学的論争
計算困難性

解決

ベイズモデルの隆盛
異種情報の統合
情報抽出, 情報検索

例 : Google

近年の発展

- (1) 哲学的 : 認識の科学から設計の科学へ
- (2) 方法論上 : 事前分布, モデル評価
- (3) 計算上 : 計算機の発達, アルゴリズム (MCMC, MCF)
- (4) 応用上 : 統計数理, 情報学, 情報検索, 人工知能, 地球物理, 統計地震学, 生物学, データ同化, 季節調整, 調査データ解析

データ中心科学の要素技術

● データ解析 ビッグデータからの深い知識獲得のための方法

統計的モデリング, ベイズ推論, 機械学習, データマイニング, テキスト検索, Web情報解析, 自然言語処理, 最適化

- ・ データ同化法
- ・ インピュテーション技術 (内外挿, 不完全データ・異常値処理)
- ・ 新NP問題の解決
- ・ 高次元空間の構造探索とモデル化
- ・ 異種情報統合による個人化技術 (サービスのテーラーメイド化)
- ・ 隠れた関係の検出, 特異性の発見, 因果推論の実現

● データ可視化 膨大な高次元データや計算結果を人間が把握できるようにするための技術

次元圧縮, 特徴抽出, パターン認識や画像処理

● データ処理技術 大量の散在するデータを処理するための技術

分散処理, 並列処理, HPC, ストリーミング計算, 巨大データベース, リンケージ技術, クラウド計算などの情報処理技術

ビッグデータ解析技術

【統計】

- A/B testing
- Classification
- Cluster analysis
- Predictive modeling
- Regression
- Spatial analysis
- Statistics
- Time series analysis
- Visualization

【データマイニング】

- Association rule learning
- Data mining
- Network analysis

【機械学習】

- Ensemble learning
- Machine learning
- Natural language processing
- Neural networks
- Sentiment analysis
- Supervised learning
- Unsupervised learning

【最適化】

- Genetic algorithms
- Optimization

【計算科学ほか】

- Pattern recognition
- Signal processing
- Simulation
- Crowdsourcing
- Data fusion and data integration

情報・システム研究機構の取り組み

(1) 現在の取り組み：

データ中心科学リサーチコモンズ

(2) 法人第3期の計画：

データサイエンス共同利用基盤施設

データ中心科学リサーチコモンズ事業の推進



本事業の特徴

- ビッグデータ時代に即した第4の科学の確立
- データ基盤、モデリング・解析基盤、人材育成の三位一体となった基盤
- 生命科学、地球システム科学、人間・社会科学における融合研究の推進
- ビッグ機構の4研究所、2直轄センターが一体となって推進する事業

データ中心科学研究基盤の構成



T型、II型人材の育成

- データ中心科学・融合研究の推進に不可欠な高度な知識
- 情報付加価値処理・データ解析・モデリング・数理科学に関する幅広い知識

- 統数研: 統計思考院
- 新領域融合研究センター
- ライフサイエンス統合DBセンター



人材育成

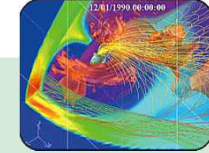


モデリング
解析基盤
整備

データ解析・
モデリングの高度化による
解析パイプライン構築

- データ構造の可視化・探索支援技術、モデリング知識基盤技術
- データ同化支援技術、シミュレーションと意思決定支援技術

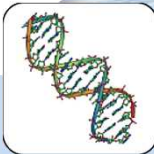
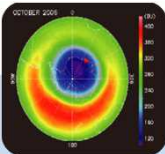
- 情報研 ■ 統数研: データ同化センター / 統計的機械学習センター
- 名大: 太陽地球環境研
- 東北大: 流体研 ■ JAMSTEC
- マックスプランク研
- ノルウェー科技大



三位一体の
基盤形成



データ基盤
整備



アカデミック・ビッグデータの統合化技術開発とデータベース化

PANSYデータ解析センターの構築

- PANSYの運用支援とデータ築盛
 - 高次解析処理と配信
 - 観測技術開拓と国際協同観測
 - 観測データ同化の推進
- 東京大学: 理学系研究科

生命科学データベースの統合化技術開発(DBCLS)

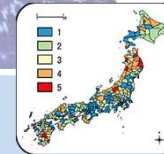
- RDF化技術によるデータベース(DB)統合の国際標準化
 - 大規模化が進む生命科学DBへの対応
 - 機構内外の文献事業との連携強化によるDBの高機能化
- 科学技術振興機構 ■ 産総研 ■ 京大化研
■ 東大 ■ 理研 ■ 韓国科学技術情報研究院

人間・社会データベースの構築

- 人間・社会データ収集・管理技術
 - データ分析・サービス合成技術
 - データプライバシー保護・匿名化・指紋技術
- 京都大学など …… 国内 25大学
■ イリノイ大学など …… 海外 10大学
■ 京都市、ソフトバンクテレコムなど …… 自治体、産業界16機関

