

4Is for Innovation: Intelligence, Integration, Interdisciplinarity, and Interaction

樋口知之 (情報・システム研究機構 統計数理研究所)

異分野融合研究： 縦と横。深掘りと水平展開

研究組織（基礎研究とNOE型研究）

統計科学技術センター
(研究基盤・研究支援)

NOE型研究組織

人材育成
組織

2軸構造体制

基幹的
研究
組織

モデリング研究系

データ科学研究系

数理・推論研究系

リスク解析戦略研究センター

データ同化研究開発センター

調査科学研究センター

統計的機械学習研究センター

サービス科学研究センター

統計思考院

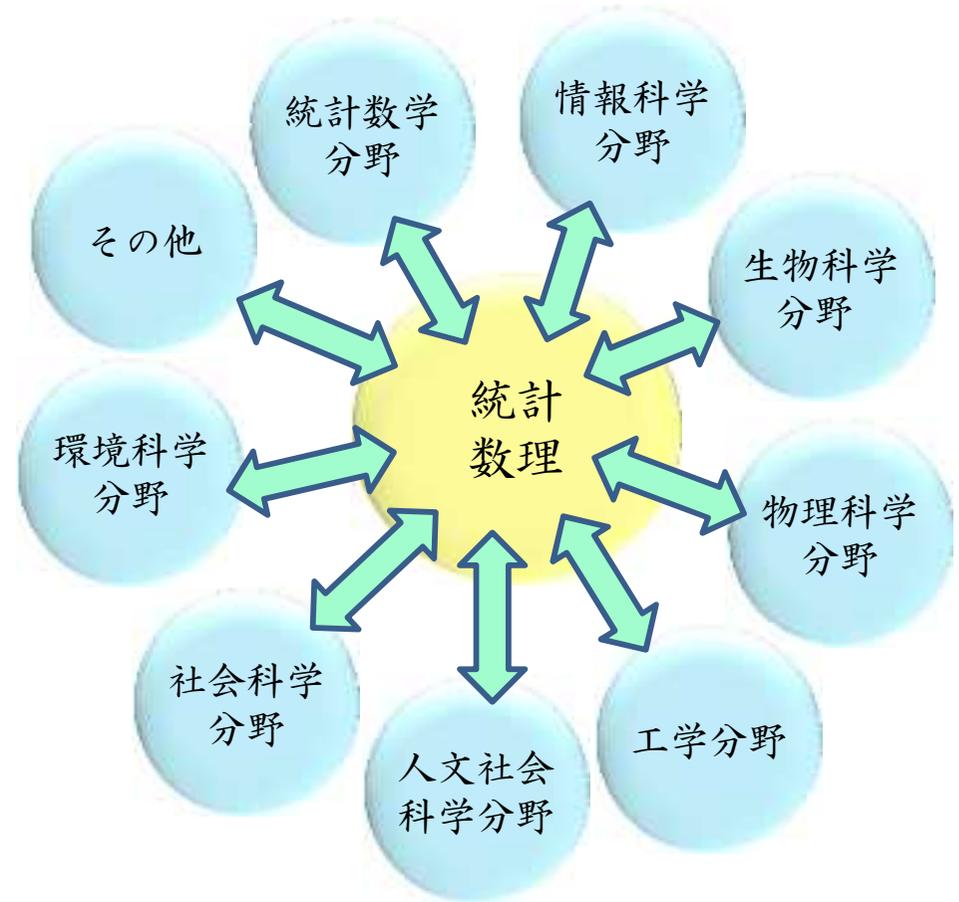
公募型共同利用・共同研究

役割

- 方法（を持った人）の提供
 - ・ 数理的方法，モデリング，データ解析法等
- 交流の場の提供
 - ・ 豊富な人的ネットワーク
 - ・ 知識交流のハブ
- 研究環境・資源の提供
 - ・ 計算機設備・ソフトウェア
 - ・ 図書・学術雑誌