全学術分野を対象

天文学、言語学、
健康医療・創薬
物質材料、金融・
経済 など



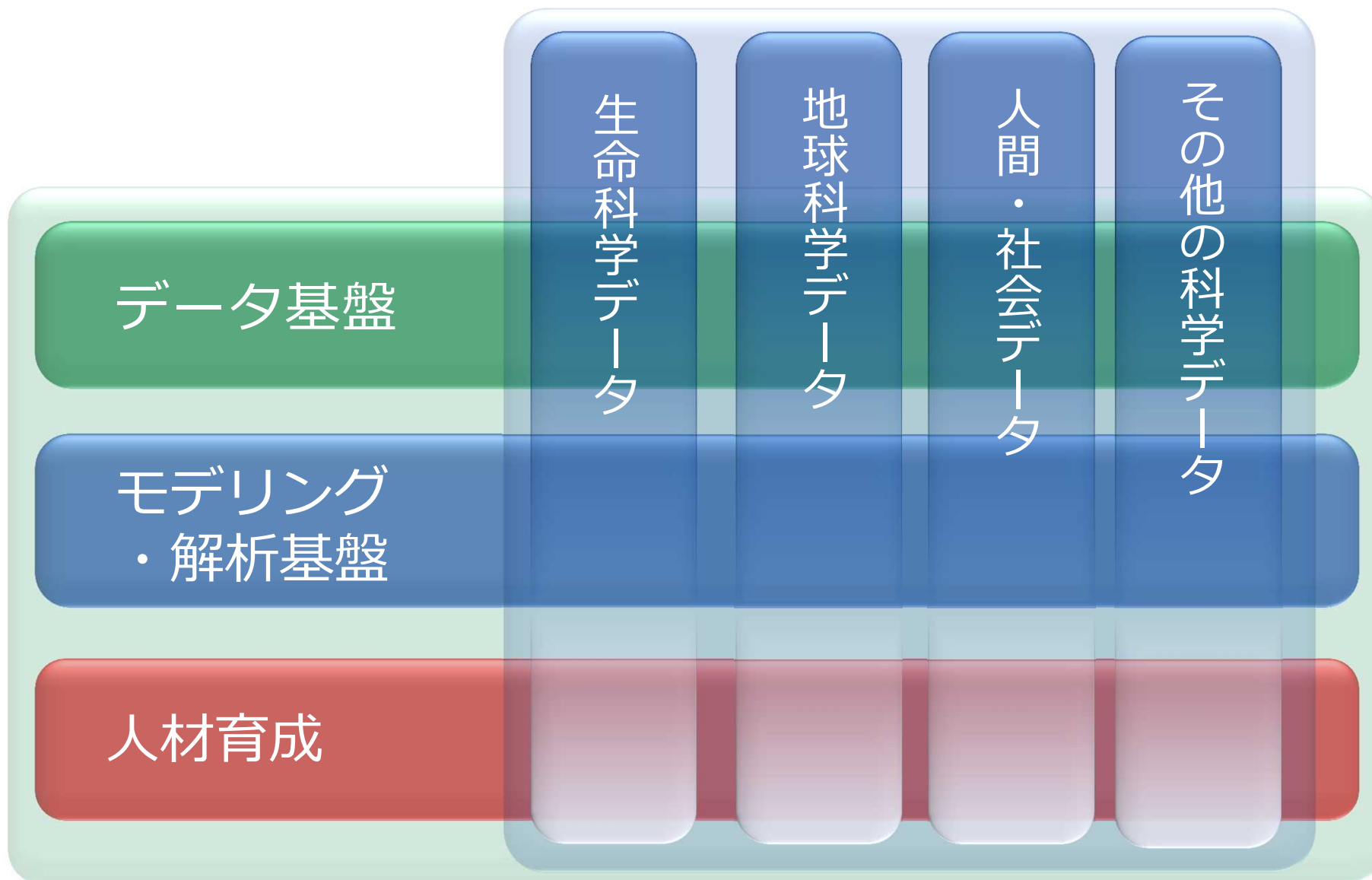
地球環境データ

生命科学データ

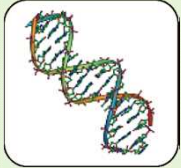
人間・社会データ

その他の領域データ

データ中心科学研究基盤

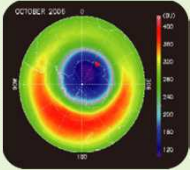


データ基盤: データベース構築



生命科学データ

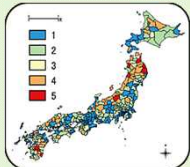
Database Center for Life Science (DBCLS)



地球科学データ

Pansy Data Analysis Center

(Program of the Antarctic Syowa MTS/IS Radar)



人間・社会科学データ

DIFI (Data Driven Information Flow Infrastructure)



他のアカデミックデータ

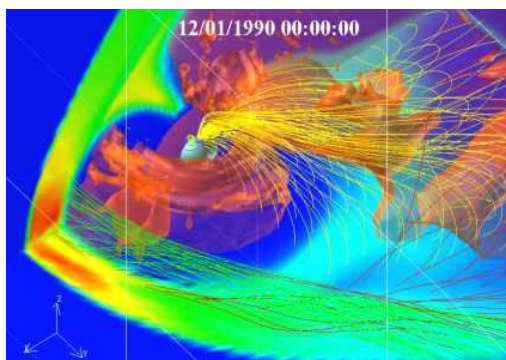
天文学, 言語学等

モデリング・データ解析基盤

データ解析の方法開発・ツール提供

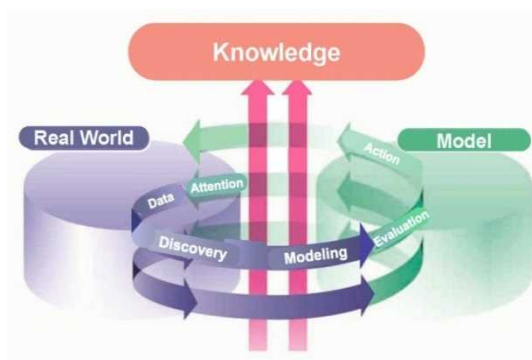
- 可視化と構造探索
- モデリングと知識発展, ベイズ推論
- 機械学習, 計算アルゴリズム
- シミュレーション・データ同化・意思決定
- e-サイエンス支援システム

Visualization of high-dimensional data and analysis results

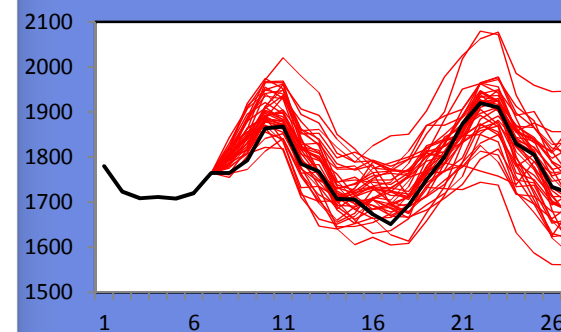


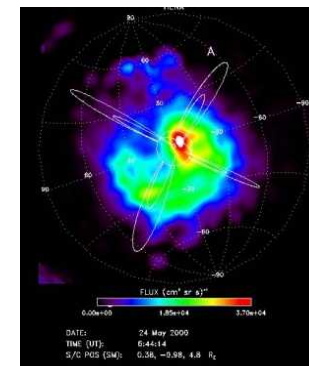
統計数理研究所
データ同化研究開発センター

Spiral of knowledge development



Ensemble simulation

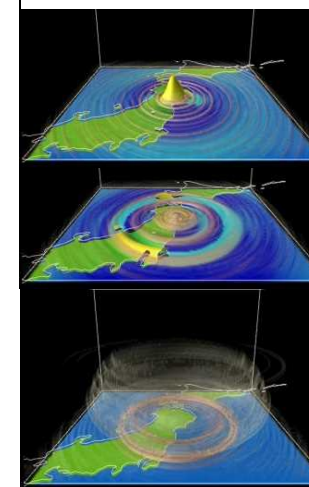




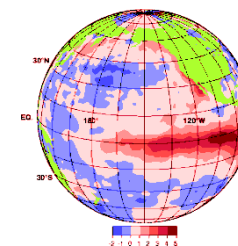
オーロラデータ同化

研究開発内容:

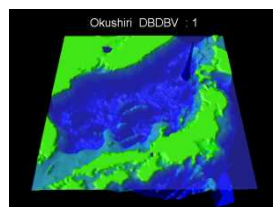
- アンサンブル・カルマンフィルタ、粒子フィルタ等のデータ同化の基本技術の研究開発
- データ同化手法の応用
 - ・ 大洋-大気系連成モデル
 - ・ 津波・微気圧データ同化 (津波予測)
 - ・ 惑星磁気圏3D構造 (宇宙天気予報)
 - ・ バイオインフォーマティクス、細胞分裂
 - ・ 感染症拡散予測・制御
 - ・ 化学プラント



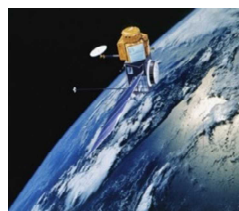
大地震に伴う大気変動のデータ同化



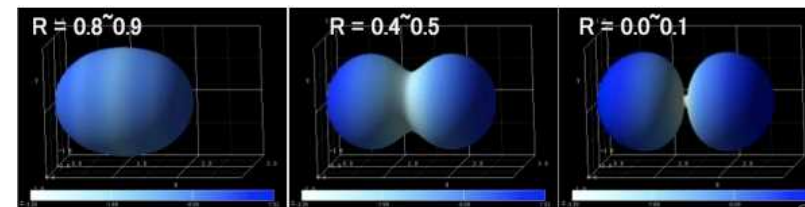
エルニーニョ予測



津波データ同化



宇宙天気予報



細胞質分裂時の細胞表層シミュレーション

ビッグデータ時代に対応する人材の育成

情報学委員会 エ-サイエンス・データ中心科学分科会（2014年9月11日）

1. はじめに
2. 海外の動向
3. 我が国の現状と人材育成に関する課題
4. ビッグデータ活用に必要な要素技術と人材育成
5. 提言
 - 提言1 データ中心科学を専門とする教育組織の設置
 - 提言2 基幹的研究組織内における恒久的なデータ解析部門の設置
 - 提言3 日本版インサイト・プログラムの早急な設置
 - 提言4 データサイエンティストの資格の制定