共同研究の特徴

- ・ 統計数理研究所との共同研究
- ・ 広範な領域の研究者との共同研究
次ページに詳述
- ・ **萌芽的な研究、新分野発掘**の重視



科学的方法論の提供と異分野交流のハブ

公募型共同利用の分類（平成26年度実績）

横軸＝統数研の基幹的研究組織／縦軸＝研究分野

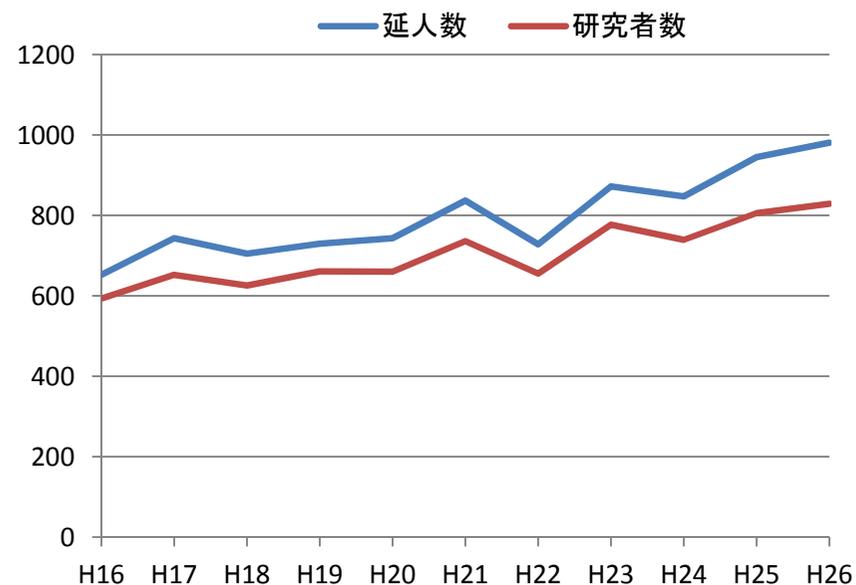
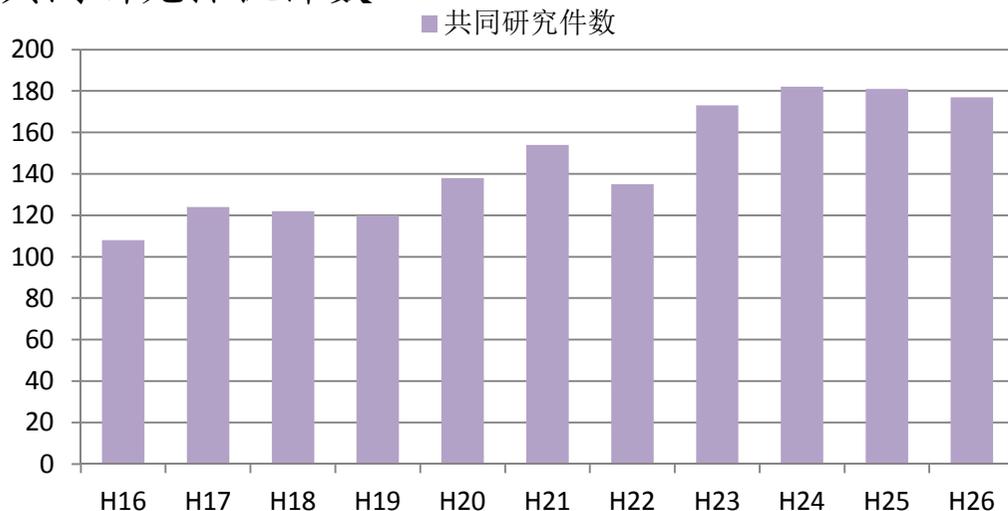
A.時空間モデリンググループ	1. 統計数学分野 (統計学の数学的理論、最適化など) 19件	2. 情報科学分野 (統計学における計算機の利用、アルゴリズムなど) 24件	3. 生物科学分野 (医学、薬学、疫学、遺伝、ゲノムなど) 34件	4. 物理学分野 (宇宙、惑星、地球、極地、物性など) 14件	5. 工学分野 (機械、電気・電子、制御、化学、建築など) 8件	6. 人文科学分野 (哲学、芸術、心理、教育、歴史、地理、文化、言語など) 14件	7. 社会科学分野 (経済、法律、政治、社会、経営、官庁統計、人口など) 38件	8. 環境科学分野 (陸域、水域など、環境データを取り扱う諸領域) 16件	9. その他 (左記以外の研究領域) 9件
B.複雑構造モデリンググループ									
C.潜在構造モデリンググループ									
D.データ設計グループ									
E.計量科学グループ									
F.構造探索グループ									
G.統計基礎数理グループ									
H.学習推論グループ									
I.計算推論グループ									
J.その他									

5/24

平成26年度実績
分類3～9 177件中 **134件** (=約75.7%)

公募型共同利用・共同研究

共同研究採択件数



年度	件数	研究者数
H16	108	593
H17	124	652
H18	122	626
H19	120	663
H20	138	660
H21	154	736
H22	135	655
H23	173	777
H24	182	740
H25	181	806
H26	177	829

平成27年度採択件数内訳 (期初)

区分	採択件数
共同利用登録	17
一般研究1	26
一般研究2	85
重点型	21
共同研究集会	17
計	166

統計思考院

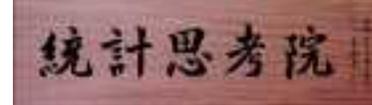
School of Statistical Thinking

- 統計数理研究所の今期中期目標・中期計画にある「統計思考力を備えたT型人材育成による融合研究の推進」を実現するため、新しい統計学の創成を目指す研究者や学生、固有分野の研究で統計学の必要性を感じた人、などさまざまな人が集い切磋琢磨しながら「統計思考」の訓練や研究をする

「統計思考“道場”」

期待される効果・目的

1. 大規模データを活用したモデリングや研究コーディネーションなど大規模データ時代に求められる**統計思考ができる人材(T型人材, モデラー, 研究コーディネータ)を共同研究の現場で育成**
2. 大学教員の**サバティカル支援事業**や**統計数理の成果の公開普及活動**の実施
3. 複雑・不確実な現象に挑戦する**新しい研究分野の創成**



共同研究スタートアップコーナー

構成員

平成27年4月1日現在

- 院長(兼) 1名
- 副院長(兼) 1名
- 教授(兼) 3名
- 准教授(兼) 1名
- 特命教授 3名
- 助教 2名
- 特任助教 4名
- 特任研究員 1名
- 外来研究員

研究支援員 等



統計思考力育成事業

統計思考院が母体となって担当しているプログラム

一般の方	公開講演会
統計数理を学びたい方	公開講座
大学院・大学院生の方	大学院連携制度
	特別共同利用研究員制度
	夏期大学院
	公募型人材育成事業
研究者の方	統計数理セミナー
研究者・その他の方	共同研究スタートアップ
他機関若手研究者の方	研究者交流促進プログラム
統計教育関連	統計教育研修
企業の方	データサイエンスリサーチ・プラザ