データサイエンティストの要件

データサイエンティストは問題の本質の把握, 定式化, データ取得, 分析, 予測, 知識獲得, 意思決定, 課題解決, 評価の全過程に関与

要素技術

- データ解析法
- データ可視化
- ビッグデータ処理技術

データリテラシー

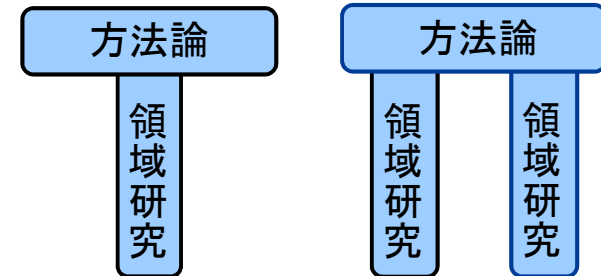
- 問題発掘、問題解決戦略立案能力
- データ収集能力, キュレーション能力(データの選択, 前処理, クレンジング)
- データの背景を見抜き, 関連するデータを見出す力
- データ分析結果の業務や事業への実装能力
- 異分野研究者・事業者との連携・コミュニケーション能力
- 研究倫理, 個人情報保護

データサイエンティストの育成方法

● データ中心科学と融合研究の推進に必要な人材

- データ解析、可視化, データ処理, モデリング、知識発展の方法
- 領域科学の知識と理解
- 課題設定, 企画立案能力
- コミュニケーション能力(異分野交流)
- 研究倫理, 個人情報保護