個別スペース

オープンスペース



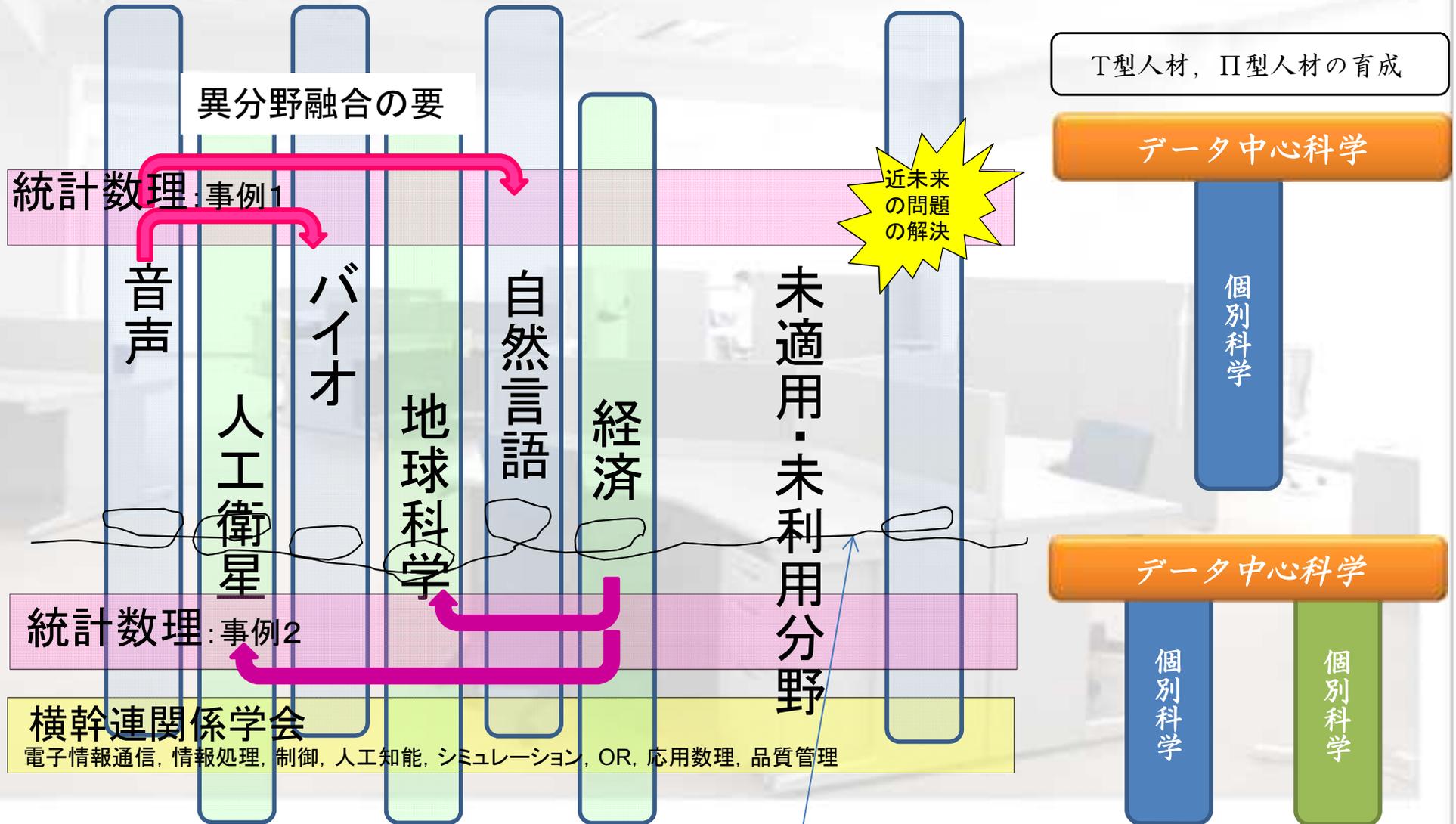
会議スペースでのTV会議等

出勤退勤札



- ◆ 統数研の教員による、統計数理の最新トピックを取り上げるセミナー（毎週水曜日に開催）
- ◆ 研究者修業の一環として、発表者からの事前ヒアリングを経て、**統計思考院特任助教**が打合せ、司会を交代で行う

統計思考院 T型人材の育成



細分化された個別科学をつなぎ融合

T型人材に必要な素養の一つの在り方



Prof. G. E. P. Box, T. Kitagawa,
R. A. Fisher and A. Kudo

統計科学のための電子図書システム
Electronic Books for Statistical Analysis

わが師わが友
統計科学の三十年
北川敏男 著
共立出版株式会社

<http://ebsa.ism.ac.jp/>

「統計科学の三十年 一わが師わが友」
(共立出版、1969)北川敏男

8章 統計学と私

“私は、幼少の頃、地理、歴史というような記述的な学問が大変好きであった。一方、数学のような論理的な思惟の学問も得意な子供であったから、妙なコンビネーションである。”

統計思考力育成事業

—統計思考力の普及・教育の場の提供—

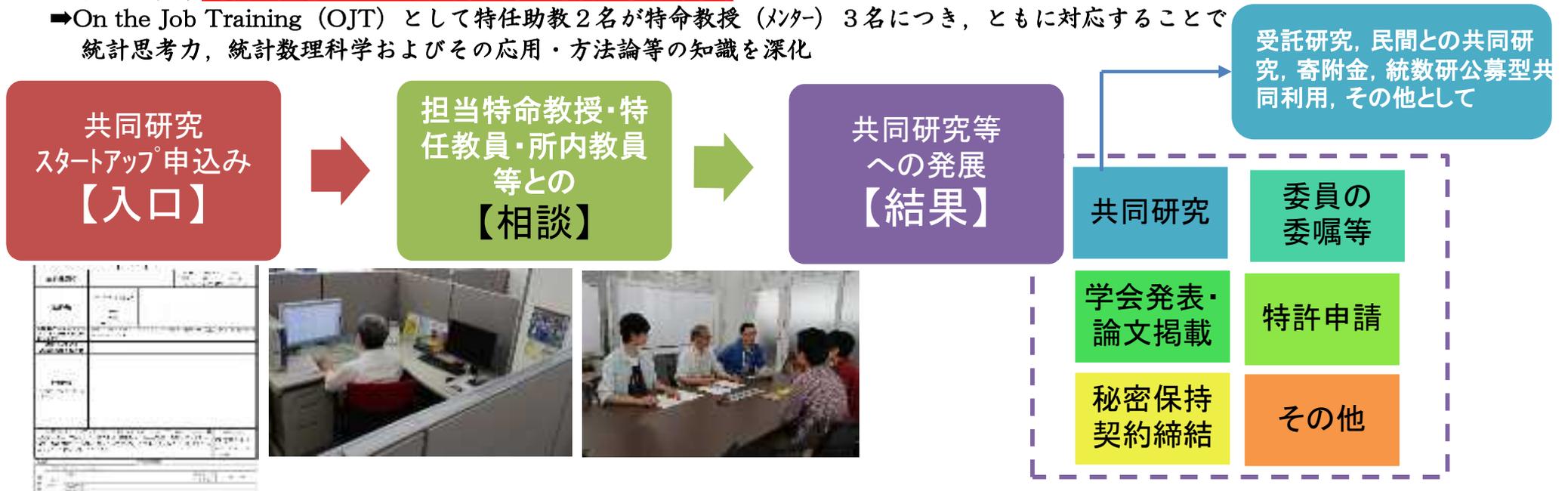
- インターネットなどの情報・通信技術や計測技術の発達に伴い、得られるデータはますます複雑化・大規模化し、データから帰納的に知識を抽出することが社会のあらゆる分野で求められており、統計的思考法が強く求められている時代
- 2011年度それまでの活動を「統計思考力育成事業」統合・拡大 統計思考力育成事業の場としての「統計思考院」を設置