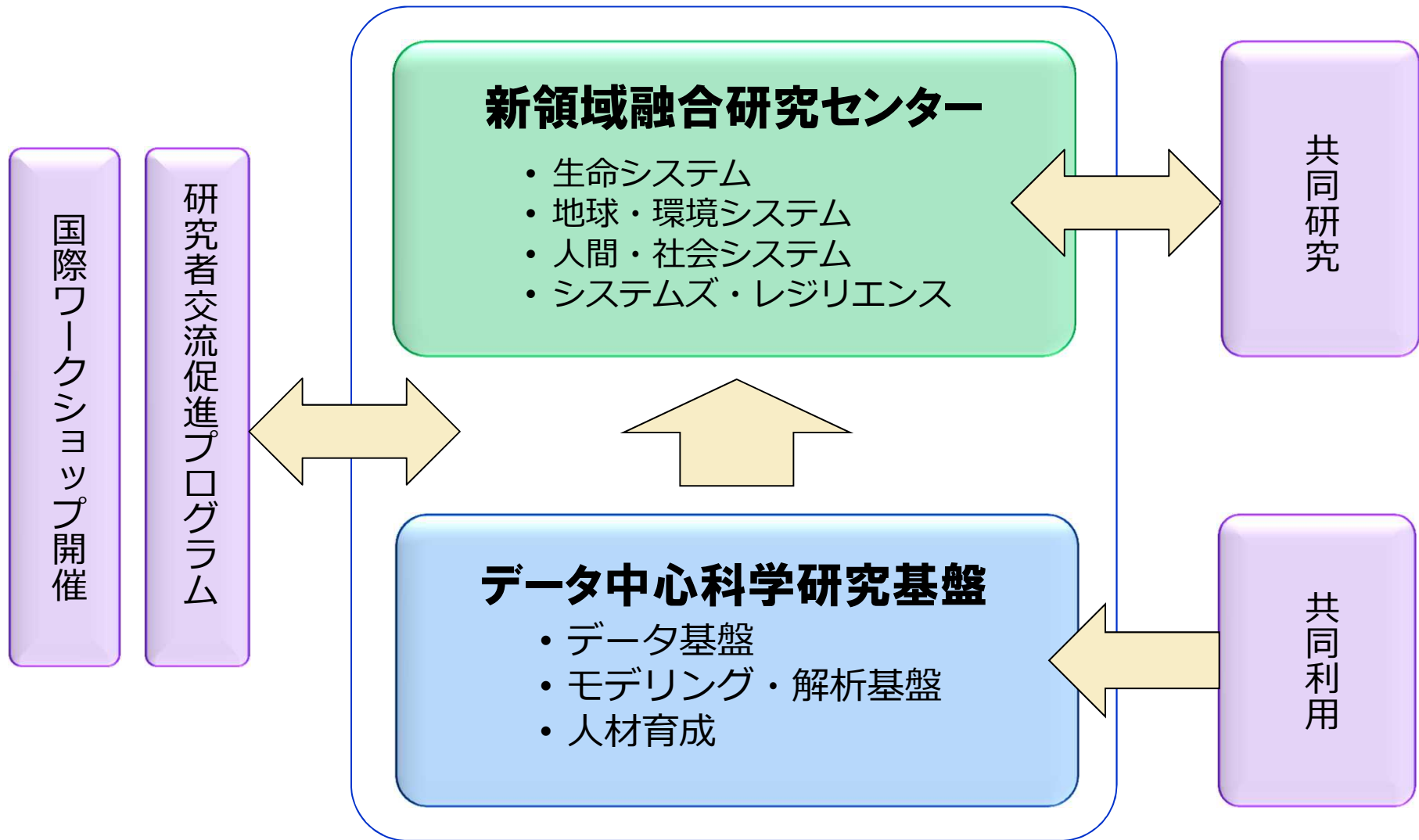
➡ **T-型, Π-型研究者**



● データサイエンティストの育成方法

- **主専攻**: データサイエンス(統計数理), **副専攻**: 領域科学
- 領域科学の博士取得者の統計・数理・情報**再教育**

データ中心科学リサーチコモンズ

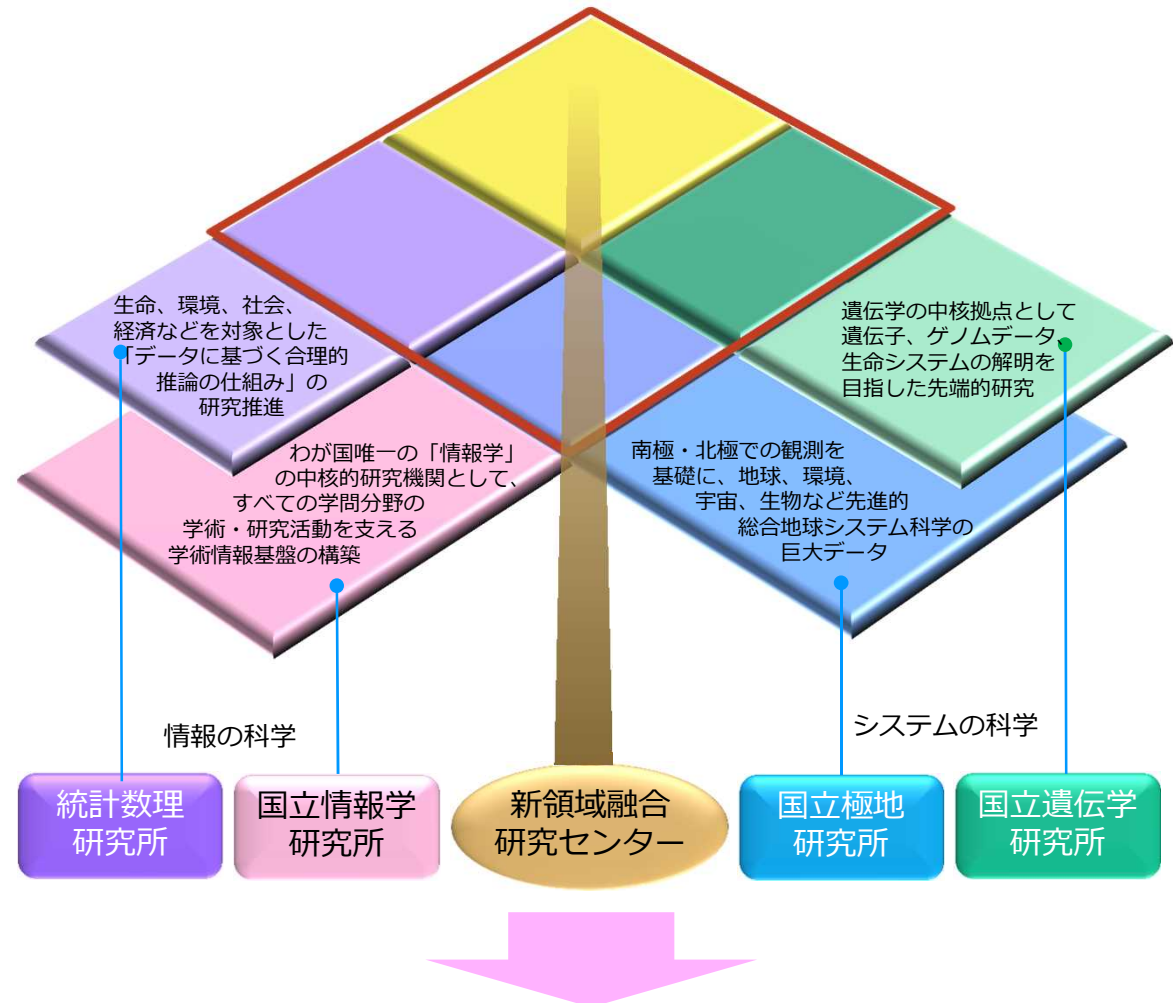


新領域融合研究センター

設置目的

極地研，遺伝研において得られる多種大量の地球科学・生命科学のデータ及び知見を**統数研**で開発されるモデリング・計算技術および**情報研**において開発される情報技術・情報基盤と結合させ，データとモデルに基づく真理の発見と予測の独創的な手法を生み出し，地球，生命，人間・社会システムの各研究分野において，**新しいパラダイムを創造**することを目指す。