統計思考力を備えた人材育成による融合研究の推進を目指して

対象	プログラム名称及び概要	
(統計数理に興味のある)一般の方	公開講演会 年1回(11月)	
統計数理を学びたい方	公開講座 年10回程度 有料 平成26年度は13講座を開催 参加人数総計は776名 平成27年度前期は6講座を開催予定	
大学生・大学院生の方	大学院連携制度 連携大学院において、統計数理に関する集中講義又は学生指導を行います。	
	特別共同利用研究員制度 他大学院学生の研究指導を行います。平成26年度は博士後期課程の学生3名を受入れ	
	夏期大学院 公開の講義を行います。年1回(夏) 平成26年度は10日間休みなしのコース開催で80名が受講	
統計数理を必要とする研究者の方	公募型人材育成事業 「統計思考力」育成のため、特に人材育成に関係する研究集会等を公募します。 * 11月初旬に公募開始予定。公募対象は研究者の方です。平成27年度はワークショップ4件が採択	
	統計数理セミナー 統数研の教員による、統計数理の最新トピックのセミナーです。 毎週水曜日 予約不要	
問題解決でお悩みの方	共同研究スタートアップ 統計数理に関わる問題について、統計専門家がその難易を見極め、解決に向けて助言を行います。 平成26年度の新規共同研究スタートアップ受付件数は41件	
他機関若手研究者の方	研究者交流促進プログラム サバティカル制度等を利用して統数研で研究をする大学教員等に対する支援制度。* 情報・システム研究機構のプログラムです。	
統計教育関連	統計教員研修 理数系教員の指導力向上のための研修等を開催しています。	
統計数理教育活動協力等	日本統計学会公式認定「統計検定」協力、 全国統計教育研究大会共催、 日本品質管理学会「科学技術教育フォーラム」共催 他	
企業の方	データサイエンス リサーチ・プラザ データサイエンス部門をスタートさせたい企業の方のために受託研究員制度(有料)を通して研究環境を提供しています。	

共同研究スタートアップ

- 問題の適切な位置付けや、課題解決のための相談に対応
- 研究者、その他の方々に統計数理に関わる問題、データ解析・統計分析等でお悩みの方に統計専門家はその難易を見極め、解決に向けて助言を行うプログラムであり、**相談の内容によって、共同研究へ発展する場合も有**
 - ⇒ On the Job Training (OJT) として特任助教2名が特命教授 (メンター) 3名につき、ともに対応することで統計思考力、統計数理科学およびその応用・方法論等の知識を深化



※ 平成26年度の共同研究スタートアップ件数: 41件

平成26年度 共同研究スタートアップ利用者の 成果等	統数研受託研究の受入れ	1件
	学会発表	6件
	論文掲載	1件
	マスコミ報道	1件

※ 統計思考院への直接の問い合わせ3件



OJTによる成果例: 統計思考院特任助教による
 “科学雑誌Newton「統計の威力」H25年12月号”
 “Newtonムック「統計と確率: ケーススタディ30」”への協力

平成26 (2014) 年度 共同研究スタートアップ受付タイトル一覧・相談者所属種別

➤ 所長への直接的問合せおよび協力:

17件 (H25年度) (内, **共同研究への発展: 2件**)

11件 (H26年度) (内, **共同研究への発展: 3件**)

No.	タイトル
1	自治体における健診データと医療費の関連について
2	平均値の差の検定方法および集計結果の効果的な表示方法について
3	ベースラインリスクの評価方法
4	人類学における数量化法の応用
5	輸入食品の統計解析
6	キャッシュフロー結果変数の決定要因の摸索
7	多変数Laplace分布の解析的取扱いについて
8	組織のネットワーク分析の相談
9	イールドカーブフォーキャストにおけるMCMCの適用について
10	企業における統計分析の活用・人材育成
11	情報量統計学的データ可視化について
12	中古オークション価格の予測について
13	多国間パワーバランス分析へのTIMSAC利用の可能性について
14	現代の財務会計に対する社会的要請に関する調査研究
15	通所サービスにおけるリハビリテーション介入効果の測定について
16	統計的システム解析ソフト ARdock の利用に関して
17	イネの網羅的遺伝子発現データの時系列解析について
18	統計的検定の方法について
19	時空間データのクラスタリングと予測の方法について
20	関連性の尺度について
21	動脈瘤手術の効果推定の方法について
22	マラウイ・ザラニヤマ森林保護区の保全に関する調査分析について
23	教育政策分析のための統計手法について
24	衛星観測データによる日射量の推定
25	広告効果の時系列分析について