運営概念図

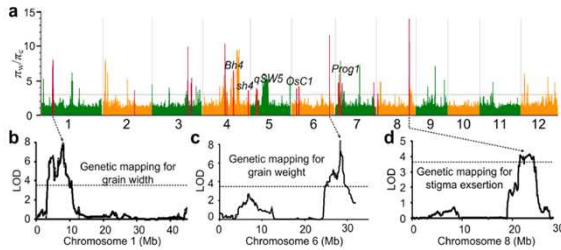


新パラダイム創成への挑戦

新領域融合研究センターのプロジェクト

生命システム学

- ゲノム関連情報の大規模生産とその情報解析手法の開発
- 遺伝的相関構造描出のための統計手法の開発と最適化
- 大量で多元的なデータの情報・統計手法を適用したゲノム機能と遺伝的ネットワーク抽出

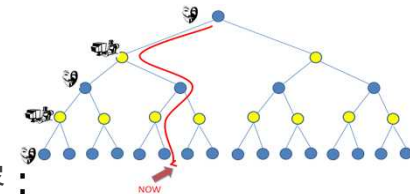


倉田のり@遺伝研

システムズ・レジリエンス

「想定外」の障害から柔軟に回復するレジリエントなシステムを設計・運用するための知識体系 (Body of Knowledge, BOK)の構築

堅牢なシステムからレジリエントなシステムへ

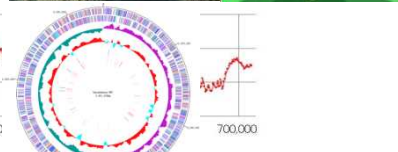
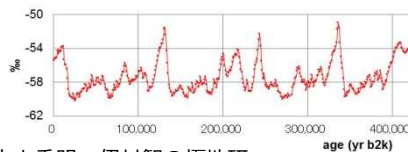
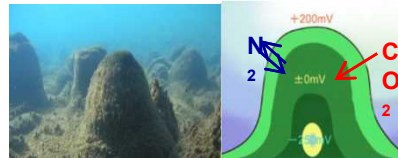


研究内容:

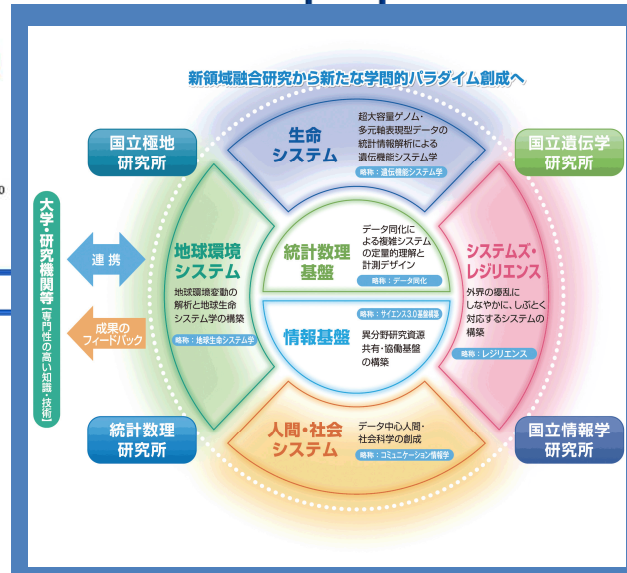
- 想定外の事象のリスクマネジメント
 - 生態系におけるレジリエンス機構
 - 社会システムにおける秩序形成
 - システムのレジリエンス性評価の計算モデル
- 丸山宏@統数研・井上克己@情報研

地球環境システム

- 地球環境変動と微生物の進化・多様性の相互作用を解明
- 環境変動下での生命の適応戦略のメカニズムを解明

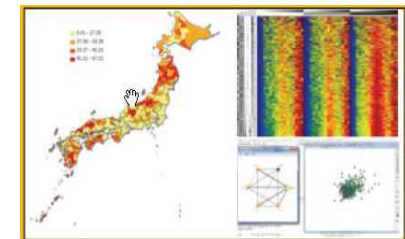
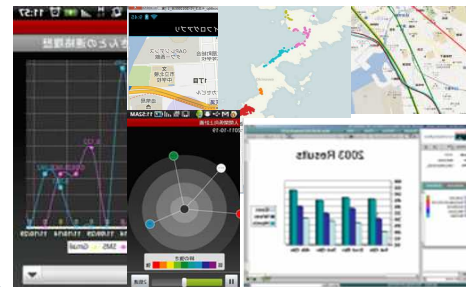


本山秀明・伊村智@極地研



データ中心人間・社会科学の創成

ネットワーク型の人間・社会データ共同利用・共同研究基盤の構築を通じて、人間・社会を適切なデータに基づいて、人やその集合体としての社会の相互情報コミュニケーションを理解・設計する新たな情報科学を創成。



曽根原登@情報研

情報・システム研究機構の取り組み

(2) 法人第3期の計画：

データサイエンス共同利用基盤施設

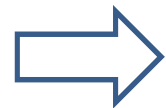
大学共同利用機関の役割りの変化

- **大学共同利用機関の本来の役割**

特定分野の研究コミュニティへの貢献

- **法人化にともなって機構が担うべき役割**

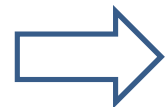
分野を超えた共同利用・共同研究、異分野融合・新分野創成の戦略的な推進



新領域融合研究センターの設置

- **法人第3期に機構が担うべき役割**

大学の機能強化への貢献



データサイエンス共同利用基盤施設の設置

法人中期目標期間と機構の取組み

法人第一期 (2004-2009)

新領域融合研究センター
およびDBCLSを設置

機構内中心の融合研究と
ライフサイエンス分野の
データ基盤構築を推進

ライフサイエンス
統合DBセンター

新領域融合
研究センター

法人第二期 (2010-2015)

データ中心科学リサーチコモンズ
事業によりデータサイエンスの三
位一体の基盤整備事業を推進

データ中心科学
リサーチコモンズ事業

ライフサイエンス
統合DBセンター

新領域融合
研究センター



法人第三期 (2016-2021)