H25年度
長野県の中小企業(ヘルスケア)
NICTのソーシャルビッグデータ研究
プロジェクトへ共同応募。採択。研究
続行中

JICA
同課題がマラウイ共和国の大統領・大臣署名
の最高レベルの政策文書に明記され、同国最
優先課題の一つとして直ちに取り組みを進め
ることが正式決定

No.	タイトル
26	天文学的モデルの識別法について
27	言語データの統計分析について
28	鋳造工程における不良要因の解析について
29	「偶然」の出来事の「珍しさ」について
30	選挙時の開票速報の仕組みについて
31	損害保険業務における統計解析法
32	湖水流入水量の統計的解析について
33	刑法犯認知件数に対する東日本大震災の影響について
34	時系列データダイナミックプレビュー用Webアプリケーションの開発と 科学データ・ソーシャルデータの融合表示
35	環境因子と疾患発症との関連性に関する研究
36	医療費・介護費の説明要因の統計的把握について
37	産業別就業者数の分析
38	需要予測について
39	価格時系列の連続性について
40	各国の生命表のモデル
41	情報量規準EICによる天文学的モデル評価について

— 相談者所属種別 —

研究機関	10
大学	10
官公庁	1
医療機関	1
民間企業	14
学協会	1
その他	4

➤ 「受託研究員制度」を利用した産学官(特に企業)への共同利用提供

ビッグデータ分析に基づく製品・サービスの差別化にはデータ分析や統計的機械学習の最先端技術が不可欠
ビッグデータ分析で最先端を目指す企業の方々に、最適な研究環境を提供

受託研究員が利用可能な統数研内資源等



⇒毎週の「統計数理セミナー」をはじめとする多くの研究集会等を通じた統数研約50名の常勤教員, 年間100名に近い国内外からの訪問研究者との交流



⇒初級～上級まで幅広いレベルの公開講座の受講

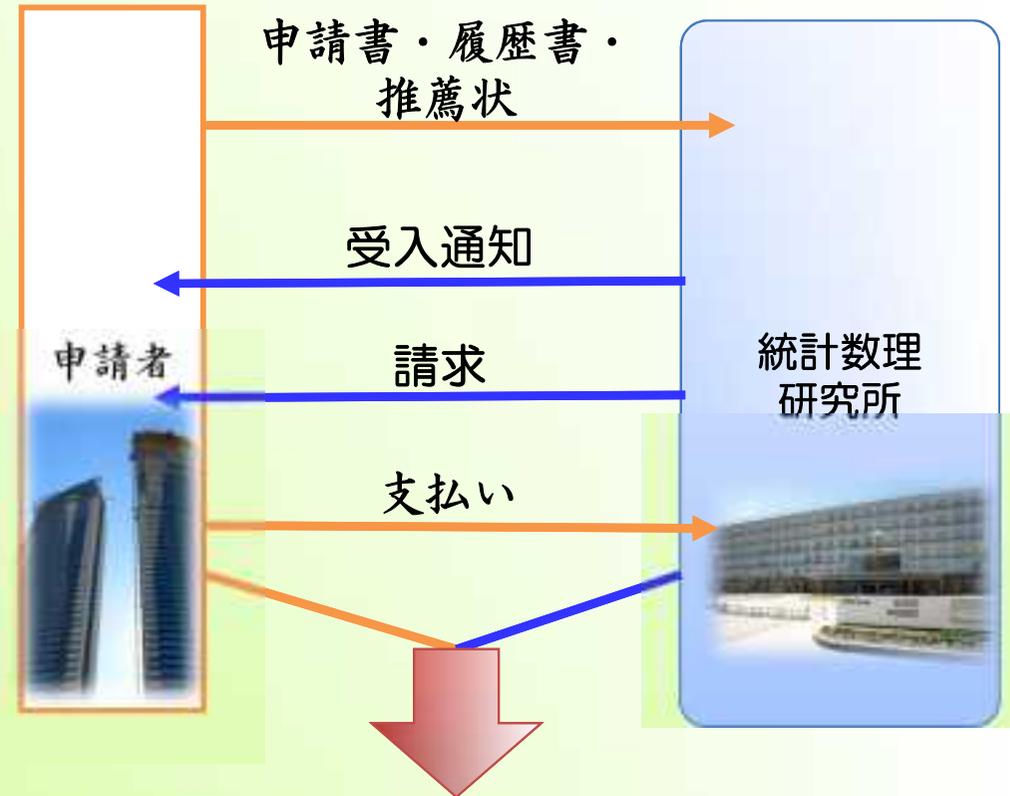


⇒統計学・数学・情報科学関係蔵書6万冊を誇るライブラリ, 電子書籍, 学術書誌情報の利活用



⇒世界最大級の共有メモリ型スーパーコンピュータシステムの利活用 *HPCIの利用登録要

受託研究員受入れフロー



受託研究員受入れ

平成27年5月より1名を受入れ

http://www.ism.ac.jp/shikoin/overview/pdf/DSRP_lo_0323.pdf

夏期大学院

- ▶ 統計数理研究所運営会議の外部委員や統計関連学会関係者からのご提案を受けて、平成18年(2007)度より全国の大学院生および学生のための夏の学校を開催しており、統計数理に係るテーマを一つ取り上げ、1~2日間の講義を開催してきたもの
- ▶ 平成26(2014)年度は、公募型人材育成事業のワークショップとのコラボレーション企画として**連続10日間**のコースとして開催(いわば、**統計数理ブートキャンプ**)

情報・システム研究機構 統計数理研究所
統計思考力育成事業 夏期大学院

感染症流行の数理モデル

夏期短期(入門)コース開講 受講生募集

非常時に現場で役立つデータサイエンティストを養成する

講師(加枠):
西浦博 (東京大学大学院医学系研究科)
Geardo Chowell (アリゾナ州立大学)
Laila Abu-Raddad (コーネル大学医学部)
Nicolas Baccar (フランス国立疫学研究所)
伊藤公人 (北海道大学人間健康学研究所)
福家寿 (東京大学大学院数理科学研究科)
若見真吾 (九州大学大学院理学研究科)
有澤三也 (統計数理研究所)
佐々木順 (総合研究大学院大学)
関井俊之 (動物衛生研究所)

日時: 平成26年8月2日(土)から11日(月)の連続10日間(休日なし)
場所: 統計数理研究所(東京都立川市緑町10-3)

登録申込みWebsite | <https://sites.google.com/vita/modelinfection>

主催: 統計数理研究所 統計思考力 東京大学大学院医学系研究科
協賛: 厚生労働省健康増進局健康政策課
協力: 文部科学省公募型事業「数学連動プログラム」



年度	テーマ
2006	時系列モデリング入門
2007	極値理論, 情報量規準と統計的モデリング
2008	赤池情報量規準と統計的モデリング
2009	実験研究および観察研究における員が推論
2010	因果のメカニズムを解きほぐす — 欠測、傾向スコア、そして統計的因果推論 —
2011	空間統計入門
2012	漸近論とその周辺
2013	情報幾何学 * 平成25(2013)年度はUstreamでも講義を配信
2014	感染症数理モデル短期入門コース

- 連続10日間(休日なし)のブートキャンプ
- 本研究分野におけるオールジャパン一流の講師陣
- 外国からの著名な講師も招へい
- 実践プログラミングまでのスキル向上を目標とするコース

【参加実績】 受講者80名 講師・チューター26名

平成27年度も同コース開催予定

研究者交流促進プログラム



- 情報・システム研究機構の支援制度を利用したサバティカル制度等を利用して統数研で研究をする大学教員等向けのプログラム
- 敷地内に併設されている“Akaike Guest House”により、遠方の研究機関からの長期滞在、内地留学も容易となっている

研究者
交流促進
プログラム

参加者募集 大学等教員
研究員募集

期待を寄せれば、拠点が変わる、成果が見える。
大学等の枠を超えた
研究者交流促進プログラムへ参加しませんか。

“Akaike Guest House” 敷地内宿泊施設

元統計数理研究所長故赤池弘次先生の業績を讃え、今後の研究活動を見守っていただきたいという願いから「Akaike Guest House」の名称がつけられた



“Akaike Guest House”
稼働率⇒約80%

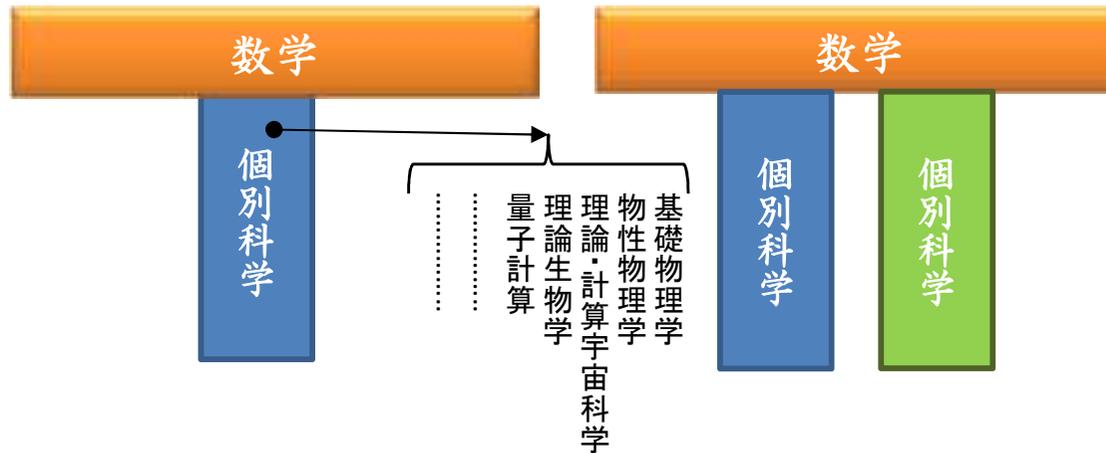


“Akaike Guest House”の稼働率の高さを鑑み、平成26(2014)年度内に増築着工。27年9月完成予定。

『数学コーディネータ』の育成: そのスキルセットとは?