戦略部門を機能強化し、当機構ならではの取組みを
立案・推進できる仕組みを確立

大学等への支援事業（データ共有、データ解析、人
材育成等）を中心とした活動を推進

データサイエンス共同利用基盤施設

データサイエンス支援

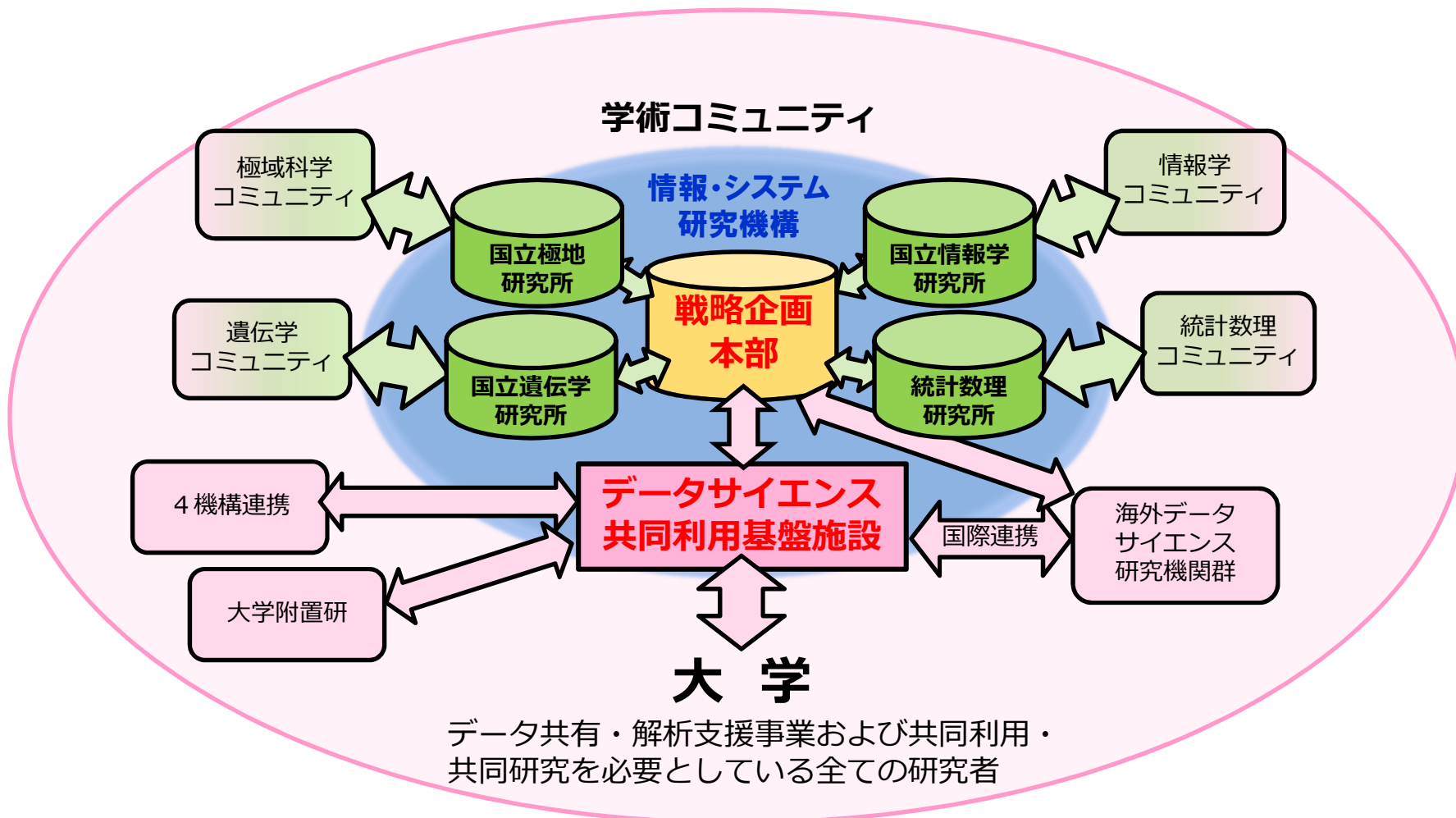
- データ共有支援
- データ解析支援
- T型・II型人材育成

戦略プログラム

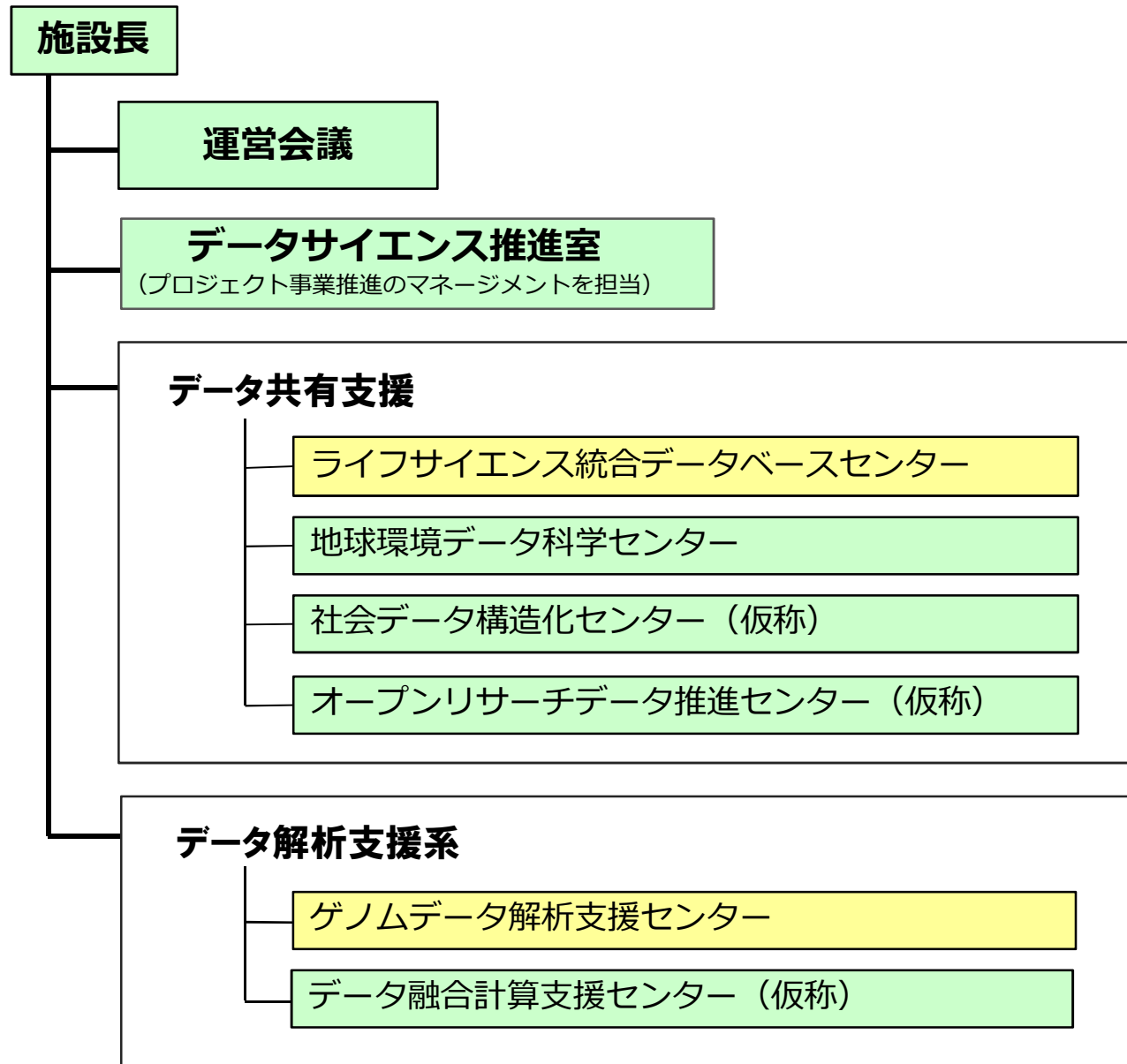
- 戦略的研究ネットワーク形成
- 未来投資型研究プログラム

データサイエンスの観点か
ら大学等の研究活動を支援

データサイエンス共同利用基盤施設



データサイエンス共同利用基盤施設



戦略プログラム

●戦略的研究ネットワーク形成

- 国際会議、MOU推進等
- データサイエンスの公募型共同利用
- 研究者交流事業
- 4 機構連携事業

●未来投資型研究プログラム

(第3期の戦略企画本部で検討)