数学コーディネータのスキル定義と育成

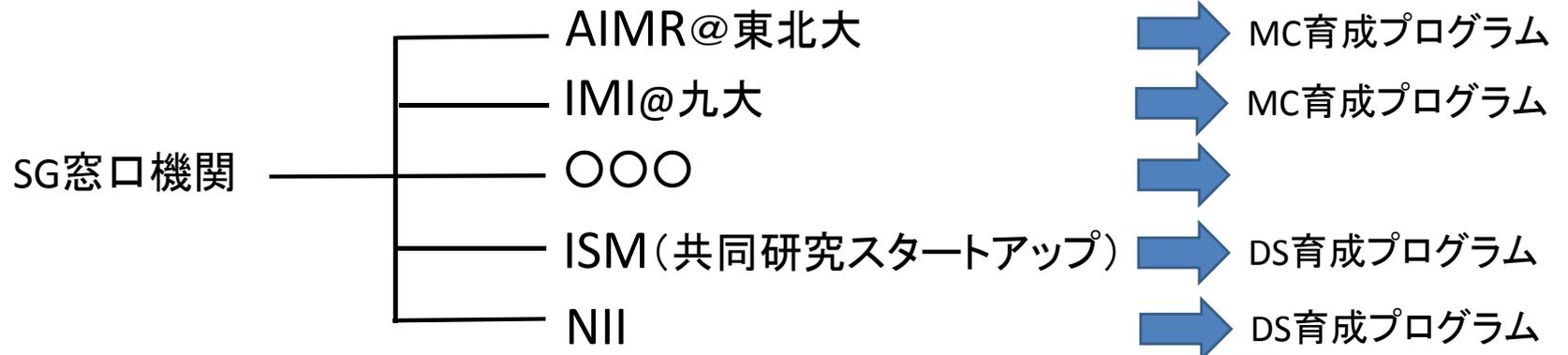
T型人材, π 型人材の育成が必須



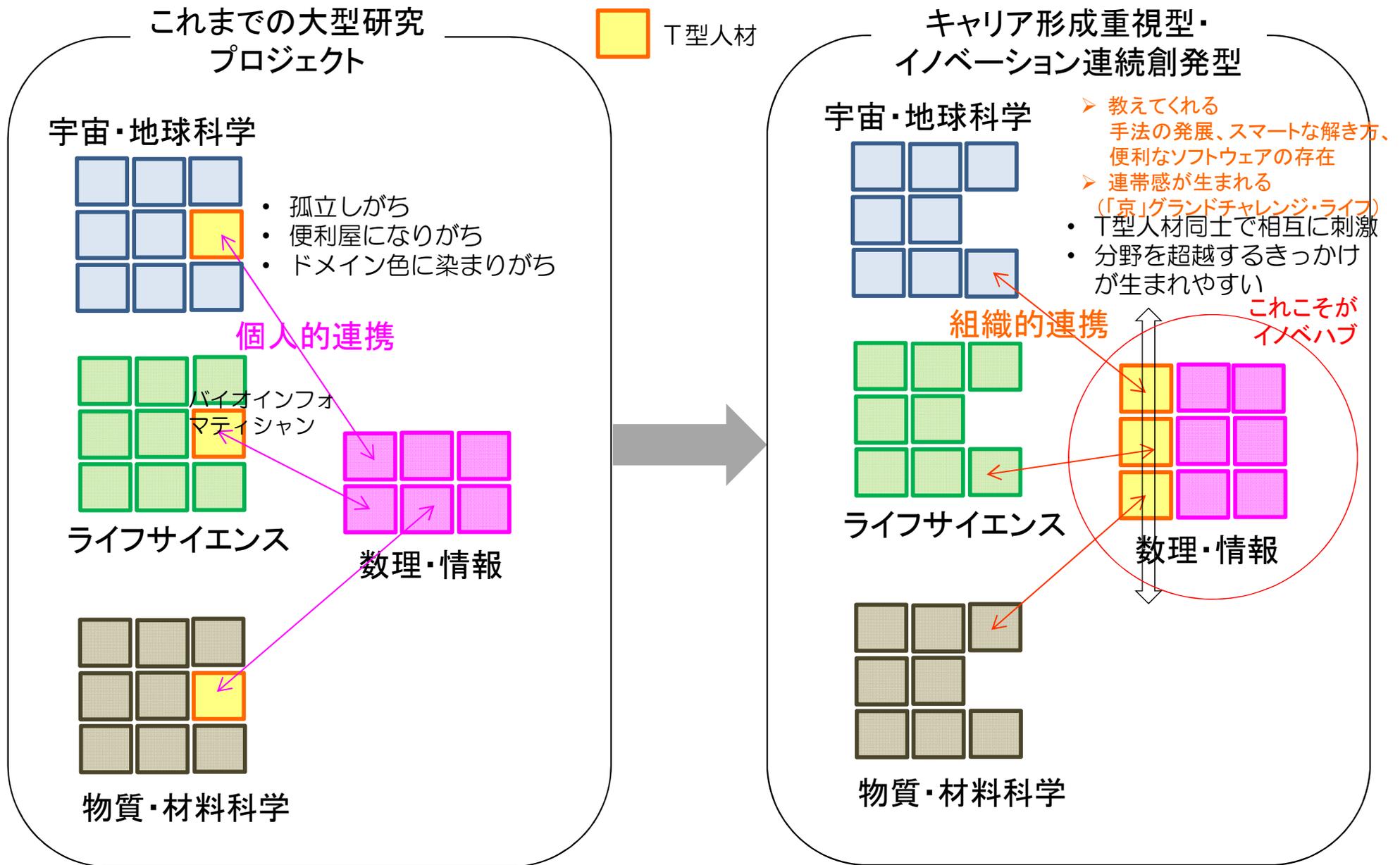
※数学コーディネータ(MC)とは、ITコーディネータ(ITC)に対する樋口(統数研)の造語

ITコーディネータとは、ITと企業経営両方の知識を持ち、経営者の経営戦略を実現するIT化支援サービスを行う専門家、もしくはその資格名のこと。ITCと略す。

スタディグループ(SG)体制の恒久化 & ワンストップサービス化



提案：これからの数学・数理科学と諸科学との協働プロジェクトの在り方

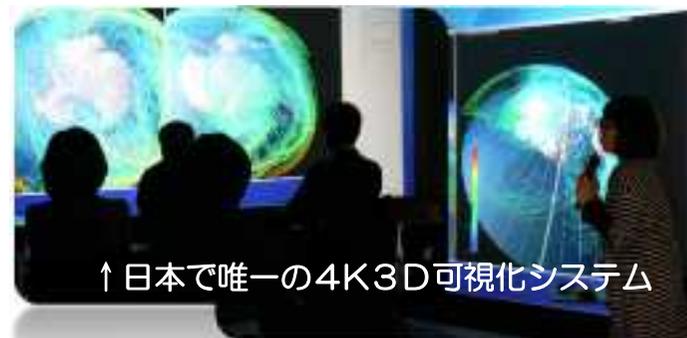


産学連携研究の実態と課題

数理＋スパコン利活用推進の視点から

統計数研の大型産学共同研究事例

- 相手先企業：企業A（業種：製造業）
 - 相手先部門：企業内研究所 研究開発部門
- 課題：
 - 「彼らが開発している機器のシミュレーションモデル構築と最適化問題及び並列計算」
 - スパコン向け並列計算プログラムの実装も目標
 - 彼らに統計数研のスパコンユーザになって頂く



↑日本で唯一の4K3D可視化システム

企業における研究部門とは(今回のケース)

- 研究部門=先行研究、R&D部門
 - 本社の設計開発部門からの問題持ち込みがトリガー
 - 自分たちで課題を設定をするわけではない。つまり、シーズ型ではない。
 - 企業全体の設計、製造計画の一部に組み込まれている
 - 共同研究する研究機関も。。
 - それを回避するために初期段階での研究者-企業側の意識の擦り合わせが非常に重要(URA)
 - 一方で、本当に企業が抱える現実の問題にあたることでそこから革新的な成果が生まれる事例も多々ある
 - 及び腰になりがちだが、それだけではよくないのではないか？
- 成果=本社設計部が試験的に使う“プロトタイプ”？
 - そうではなく、本社設計部が使う本番用プログラム
 - よって、将来的には本社設計部が統数研スパコンのエンドユーザの可能性
 - そのため、すべての要求が実装されない限り使ってもらえない！
 - スパコン利用による資金獲得には即つながらない
 - Windowsユーザをいかにスパコンユーザにするか教育的支援が必須
 - 要求のシビアさ

企業における研究部門とは(つづき 2)

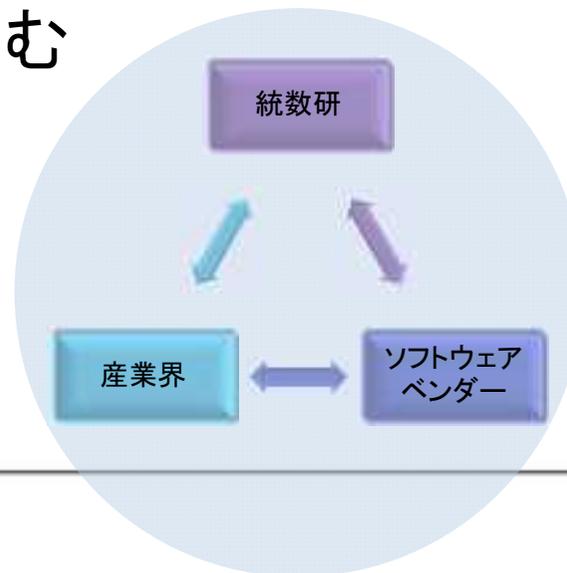
- 企業は**部門間の密な連携**で活動している
 - 研究フェーズ、開発フェーズを明確には分離不可能
 - 本社設計部門が先行**研究部門を部分的に追い越す**こともある
 - 研究開発部門だから持ち込まれる課題が革新的 ←間違った認識
 - 設計製造部門だから革新的でない ←間違った認識
 - 高等教育機関が産学連携の枠組みでどう関与していくか
 - HPC(スパコン)利活用の世界でも明確なスタンスはまだないようにみえる

見えてきた課題と解決に向けた方策

- ソフトウェアライセンスの問題 (モデリングのプラットフォームが汎用ソフトに依存)
 - 商用ライセンスの企業からの持ち込みが基本
 - 高額なスパコン向け(大規模CPU数)料金体系
 - ユーザが利用するソフトウェアのベンダーと積極的にコンタクトをとり、確認協力の関係を構築(産学連携支援URA等の役割)
 - (そもそもスパコンでは動作しない場合がほとんどだが)
 - ↑の構築が、昨年度の数学協働PのSGで非常にうまくいった他例あり
 - アイシンAW
 - 自動車用オートマチックトランスミッションのギャノイズばらつきの要因究明
 - データ分析(品質管理)+ソフトウェアベンダ
- Win-Win関係構築の困難さ
 - 理論的応用研究と計算機利用支援の両輪の形が望ましい

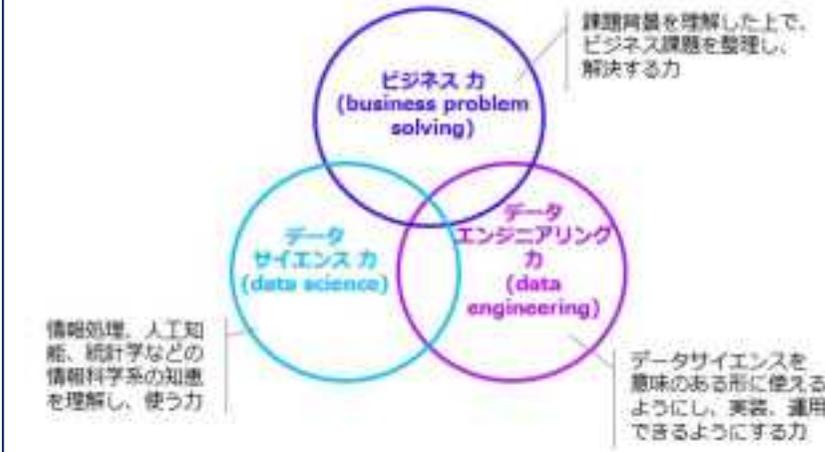
今後の展望

- 資金獲得のためには大型案件が望ましい
 - 計算機利用の観点からは実装へのコミットを放棄すべきでない
- Win-Win-Win relationshipの可能性
 - 統数研が得意とする**コア・コンピタンス**に注力したい
 - 統数研はデータ解析、モデリングの先端的知識を提供するコンサルのように振る舞う
 - **ソフトウェアベンダーとの積極的なコラボレーション**で企業の課題解決に取り組む



数理アカデミアと産業界との密な連携を阻む3つの谷

データサイエンティストに求められるスキルセット



データサイエンティスト協会が定めたスキルセット (2014年12月) 大規模に展開できない

3の谷が最初にくることも多い

3の谷
企業のニーズにあっていない
企業戦略とのすりあわせ

2の谷
計算をスケール化 (高速化
大規模に展開) できない

1の谷
抱える課題を数理的な課題として設定できない



数学の見える化

データサイエンティスト
数学コーディネータ

数学の使える化

サイエンティフィックプログラマー
計算エンジニア
(ソフトウェアベンダー)

価値の共有化

産学共同研究
担当URA